Studies in Universal Logic

Arnold Koslow
Arthur Buchsbaum
Editors

The Road to Universal Logic

Festschrift for 50th Birthday of Jean-Yves Béziau
Volume I

Birkhäuser

Birkhäuser

Studies in Universal Logic

This series is devoted to the universal approach to logic and the development of a general theory of logics. It covers topics such as global set-ups for fundamental theorems of logic and frameworks for the study of logics, in particular logical matrices, Kripke structures, combination of logics, categorical logic, abstract proof theory, consequence operators, and algebraic logic. It includes also books with historical and philosophical discussions about the nature and scope of logic. Three types of books will appear in the series: graduate textbooks, research monographs, and volumes with contributed papers.

Arnold Koslow • Arthur Buchsbaum
Editors

The Road to Universal Logic

Festschrift for 50th Birthday of Jean-Yves Béziau
Volume I

Birkhäuser

Editors
Arnold Koslow
City University of New York
New York, NY
USA

Arthur Buchsbaum
Federal University of Santa Catarina
Florianópolis, SC
Brazil

ISSN 2297-0282 ISSN 2297-0290 (electronic)
Studies in Universal Logic
ISBN 978-3-319-10192-7 ISBN 978-3-319-10193-4 (eBook)
DOI 10.1007/978-3-319-10193-4
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014953260

Mathematics Subject Classification (2010): 03B22, 03B42, 03B62, 03G30, 03A05, 01A99

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

These two volumes gather together the tributes of a distinguished group of colleagues and friends in honor of Professor Jean-Yves Beziau on his fiftieth birthday.

The articles in each of the two volumes (of which this is the first) fall, broadly speaking, into four categories:

1. those concerned with universal logic,
2. those concerned with hexagonal and other geometrical diagrams of opposition,
3. those concerned with paraconsistency, and
4. current work not directly connected to the work of Jean-Yves Beziau.

With these contributed papers, we want to record our gratitude for the intellectual and organizational work of Jean-Yves in uncovering a golden tradition of logical thought, and his constant encouragement to all of us to insure that tradition will continue and flourish. Many thanks, Jean-Yves. Our heartfelt thanks on this your fiftieth birthday.

With the possible exception of the last category, there are three subdivisions of universal logic as conceived by Jean-Yves Beziau. In order to understand this project, we can do no better than to recall the way in which universal logic was compactly described by Beziau in the preface to what is probably the defining collection on the subject,[1] and to expand upon it, briefly:

(i) [**Beyond particular Logical Systems**] "Universal logic is a general study of logical structures. The idea is to go beyond particular logical systems to clarify fundamental concepts of logic and to construct general proofs." (p. v)

(ii) [**Comparison of Logics**] "Comparison of logics is a central feature of universal logic." (p. v)

(iii) [**Abstraction and the central notion of Consequence**] "But the abstraction rise is not necessarily progressive, there are also some radical jumps into abstraction. In logic we find such jumps in the work of Paul Hertz on *Satzsysteme* (Part 1), and of Alfred Tarski on the notion of a *consequence operator* (Part 3). What is primary in these theories are not the notions of logical operators or logical constants (connectives and quantifiers), but a more fundamental notion: a relation of consequence defined on undetermined abstract objects that can be propositions of any science, but also data, acts, events." (p. vi)

[1] Beziau [2].

(iv) [**Beyond Syntax and Semantics**] "In universal logic, consequence is the central concept. But this consequence relation is neither syntactical (proof-theoretical), nor semantical (model-theoretical). We are beyond the dichotomy syntax/semantics (proof theory/model theory." (p. vi)

There are of course other themes that are characteristic of Universal Logic, but it seems evident to me that the first observation—(i) [**Beyond particular Logical Systems**]—indicates clearly that universal logic does not advocate a unique logical system that is the one correct, most expressive, accurate, and useful logical structure. Universal logic includes in its domain a host of logical structures in all their variety. But universal logic is not simply a catalogue of all advocated or imagined logical structures, all logical possibilities, as it would have all the utility of a telephone book that is useful for certain problems, but cognitively dumb.

It is the second observation—(ii) [**Comparison of Logics**]—which adds intellectual content to the project. Comparison is indeed central to universal logic, but not comparisons of a vapid kind. What is intended are comparisons that not only note the difference between logical structures, but explanations of why there are those differences in a way that reveal their different logical character. That is, the second observation suggests that not only are comparisons offered, but that there may be many different ways of ordering those logics, and one cannot take for granted that those orderings or comparisons are coherent when taken together. This kind of issue is nicely illustrated when we think of a paper now commonly referred to as "Beziau's translation paradox".[2] Simply put, two logical systems K (classical propositional logic), and $K/2$ are described. Two orderings or relations are proved to hold: that K is an extension of $K/2$ and also that there is a faithful translation of K into $K/2$. So there are two orderings. The first seems to indicate that K is clearly the stronger logic, yet the second result seems to say otherwise (that there is within $K/2$ a faithful translation of classical propositional logic). Each of the two orderings seems to measure the strength of one logic over another. According then to Beziau's concept of universal logic, comparisons are a central task, but it is also a task of universal logic to figure out what to do when the orderings seem to go in different directions.

Beziau has suggested that it is like the so-called Galilean "paradox", which notes that there are more square natural numbers than there are natural numbers, and also notes that those two collections are evenly matched. It is not that Galileo's solution is recommended for the Beziau example. That is not a possible way out, since Galileo thought that, in the case of infinite collections, the notion of "is larger than" just doesn't apply. The intended similarity, as we see it, is that in both cases there are two ways of explaining the notion of one collection having more members than another, and one logic being more powerful than another. The two ways give opposing verdicts, and the resolution of this situation, Beziau maintains, is a task that lies squarely within the province of universal logic.

We mentioned that the study of Hexagonal logics of opposition falls squarely within the province of universal logic, for they provide a good example of finite logical systems, with a specified particular implication relation between their sentences (taken pairwise). In fact there is a growing literature which considers consequence relations on finite geometrical arrays of different dimension. All belong comfortably within the project that is universal logic.

[2] Beziau [1].

We also mentioned that paraconsistent logics are included in the program. That should be obvious if one considers the various consequence relations to be found in that branch of logic. Also we need to mention the beautiful studies of Dov Gabbay in which he proposed the study of *restrictive access logics* as an alternative to paraconsistent logics that is an extension of classical logic.[3]

These restrictive access logics can be described by using a substructural consequence relation, where there is a modification of the Gentzen structural conditions on implication. It then becomes an interesting problem to see what features the logical operators have will have as a consequence.[4] The examples of paraconsistent and restrictive logics lie well within the province of present day logic.

In contrast, what is interesting and novel is that Beziau's observation's in (iv) [**Beyond Syntax and Semantics**] permits the extension of the program beyond the more traditional range of contemporary logical systems. As he stated it, not only can we have the notion of consequence for scientific propositions, and non-propositional, non-sentential objects including, data, acts, and events, but we do now add pictures (perhaps mathematical diagrams), and even the epistemic notion of states of belief for which consequence relations exist, and the possibility of logical operators acting on pictures as well as states of belief. We are concerned with consequence relations that are beyond the semantical or proof-theoretical.

The case for a consequence relation between pictures has recently been forcefully made by Jan Westerhoff. Here, compactly, is the claim:

"I will describe an implication relation between pictures. It is then possible to give precise definitions of conjunctions, disjunctions, negations, etc. of pictures. It will turn out that these logical operations are closely related to, or even identical with basic cognitive relations we naturally employ when thinking about pictures."[5]

This example with its particular consequence relation, and the pictures it relates, is an extension well beyond the usual restriction of logic to syntax and semantics. It illustrates the broad implications of Beziau's observations in (iv) and the fertility of the project of universal logic. It is not business as usual.

Finally we will briefly describe another case due to Peter Gärdenfors,[6] who developed a logic of propositions upon the basis of a theory about belief revision. His results can be recast in such a way that they also follow as a case where he defines propositions as special kinds of functions, and also defines a special relation among those functions that turns out to be a consequence relation. The result is fascinating: the conjunction of functions turns out to be the functional composition of functions, and Gärdenfors' special relation among the functions is a consequence relation provided that functional composition is commutative and idempotent.

More exactly, (1) let S be a set of states of belief of some person. (2) Let P be a set of functions from S to S (called propositions) which is closed under functional composition. (3) For any members $f_1, f_2, \dots, f_n$ and g in P, let (G) be the condition

[3] D.M. Gabbay and A. Hunter [4].

[4] Private communication from D. Gabbay, 2005.

[5] Westerhoff, J. [6]. The implication relation proposed for pictures is similar to one that Corcoran [3] proposed for propositions, as noted by Westerhoff.

[6] Gärdenfors [5].

that

$$f_1, f_2, \ldots, f_n \Rightarrow g \quad \text{if and only if} \quad gf_1 f_2 \ldots f_n = f_1 f_2 \ldots f_n$$

(the concatenation of two functions here indicates their functional composition).

In particular, for any two propositions (functions) f and g, f implies g ($f \Rightarrow g$) if and only if $gf = f$. It is easy to prove that the relation (G) is a consequence condition if and only if functional composition is commutative and idempotent. The logic of these propositions has been shown by Gärdenfors to be Intuitionistic, and his consequence relation (G) is clearly epistemic. Again, it is not logic as usual, but it is just one more case of the fruitfulness of the ideas that the project of universal logic embodies.

References

1. Beziau, J-Y.: Classical negation can be expressed by one of its halves. Log. J. IGPL **7**(2), 145–151 (1999)
2. Beziau, J-Y.: Universal Logic: An Anthology, from Paul Hertz to Dov Gabbay. Birkhäuser, Boston (2012)
3. Corcoran, J.: Information-theoretic logic. In: Martinez, C. (ed.) Truth in Perspective, pp. 113–135. Ashgate, Aldershot (1998)
4. Gabbay, D.M.: Restrictive access logics for inconsistent information. In: ECCSQARU. Lecture Notes in Computer Science, pp. 137–144. Springer, Berlin (1993)
5. Gärdenfors, P.: The dynamics of belief as a basis for logic. Br. J. Philos. Sci. 1–10 (1984)
6. Westerhoff, J.: Logical relations between pictures. J. Philos. **102**(12), 603–623 (2005)

New York, USA — Arnold Koslow
Florianópolis, Brazil — Arthur Buchsbaum

Contents

Contributors

Musa Akrami Department of Philosophy of Science, Science and Research Branch of Tehran, Islamic Azad University, Tehran, Iran

Peter Arndt Fakultät für Mathematik, Universität Regensburg, Regensburg, Germany

Arnon Avron School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel

Marcelo Esteban Coniglio Department of Philosophy and Centre for Logic, Epistemology and the History of Science (CLE), State University of Campinas (UNICAMP), Campinas, Brazil

John Corcoran Philosophy, University at Buffalo, Buffalo, NY, USA

Lorenz Demey Center for Logic and Analytic Philosophy, KU Leuven, Leuven, Belgium

Răzvan Diaconescu Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania

Martín Figallo Department of Mathematics and Institute of Mathematics (INMABB), National University of the South (UNS), Bahía Blanca, Argentina

Brian R. Gaines University of Victoria, Victoria, Canada; University of Calgary, Calgary, Canada

René Guitart IMJ-PRG Université Paris Diderot, Paris, France

Idris Samawi Hamid Philosophy, Colorado State University, Fort Collins, CO, USA

James Hawthorne Department of Philosophy, University of Oklahoma, Norman, OK, USA

Wilfrid Hodges Herons Brook, Sticklepath, Okehampton, Devon, England, UK

Laurence Horn Departments of Linguistics and Philosophy, Yale University, New Haven, USA

Jean-Louis Hudry Departamento de Filosofia, Universidade de São Paulo (USP), São Paulo, SP, Brazil

Lloyd Humberstone Department of Philosophy, Monash University, Clayton, Victoria, Australia

Dale Jacquette Institut für Philosophie, Abteilung Logik und theoretische Philosophie, Unitobler, Universität Bern, Bern 9, Switzerland

Srećko Kovač Institute of Philosophy, Zagreb, Croatia

Marcus Kracht Fakultät Linguistik und Literaturwissenschaft, Universität Bielefeld, Bielefeld, Germany

Javier Legris IIEP-BAIRES, Facultad de Ciencias Económicas, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina

Wolfgang Lenzen University of Osnabrück, Osnabrück, Germany

David Makinson Department of Philosophy, Logic and Scientific Method, London School of Economics, London, UK

Daniel Parrochia University Jean Moulin, Lyon, France

João Rasga Dep. Mathematics, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal; SQIG, Instituto de Telecomunicações, Lisboa, Portugal

Stephen Read Arché Research Centre, University of St Andrews, St Andrews, Scotland, UK

Juan Redmond Instituto de Filosofía, Universidad de Valparaíso and Conicyt, Valparaíso, Chile

Amílcar Sernadas Dep. Mathematics, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal; SQIG, Instituto de Telecomunicações, Lisboa, Portugal

Cristina Sernadas Dep. Mathematics, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal; SQIG, Instituto de Telecomunicações, Lisboa, Portugal

Hans Smessaert Research Group on Formal and Computational Linguistics, KU Leuven, Leuven, Belgium

Vladimir L. Vasyukov Institute of Philosophy, Russian Academy of Sciences, Moscow, Russia; National Research University Higher School of Economics, Moscow, Russia

Jan Woleński Sucha Beskidzka, Poland

Ibn Sīnā's Two-Partite Versus Nine-Partite Logicography

Musa Akrami

Abstract A tradition of writing and teaching logic came into existence in Islamic world on the basis of Aristotle's treatises both on logic and on topics related to logic, the most apparent manifestation of which was to represent logic in the form of a nine-partite system of logicography (according to eight treatises of Aristotle and Porphyry's *Isagoge*). Ibn Sīnā, as the most distinguished logician of the Islamic world, could combine both Aristotelian and Stoic legacy in logic with his own critical reflections on logic, first philosophy, and the relation between these two disciplines. Accordingly, he, as the most voluminous author in the field of logic, has presented both many books in the framework of Aristotle's work on logic and some different books, the most important of which is *al-Ishārāt wat-Tanbīhāt: Mantiq* (*Remarks and Admonitions: Logic*).

In this book, Ibn Sīnā presents his early project in textbooks of logic according to his own conception of logic in its definition, relation with first philosophy, metaphysical foundations, tasks, topics or subject matters, and the appropriate structure of textbooks to manifest logic as it is or as it must be. Accordingly, *Ishārāt* became the manifestation of representing logic in an important non-Aristotelian manner: a manner that has been called two-partite system of logicography, with prevalence particularly in Eastern districts of the Islamic world.

In this paper, we will speak about Ibn Sīnā's innovations and achievements in logic as well as their advantages, all relying on some points taken from the history of logic.

Mathematics Subject Classification Primary 01A30 · Secondary 03A05

Keywords Aristotelian logic · Nine-partite logic · Ibn Sīnā's innovations in logic · Relation of logic and metaphysics · Two-partite logic

1 Introduction

Aristotle's books, in particular those on logic, were translated from Greek and/or Syriac into Arabic during the great movement of translation in the Islamic world.

Six books were accepted as related directly to logic. Two other books and a book written by Porphyry, on the basis of Aristotle writings, were added to the first six books so that the number of Aristotelian books on logic came to nine. Early Muslim logicians and philosophers received and accepted these nine books as the basis of both writing and teaching logic as an important tradition or style of logicography.

Ibn Sīnā adopted this style in his early works while he was taking a critical attitude towards the content and structure of logic in received tradition. He could make changes

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_1

both in the content and in the structure of logicography on the basis of some justifications, so that a two-partite style of logicography was brought out, with some advantages, alongside the traditional style. A great number of logicians, particularly in Western and Persian districts, adopted two-partite style, though there have been other logicians up to now who have made use of either nine-partite system or a combination of both systems.

Various justifications, from metaphysical to educational ones, have been set forth for the two-partite style of logicography. We will try to give a historical report as well as both logical and metaphysical foundations of Ibn Sīnā's against orthodox Aristotelian tradition in the Islamic world.

2 Aristotle's Books on Logic

Aristotle's *Organon*, having been received such a name in Byzantine Era, consists of 6 treatises: (i) *Categories*, (ii) *On Interpretation*, (iii) *Prior Analytics*, (iv) *Posterior Analytics*, (v) *Topics*, and (vi) *Sophistical Refutations*.

While Alexander of Aphrodisias (fl. 200 AD), the Peripatetic philosopher and the most leading commentator on the Aristotle's works, had not regarded the treatises *Poetics* and *Rhetoric* among Aristotle's writings on logic, Neoplatonist philosophers such as Ammonius Hermiae (c. 440–c. 520 AD) considered them as two books concerning logic. Accordingly, the number of Aristotle's books on logic was amounted to eight and included (vii) *Poetics* and (viii) *Rhetoric*.

Muslim logicians received such a legacy in logic and logicography with 8 parts. Thus, according to Fārābī, logic has 8 parts, as Ibn Nadim, in a section concerning the parts and order of Aristotle's books on logic, speaks of 8 books (with their names) (Ibn Nadīm [7], pp. 453–454).

These treatises were translated form Greek and Syriac into Arabic by translators such as Ḥunayn Ibn Isḥāq, Isḥāq Ibn Ḥunayn, YaḥyāIbn 'Udayy, Abū Bishr Mattā Ibn Yūnus (Ibn Nadīm [7], 454–456).

The Neoplatonist philosopher Porphyry of Tyre (c. 234–c. 305), gathering some matters scattered in Aristotle works on logic (mainly from Demonstration and Dialectic), wrote an "introduction" to philosophy and logic that was called *Isagoge* (= Introduction) by him. Its Latin translation was the standard textbook on logic throughout the Middle Ages (Barnes [2], ix). This book was translated into Arabic by Ayyūb Ibn Qāsim Riqqī under the title *Isagoge fi al-Madkhal Ilā al-Kutub al-Mantiqīyyah* (Ibn Nadīm [7], 445 and 462). Muslim logicians added Porphyry's *Isagoge* to the eight Aristotle's treatises, so that the number of the books related to Aristotelian books on logic, rooted directly in Aristotle's legacy, has reached nine and included (ix) *Isagoge*. Thus nine-part/nine-partite logic, with 9 books listed above as its corpus, was accepted among Muslims as standard corpus of logic and logicography.

3 Logic in Ibn Sīnā's *Shifā* (= *Healing*)

In logic, Ibn Sīnā has made use of both Aristotle's and Stoics' legacy. Moreover, he has some contemplations, commentaries, and innovations (both in the contents and structures) of his own.

Ibn Sīnā has written about 15 treatises, the most important of which are

1. *ash-Shifā'—al-Mantiq, al-Madkhal* (*Healing: Logic, Isagoge*),
2. *ash-Shifā'—al-Mantiq, al-Qiyas* (*Healing: Logic, On the Syllogism*),
3. *ash-Shifā'—al-Mantiq, al-Burhan* (*Healing: Logic, On Demonstration*),
4. *an-Najāt—al-Mantiq* (*Deliverance: Logic*),
5. *Dāneshnāme 'Alāyī—Mantiq* (*'Alāyī Encyclopedia, Logic*),
6. *al-Ishārāt wat-Tanbīhāt—Mantiq* (*Remarks and Admonitions: Logic*), *and*
7. *Mantiq al-Mashriqīyyīn* (*the Logic of the Orientals*).

One may find in the order of the above books a development of logicography from a system based on the nine books listed above to the two-partite logicography style in which there is a structural evolution in presenting logic.

It must be noted that there are eight features in Muslims' scientology, called "eight headlines", that are important in discussing the characteristics of a science and its differences with other sciences. One of the eight headlines is "order of the sections". This feature relates to the structure of the matter presented as the corpus of a science. Accordingly, any change in a traditional structure of a science may be seen as an innovation.

In *Shifā*, as a free representation and commentary of the nine-partite logic of Aristotle, Ibn Sīnā introduces logic in some "Techniques", each having some "Articles" with some "Chapters".

- ***Technique 1***, the "*Introduction*" (*Madkhal/Isagoge*), contains 3 articles as follows:
  - ***Article 1***, with 9 chapters, concerning discussions about sciences and logic, usefulness of logic, the subject matter of logic, definition of a simple word and a composite word, the essential and the accidental, essence, the types of the universal simple word, and genus;
  - ***Article 2***, with 14 chapters, concerning relations of genuses, the natural and the rational and the logical, and common accident;
  - ***Article 3***, with 4 chapters, concerning the similarities and differences between the five universals.
- ***Technique 2***. This section, under the title "categories" (= *maqūlāt*), contains 7 articles as follows:
  - ***Article 1***, with 6 chapters, concerning the purpose of the categories, relations between the different words, accident, and both accident and substance regarding two different aspects;
  - ***Article 2***, with 5 chapters, concerning the kind and the basis of division of the universal, and the number of the categories;
  - ***Article 3***, with 4 chapters, concerning the first, second and third substances, universal and particular substances, and quantity;
  - ***Article 4***, with 5 chapters, concerning quantity in accident, the properties of the quantity, and study of correlation (= *muzāf*);
  - ***Article 5***, with 6 chapters, concerning quality and its types, passitivities;
  - ***Article 6***, with 14 chapters, concerning the types of the forth genus of quality, accidents of quality, criticisms, where (or place), when (of time), and other categories;
  - ***Article 7***, with 4 chapters, concerning opposites, criticisms on the oppositions, and contraries.

- ***Technique 3***. This technique, "On Interpretation", contains 2 articles as follows:
  - ***Article 1***, with 10 chapters, concerning the knowledge of relevance between the affairs and the conceptions, definition of the simple words and the composite words, study of noun, word, statement, definition of a proposition, definitive discourse, the First indivisible, the types of the quantified/determinate propositions, on the truth and falsity of the quantified/determinate propositions, and contradiction;
  - ***Article 2***, with 14 chapters, concerning dyadic and triadic propositions, on the validity of these relations between quantified contradictories, specific propositions, opposition between the affirmative and negative propositions.
- ***Technique 4***. This section, under the title "syllogism" (= *qiyās*), contains 9 articles as follows:
  - ***Article 1***, with 7 chapters, concerning the form of the syllogism, logic as a tool of philosophical sciences, on affirmation and negation, necessity and contingency and impossibility, contradictions between the premises, general absolute syllogism, criticisms;
  - ***Article 2***, with 14 chapters, concerning conversion of premises, conversion of the absolutes, conversion of the necessaries and the contingents, and conjunctive syllogisms and their three forms;
  - ***Article 3***, with 5 chapters, concerning complex syllogisms, and contingent universal premise and its conversion;
  - ***Article 4***, with 6 chapters, concerning possible syllogisms of the first form, complex syllogisms of the first form, possible syllogisms of the second form, complex syllogisms of the second form, and possible simple and complex syllogisms of the third form;
  - ***Article 5***, with 5 chapters, concerning conditional syllogisms, disconjunctive conditionals, simple and singular concepts in the conditionals, and the negative universal in the conditionals; and composite conditional combinations, universal and particular;
  - ***Article 6***, with 6 chapters, concerning syllogisms made by conjunctive conditional in three forms, syllogisms made by conjunctive and disconjunctive, and syllogisms made by conditional categorical in three forms;
  - ***Article 7***, with 2 chapters, concerning correlation of conjunctive conditionals, the disconjunctive conditional premises and opposition among some of them;
  - ***Article 8***, with 3 chapters, concerning the definition of exceptive syllogism and its types, and syllogism per impossible;
  - ***Article 9***, with 24 chapters, concerning the syllogism that its meaning is not complete unless being universal and affirmative, analysis of the syllogisms, situations preventing the analysis according to the form of the syllogism and the forms of the premises, induction, true premises implying true conclusion, demonstration in circle, conversion of syllogism, syllogisms made by opposing premises, petition of principle, conversion of the conclusions, on induction, and on analogy.
- ***Technique 5***. This section, under the title "demonstration" (= *burhān*), contains 4 articles as follows:
  - ***Article 1***, with 12 chapters, concerning the place of the book "*demonstration*", foundations of deductions, from the knowns to the unknowns, certain knowledge, the validity of the premises of the demonstration;

- ***Article 2***, with 10 chapters, concerning the foundations of demonstration and their universality and necessity, essential predicates, subject matters of the sciences, differences and similarities of the sciences, and relations of the premises of a demonstration;
- ***Article 3***, with 9 chapters, concerning the difference of mathematical and nonmathematical sciences with dialectic, universal affirmative demonstration, difference and similarities of the sciences in principles and subject matters;
- ***Article 4***, with 10 chapters, concerning "definition", relation of definition to demonstration, and inclusion of causes as middle terms of the demonstrations.

- ***Technique 6***. This section, under the title "dialectic" (= *jadal*), contains 7 articles as follows:
  - ***Article 1***, with 10 chapters, concerning the knowledge of dialectical syllogism and its usefulness, the reason for its name, its definition, distinguishing dialectical syllogisms, the parts of dialectical syllogism, and generally accepted premises in dialectic;
  - ***Article 2***, with 6 chapters, concerning the position of proof and falsification on the basis of the position itself or external affairs;
  - ***Article 3***, with 4 chapters, concerning genus;
  - ***Article 4***, with 3 chapters, concerning the property of dialectical syllogism and applying common positions in the property;
  - ***Article 5***, with 5 chapters, concerning the first conditions for delimitation and definition, proving the definition and falsifying the property;
  - ***Article 6***, with 1 chapter, concerning identity, otherness;
  - ***Article 7***, with 4 chapters, concerning the quests of the one who asks of syllogism and induction.

4 Logic in Ibn Sīnā's *al-Ishārāt wat-Tanbīhāt* (= *Remarks and Admonitions*)

If we agree that it is possible to distinguish two periods of Ibn Sīnā's writings, in particular on logic, we should confirm that *al-Ishārāt wat-Tanbīhāt* (= Remarks and Admonitions) is his most important book belonging to the second period, although no one can give an exact date for writing of this book. While, as we said, logic in *Shifā* had been presented according to Aristotelian tradition, *Ishārāt* manifests the author's innovations in both division of logic and its topics. Ibn Sīnā briefly presents his own views in some sections under the title "*Ishārāt*" (= Remarks), while he gives some critical points concerning the views of other thinkers in sections called "*Tanbīhāt*" (= Admonitions).

Ibn Sīnā's al-Ishārāt wat-Tanbīhāt (= Remarks and Admonitions) has four parts: Logic, Physics, Metaphysics, and Sufism.

The book on logic, being the first book in the order, consists of 10 sections called "way/rightway/method" (nahj). We write its contents according to its English translation (Ibn Sīnā [12], pp. viii–xi), accepting the translator's word for "nahj" (i.e. "method"):

- **The First Method**, concerning the Purpose of Logic, containing 16 remarks on the knowledge of the composite as requiring knowledge of single elements, the logician's need for taking into consideration universal language, conception and assent, the logician's need for knowing the principles of the explanatory phrase and proof, the expression as a sign for the concept, the predicate, the essential, the accidental, the concomitant accidental, the separable accidental, the constitutive essential, the non-constitutive concomitant, the non-concomitant accidental, the essential in another sense, that which is stated as the answer to the question "what is it", the various types of that which is stated as the answer to the question, "what is it?";

- **The Second Method**, on the Five Simple Terms, the Definition and the Description containing 8 remarks concerning that which is stated as the answer to the question "what is it?" as "genus", and that which is stated as the answer to the question "what is it?" as "species", the arrangement of genus and species, the difference, property and the common accident, the description of the five [terms], definition, description, the types of errors that occur in the identification of things by definition and description;
- **The Third Method**, on Assertive Composition containing 10 remarks concerning the types of propositions, affirmation and negation, singularity, indefiniteness and definiteness, the judgment of the indefinite proposition, the definiteness and indefiniteness of conditional propositions, the composition of conditional propositions from predicative ones, equipollence and positiveness, conditional propositions, the dispositions that accompany propositions, and that giving them specific judgments in definiteness and in other cases, the conditions of propositions;
- **The Fourth Method**, the Matters and Modes of Propositions containing 8 remarks concerning the matters of the modes of propositions, and the difference between an absolute and a necessary proposition, the mode of possibility, principles and conditions for the modes, the determination of the universal affirmative in the modes, the determination of the universal negative in the modes, the determination of the two particular propositions and the modes, the implication of modal propositions;
- **The Fifth Method**, on the Contradiction and Conversion of Propositions, containing 5 remarks concerning the contradiction between absolute propositions, and the determination of the contradictory of absolute and concrete propositions, contradiction in the remaining modal propositions, the conversion of absolute propositions, the conversion of necessary propositions, the conversion of possible propositions;
- **The Sixth Method** containing one remark concerning propositions, with respect to those involving assent, and similar ones;
- **The Seventh Method**, on the Beginning the Second Composition of Proof containing 7 remarks concerning the syllogism, induction and analogy, the syllogism, the conjunctive syllogism, the various types of predicative conjunctive syllogisms, the first figure, the second figure, and the third figure;
- **The Eighth Method**, on Conditional Syllogisms, and on What Follows the Syllogism, containing 4 remarks concerning conditional conjunctive syllogisms, the syllogism of equals, repetitive conditional syllogisms, the syllogism by contradiction;
- **The Ninth Method**, in which a Brief Explication of the Demonstrative Science is given, containing 6 remarks concerning the various types of syllogisms, with respect to their matters and their production of assent, the syllogisms and the demonstrative inquiries, the subjects, principles, questions [and transference of demonstrations] in the sciences, the correspondence of the sciences, causal demonstration and factual demonstration, the questions [in the sciences];
- **The Tenth Method**, On Fallacious Syllogisms.

5 A Comparison Between Logic in *Shifā* and Logic in *Ishārāt*

One may regard the first method of *Ishārāt* as equivalent to the "Introduction" (= *al-Madkhal* or *Isagoge*) of logic in *Shifā*. Of course, some topics have been explained in *Shifā* in detail (such as the Five Universals), while there is not enough place for the words, the essential, and the accidental. These topics have been broadly studied in *Ishārāt*.

In *Shifā*, Ibn Sīnā pays attention to the fact that there are some topics in the foundation of logic that are not parts of logic but they are parts of first philosophy. We emphasize here on first philosophy as the study of being qua being or, broadly speaking, the study of the essences as they exist in the mind or external to the mind. There have been three different attitudes towards the relation between logic and philosophy: Different followers of Aristotelian school of logic believed that logic is a tool for philosophy; the Stoics held that logic is a part of philosophy; and, according to Platonists, logic is a part of philosophy and, at the same time, a tool for philosophy.

He, in the *Introduction* (= *Madkhal/Isagoge*), speaks of the precursor's habit or practice to make long the foundations and preliminaries of logic with some topics that do not belong to logic but to first philosophy. He also speaks of another book (other than *Shifā*, i.e. *Falsafat ol-Mashrighīyyah/Philosophy of the Orientals/Easterners' Philosophy*) of his own having been written in which, contrary to his sympathetics (i.e. Peripatetic philosophers), the philosophical problems have been brought forth for discussion in accordance with the nature of the matter, avoiding the Peripatetics' method (Ibn Sīnā [9], *Shifā, al-Madkhal*, 9–12). In another place, he insists that he has avoided mentioning such problems, bringing them in their own appropriate place (Ibn Sīnā [13], *Burhān* (= *Demonstration*), p. 10). Ibn Sīnā has no commitment throughout the work to this view, so that he studies the categories in the techniques on logic. He goes on to say that he has another book in which he has presented philosophy according to his own specific view. He says explicitly that *Shifā* is more extensive and more sympathetic towards the Peripatetics (*ibid*). He is aware of the difference between his own style and predecessors' manners in writing the books on philosophy and logic: he wants to postpone some discussions concerning the universal affirmative proposition to the technique on syllogism according to habit or custom, although it is better to be stated in the third technique.

Moreover, it must be emphasized that while Ibn Sīnā's approach towards logic in *Shifā* is material, he has a formal approach towards logic in *Ishārāt*.

Even in his Persian book under the title *Dāneshnāme 'Alāyī* (= *'Alāyī Encyclopedia*), Ibn Sīnā has put the section on "Definition" before the section on "Propositions", and his discussions concerning Dialectic, Rhetoric, and Poetics are very short (similar to corresponding discussions in *Najāt* (= *deliverance*)).

6 Ibn Sīnā's Reasons on Changing the Structure of Logic and the Style of Logicography

One's method in logicography is based on his/her definition of logic as Ansārī has said: "if you know why you should read logic, then you would know how you should read logic" (Ansārī [1], p. 320). We mention here only some main reasons for choosing two-partite logic and logicography.

6.1 Dividing Knowledge into Conception and Assent

Following al-Fārābī in his Uyūn ol-Masā'el, Ibn Sīnā proceeds by dividing knowledge into two types: (i) conception (*tasawwur*), being the mere imaging or grasping of an object without any judgment (the result of conception being called "concept", with two types, simple and composite), and (ii) assent (*tasdīq*), presupposing conception.

Such a division is useful for discussing the purpose of logic. One may find this division even in *Shifā*. It must be said that if knowledge has two types, ignorance too has two types, ignorance in relation to conception and ignorance in relation to assent. The purpose of logic is to transfer from ignorance in one of the forms of conception and assent to knowledge in one of those two forms (*Burhān*, p. 18).

Such an approach reaches its perfection in *Ishārāt*. It is in this work that Ibn Sīnā gives appropriate names according to what is customary: "*It is customary to call the thing by means of which the sought concept is attained "an explanatory phrase", which includes definition, description, and what resembles them; and to call the thing by means of which the sought assent is attained "proof", which includes syllogism, induction, and their like*." (*Ishārāt 1984*, *p. 49*)

According to such a view, concept brings out of concept and assent out of assent. This is the principal basis of two-part logicography, based on "definition" and "proof" as two types of thought. Ibn Sīnā's Isagoge, and, indeed, Isagoge in the Islamic tradition of logic, is indeed the introduction to the logic of definition, while the study of propositions is introduction to the logic of proof.

6.2 Study of Categories as a Part of First Philosophy

Ibn Sīnā explicitly says in the *Categories* of *Shifā* that inclusion of the categories in logic is not correct since they belong to (i) first philosophy according to the quality and existence, (ii) natural philosophy or physics (as neighbor of first philosophy) according to their establishment in human mind, and (iii) lexicography according to the words used to refer to them (Ibn Sīnā [10], *Shifā: al-Maqūlāt*, 5–8).

In *Ishārāt*, Ibn Sīnā argues that though the First Teacher (i.e. Aristotle) opens his teachings with ten categories, they are not among the subject matter of logic. Indeed, they are first intelligibles, while the subject matter of logic is second (logical) intelligibles.

Ibn Sīnā says that the subject matter of logic are mental subjects, having no external correspondents, so that they are second intelligibles and predicates without any existence for them in the external world.

Logic, therefore, pays no attention to individuals (with external or mental existence) and the essences of the existents. It concerns the second (logical) intelligibles and mental concepts with the names such as predicates, subjects, universals, and particulars. Logic, of course, speaks of some meanings of the words, without any necessity of studying such matters as its task (Ibn Sīnā, *Shifā*: *Theology*, pp. 10–11, *Introduction*, pp. 22–23). He even insists that speaking of categories in logic is to be considered a mistake (Ibn Sīnā [11], *Ishārāt* 1971, Part 1, p. 43).

It must be said that Tūsī, in his commentary on *Ishārāt*, takes the nine-partite system in spite of his referring to the views of the moderns (i.e. logicians such as Ibn Sīnā, in opposition to Aristotle as the precursor) in not regarding the specification of the natures of the universals and study of the objectivity of the existents (either substance or accident) as belonging to logic.

Yet, he insists that the art of definition and obtaining the premises of deductions is not possible without having conception of the categories and distinguishing the categories from each other. He is speaking as if the founder of the logic (i.e. Aristotle) himself has determined the place of the categories (as represented in the treatise *Categories*) to be the first book of the logic and logicography. Moreover, he emphasizes on the usefulness of knowing the categories in giving examples to make the explanation of a problem easy (Tūsī [16], p. 42).

Any way, it is evident that Tūsī, in spite of accepting the nine-partite style of logicography, agrees that categories are among metaphysical foundations of logic (i.e. they are not genuine logical problems), though so important that must be put at the beginning of logic and logicography. Categories, as first intelligibles or natural universals as the higher genuses and accidents of the existents, are not the problems of logic but some of the foundations for it. Their entrance into logic depends upon the view of a logician concerning the inclusion of them in first philosophy or in any introduction to logic.

While following the precursors' method of logicography in *Shifā*, Ibn Sīnā discusses substances and accidents in the "*Theology*" (i.e. the section concerning first philosophy or metaphysics) of *Shifā*, too. It is in this book that he speaks explicitly of the fact that discussing such problems is not the task of the logician so that such undertaking for logicians is a deviation in his/her due course.

Finally, we add the words of Ibn Khaldun, in justifying the method of the moderns in eliminating the categories in their logicography: *the logicians consider the categories accidentally not essentially* (cf. Ibn Khaldun [6], vol. 2, pp. 1024–1028).

7 Two-Partite Logic Versus Nine-Partite Logic in the Islamic Tradition

The nine-partite logic and logicography, the corpus of which consists of the eight treatises of Aristotle and Porphyry's Isagoge, became an important tradition in writing and teaching logic in the Islamic word. It was adopted by Fārābī, Ibn Sīnā (before *Ishārāt*), Bahmanyār (Ibn Sīnā's disciple), Ibn Rushd, Ikhvān os-Safā, Tūsī, Qutb od-Dīn Shīrāzī, Dashtakī, Ansārī, and many other logicians as the principal tradition of Aristotelian logic in the Islamic world. This is the very tradition that Rescher has called the School of Baghdād (Rescher [15], p. 14).

Against such a tradition, the new approach towards the structure and contents of logic was extensively welcomed particularly in the Eastern part of the Islamic world (e.g. Persian world). The typical topics of the textbooks of logic, showing the style of lexicography and the structure of the texts, may be introduced as follows:

1. A compendious knowledge of logic
2. Study of the words
3. Isagoge of the five universals (an introduction to the logic of definition)
4. Definition
5. Propositions (an introduction to the logic of argumentation)

6. Syllogism (proof in general)
7. The five figures (i.e. demonstration, dialectic, sophistry, rhetoric, poetics)
8. Situation of sciences (logic of science and scientology)

Ibn Khaldun speaks of the new system of logicography attributing it to Fakhr Rāzī (Ibn Khaldun [6], p. 492). It is evident that he is wrong in such an attribution. His report shows the attitude of the famous theologians, philosophers, and logicians such as Rāzī.

In addition to Rāzī, the new system attracted Ghazālī, Suhravardī, Urmawī, Khunjī, Abharī, Kātebī, Qutb o-Ddīn Rāzī, Allāmeh Hellī, Taftāzānī, and Mullā Sadrā.

Three of the logicians belonging to such "moderns" have had celebrated place in establishing and transferring bipartite style of logicography:

(i) Afzal od-Din Khunjī (590–646). Khunjī has written some important famous books such as *"Kashf ol-Asrār an Qumūz al-Afkār"* (= *Disclosing the Secrets of the Complexities of Thoughts*). According to Farāmarz Qarāmalekī, Fakhr Rāzī, with his critical commentary on Ishārāt (called *Al-Enārāt fī Sharh al-Ishārāt* (= *Clarifications in Commentary on al-Ishārāt*), one of the many commentaries on Ishārāt and commentaries on commentaries on Ishārāt), is the leading intermediate between Ibn Sīnā and Khunjī. (Farāmarz Qarāmalekī 1373/1994, 46)

(ii) Serāj od-Din Urmawī (594–682). Urmawī has a book, called *Bayān ol-Haqq va Lesān os-Sedq* (= *Expression of the Right and Language of the Truth*), to give an explanatory report of Khunjī's book. He has written a short book called *Matāle' ol-Anvār* (= *Rising Place of the Lights*).

(iii) Najm od-Din Kātebī Qazwīnī (600–675) has written some commentaries on Khunjī's and Rāzī's books as well as some other books on logic including a very famous textbook under the title *Resāle-ye Shamsīyyeh* (= *Solar Treatise*) that has been read and taught for several centuries.

It is true that from the 7th (13th) century (after the Hejira, i.e. 13th century A.D.) onward two-part system of logicography became dominant in Iranian schools as the manifestation of the school in logic following Ibn Sīnā for which Nicholas Rescher chooses the name "Eastern School" (following Ibn Sīnā himself in calling the book of his late career *Mantiq al-Mashriqīyyīn* (= *Logic of the Orientals*)) in opposition to "School of Baghdad" or "Western School" with its nine-partite logic (Rescher [15], pp. 15–17). The school of Baghdad has been founded by Abu Bishr Mattā Ibn Yūnus, the translator of *Posterior Analytics* and the principal teacher of Fārābī. Accordingly, Rescher writes:

> *Ibn Sīnā's call to study logic from independent treatises rather than via the Aristotelian texts met with complete success in Eastern Islam, where after the demise of the School of Baghdad, the formal study of Aristotle's logical writings came to an end.* (*This abandonment of Aristotle may have been a requisite for the survival of Greek logic in Islam; a discipline that demanded study of works of an alien philosopher could probably not have survived.*) *Only in Muslim Spain did the tradition of Aristotelian studies of the School of Baghdad manage—for a time—to survive.* (Rescher [15], p. 16)

Mohammasd Taqī Dāneshpazhūh has reported that Vattier, the Latin translator of Ibn Sīnā's Najāt in the 17th century, has spoken of the two systems of logicography (Qarāmalekī [5, p. 43]).

In recent years, some books on logic have been written in the framework of some combination of both systems of logicography without significant insistence on the history of the development of logicography. Hasan Malekshāhī [14], and Ahad Farāmarz

Qarāmalekī [3–5] have paid specific attention to the differences of the two systems of logicography.

8 Main Features and Advantages of Ibn Sīnā's Two-Partite Methodology in Logicography

One may sum up the main features of two-partite logicography as follows:

1. It regards logic as an independent art or technique not as a tool of sciences.
2. On the basis of division of thought into two parts (i.e. conception and assent), logic was divided into two parts: (i) logic of definition and (ii) logic of proof (of argument).
3. Matters concerning science were eliminated from "demonstration" and were attached to logic independently.
4. The detailed discussion on the Five figures (i.e. demonstration, dialectic, sophistry, rhetoric, and poetics) was eliminated, the elimination being justified on the basis of two principal end of logic: (i) obtaining truth and (ii) avoiding error ([8], Najāt, Arabic, p. 93).
5. There is no place for categories (as some parts of metaphysical foundations of Aristotelian logic) in such a logic because the foundations of a science as logic are not among its problems (being, in fact, the problems of a higher order science). The problems of each science are the rules and essential accidents of the subject matter of that science. According to Ibn Sīnā, the study of categories belongs to first philosophy. Categories are absent from Ibn Sīnā's *an-Najāt* (= *Deliverance*) Uyūn ol-Hekmah (= *Sources of the Sophia*), and *Mantiq al-Mashriqīyyīn* (= *Logic of the Orientals*). Ghzālī too regards the study of categories as a part of theology (i.e. general theology or first philosophy) in his *Maqāsid ol-Falāsefeh* (= *Aims of the Philosophers*), although he puts categories at the end of his *Me'yār ol-Ilm* (= *Criterion of the Knowledge*), a task that is followed by Rāzī in his *Resāle-ye Kamālīyyeh* (= *the Perfection Treatise*). It is interesting to know that while Ibn Sīnā has written the logic section of Shifā in accordance with the nine-partite system of logicography, he speaks of categories in philosophy section of Shifā, too.
6. The place of "definition" has been changed from the section on "demonstration" and "dialectic" in *Shifā* to a specific place before "propositions" or "interpretation" (as a part of the "concepts", after the Five Universals) in *Ishārāt* (in a more systematic detailed and separated form).
7. Dialectics, Poetics and Rhetoric were eliminated. They found independence in combination with demonstration and sophistry under the title of the "Five Figures" (= *Senā'āt-e Khams*) as material logic (the Five Figures either have been eliminated, or have been studied at the end of the books on logic).
8. Some new definitions of logical concepts and issues such as "essential", "three propositions", "the quantified" and some others were represented.
9. There is an emphasis on the role of the words on the basis of the role of the language as reflection of the mind.
10. "Definition" attained independence.
11. Scientology, being distinguished from formal logic, was presented as an appendix.

On the basis of what was said, we may give a hint of some advantages of Ibn Sīnā's deviation from the tradition of the Aristotelians or the followers of the School of Baghdād in their representation of both contents and structure of logic and related textbooks:

I Emphasis on Formal Identity of Logic
II New Attitude Towards concepts of logic, particularly "Definition", on the Basis of New Foundations
III Separation of Formal Logic from Material Logic: demonstration, dialectic, sophistry, rhetoric, and poetics (as the five figures) are the material logic of the proof. Accordingly, they must be separated from formal logic.
IV Separation of Logic from First Philosophy
V While Ibn Sīnā, in some places of his books (in particular his early works), speaks of logic as a tool of philosophy, he regards it as an independent discipline, though with its role in movement from the known to the unknown.
VI Having good justification for readiness and fluency of both teaching and learning of logic, on the basis of natural manner of thinking.

References

1. Ansārī, M.I.J.: Tohfat al-Salātin. Mirās-e Maktūb, Tehran (1386/2007). Critical Edition by Ahad Farāmarz Qarāmalekī
2. Barnes, J.: Introduction to Introduction by Porphyry. Clarendon Press, Oxford (2003)
3. Faramarz Qarāmalekī, A.: Al-Ishārāt wa al-Tanbihat, beginning of two-part logic. Āyīne-ye Pazhuhesh **24**, 38/630–50/642 (1373/1995) (in Persian)
4. Faramarz Qarāmalekī, A.: The Role of Ibn Sīnā in Development of Logic. In: Proceedings of International Conference of Ibn Sīnā. Hamadan University Press, Hamadan (1383/2005) (in Persian)
5. Faramarz Qarāmalekī, A.: Structural development in Ibn Sīnā's logicography and its historical influence in Islamic Era. Hekmat-e Sīnavī (Meshkāt ol-Anvār) **13**, 41–60 (1388/2009) (in Persian)
6. Ibn Khaldun: Muqaddamah (= Introduction). Bongah-e Tarjemah va Nashr-e Ketab, Tehran (1352/1973), translated into Persian by Mohammad Parvīn Gunābādī, 2 vols.
7. Ibn Nadīm: Al-Fihrist. Asātīr Publishing Co. with the cooperation of the International Centre for Dialogue Among Civilizations, Tehran (1381/2002). Translated into Persian by M.R. Tajaddod
8. Ibn Sīnā: Najāt (= Deliverance). Matba'at os-Sa'ādeh, Cairo (1357/1938)
9. Ibn Sīnā: Shifā, al-Madkhal (= Introduction). Matba'at ol-Amīrīyyah, Cairo (1371/1952)
10. Ibn Sīnā: Shifā, al-Maqūlāt (= Categories). Matba'at ol-Amīrīyyah, Cairo (1379/1959)
11. Ibn Sīnā: al-Ishārāt wat-Tanbīhāt, 2nd edn. (1971). Ed. Sulayman Dunya
12. Ibn Sīnā: Remarks and Admonitions, Part One, Logic. Pontifical Institute of Medieval Studies, Toronto (1984), tr. by Shams C. Inati
13. Ibn Sīnā: Burhān (= Demonstration). Fekr-e Rūz, Tehran (1373/1994), translated into Persian by M. Qavām Safarī
14. Malekshāhī, H.: Tarjomeh va Sharh-e Al-Ishārāt wat-Tanbīhāt (= Remarks and Admonitions: Translation and Commentary). Soroush, Tehran (1367/1988), 5th printing, 1388/2009
15. Rescher, N.: Studies in the History of Arabic Logic. Pittsburgh University Press, Pittsburgh (1963)
16. Tūsī, N. od-D.: Asās ol-Iqtibās. Nashr-e Markaz, Tehran (1374/1995), critical edition with commentaries by 'Abdollāh Anvār

M. Akrami (✉)
Department of Philosophy of Science, Science and Research Branch of Tehran, Islamic Azad University, Tehran, Iran
e-mail: musa.akrami@srbiau.ac.ir

Homotopical Categories of Logics

Peter Arndt

Abstract Categories of logics and translations usually come with a natural notion of when a translation is an equivalence. The datum of a category with a distinguished class of weak equivalences places one into the realm of abstract homotopy theory where notions like homotopy (co)limits and derived functors become available. We analyze some of these notions for categories of logics. We show that, while logics and flexible translations form a badly behaved category with only few (co)limits, they form a well behaved homotopical category which has all homotopy (co)limits. We then outline several natural questions and directions for further research suggested by a homotopy theoretical viewpoint on categories of logics.

Keywords Logics · Categories · Higher categories · Abstract homotopy theory

Mathematics Subject Classification (2000) Primary 03B22 · Secondary 55U35

1 Introduction

In his opening lecture at Unilog 2010, Jean-Yves Béziau named the following as the main questions of Universal Logic:

1. What is a logic?
2. What is a translation between logics?
3. When are two logics equivalent?
4. How to combine logics?

Lots of different answers to these questions have been proposed over time, with a recent increase of activity spurred by the contests of the Unilog conference series.

The consideration of categories of logics is a way of evaluating and comparing such answers. First, observe that answering questions 1 and 2 usually results in a category whose objects are logics and whose morphisms are translations. One then gets tentative answers to questions 3—two logics might be called equivalent if they are isomorphic in that category—and 4—a combination of logics may be seen as the formation of a colimit in this category, following [47]. However, these answers to questions 3 and 4 are rarely satisfying.

To see this, let us place ourselves in the setting of Hilbert systems, i.e. formal languages generated by some primitive connectives and variables and endowed with a consequence relation. A *strict translation* is a map of the formal languages sending generating connectives to generating connectives and preserving consequence, while a *flexible translation*

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_2

may send generating connectives to more complex formulas (both are required to map n-ary connectives to n-ary connectives).

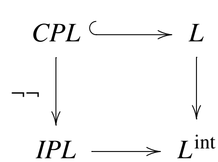
Now for question 3 consider the two presentations of classical propositional logic $CPL_1 := \langle \wedge, \neg \mid \text{rules} \ldots \rangle$ and $CPL_2 := \langle \wedge, \neg, \vee, \to \mid \text{rules} \ldots \rangle$. Clearly, one would say that both are presentations of "the same" logic, since the connectives $\vee$ and $\to$ appearing additionally in the second logic, are expressible, up to logical equivalence, by compositions of $\wedge$ and $\neg$, and do not need to be present as primitive symbols. Indeed, the inclusion of formal languages $CPL_1 \to CPL_2$ has the property that it is conservative, i.e. inferences hold in the target logic if and only if they hold in the domain logic, and that any formula of the target logic is logically equivalent to one in the image. We will call a translation with these properties a *weak equivalence*.

But there can be no isomorphism between these logics. There is no strict translation at all from CPL_2 to CPL_1, since it would have to map the binary connectives $\vee$ and $\to$ to the only binary connective $\wedge$ of CPL_1, but such a map cannot preserve consequence, since the connectives satisfy different rules. We do have a flexible translation from CPL_2 to CPL_1 which maps, for example, $- \vee -$ to the derived formula $\neg(\neg(-) \wedge \neg(-))$. But this cannot be part of an isomorphism since going back via the inclusion results in the map $CPL_2 \to CPL_2$ which sends $- \vee -$ to $\neg(\neg(-) \wedge \neg(-))$ – the formulas are logically equivalent, but not equal, and an the composition of an isomorphism with its inverse has to give the identity.

Thus we found that a sensible notion of equivalence is an extra notion, and does not emerge from the categorical structure.

For question 4 about combining logics, the answer that the combination of logics should be a colimit is often a good one where it applies, but this is only the case for a restricted class of diagrams of logics. Essentially, only colimits of diagrams of strict morphisms exist and behave well.[1] This includes a lot of cases from practice, but it would be even nicer to be able to combine logics along flexible morphisms.

Consider, for example, a modal extension of classical propositional logic, presented by $L := \langle \wedge, \neg, \vee, \to, \Box, \Diamond \mid \text{rules} \ldots \rangle$. It receives an inclusion of classical propositional logic $CPL := \langle \wedge, \neg, \vee, \to \mid \text{rules} \ldots \rangle$. Now we might be interested in what happens if we make the underlying propositional logic of L intuitionistic by removing the law of excluded middle from its rules. For example, we might ask whether properties like algebraizability or the validity of a metatheorem of deduction will still hold and how to construct a semantics for the new logic. Such questions have been amply addressed in the theory of fibring of logics, so we could try to express our "intuitionistified" logic L^{int} as a fibring, i.e. a pushout, of logics:



The idea is that we embed the classical propositional sublogic of L into intuitionistic logic IPL along the double negation translation (the left vertical arrow), and glue the extra, not

[1] See Example 2.8 for a colimit of flexible morphisms which does exist, but does not behave right.

doubly negated, layer that intuitionistic logic has in comparison to *CPL*, to the modal logic *L* while maintaining the place that was formerly occupied by *CPL*.

The problem is that this colimit does not exist in the category of Hilbert systems and flexible translations, and indeed very few colimits exist there.

The conclusion that we draw from these observations is that it is better to regard questions 1, 2 and 3 as fundamental questions: We should first ask for notions of logic, translation and weak equivalence. Then we have a category with an additional structure, a distinguished class of morphisms given by the weak equivalences. As harmless as it looks, this has vast implications: Such a pair consisting of a category and a class of morphisms, also called a relative category, is all one needs, to do an abstract form of homotopy theory.

The usual categorical notions and constructions can now be accompanied with their "derived" versions. For example, there is the notion of homotopy colimit: Usual colimits do not need to preserve weak equivalences, i.e. given two weakly equivalent diagrams their colimits need not be weakly equivalent. A homotopy colimit can roughly be thought of as the best approximation of a colimit construction which preserves weak equivalences. One could argue that, if one devises a logically meaningful construction of a new logic from some given other logics, then one would like equivalent inputs to lead to equivalent outputs and that thus the derived notions are the better ones. Maybe more importantly, homotopy (co)limits can exist where (co)limits do not exist. Indeed, for Hilbert systems all homotopy colimits do exist.

Another benefit from working with relative categories is this: Relative categories are commonly regarded not as the important objects in themselves, but rather as *presentations* of a so-called $(\infty, 1)$-category. As an analogy, in group theory one can have different presentations by generators and relations of a group, and these can be useful for answering different questions about the group, while the actual object of interest is still (the isomorphism class of) the group itself. Analogously there can be different relative categories which are presentations of the same $(\infty, 1)$-category, and when asking questions whose answers are invariant under equivalence, there is no harm in switching to a better suited presentation. There is a theory of $(\infty, 1)$-categories, very much parallel to usual category theory, where one studies such invariant properties and the $(\infty, 1)$-categories of logics that we consider here have much better properties in this realm, than in the usual category theoretical world where they arose.

Overview of the Article. In this article, we explore a bit of the homotopy theoretical perspective on logics that we have hinted at. In Sect. 2, we review usual categories of logics and pin down, what is the problem with categories of flexible morphisms. In Sect. 3, we give a quick tour through some concepts of abstract homotopy theory to explain the setting in which we wish to study categories of logics. Here we can only serve some rough ideas of a huge area, but we need very little of the full scope of abstract homotopy theory and the next chapter—particularly Sect. 4.2—can be read with very few prerequisites. The main ingredients that will be used from here are simplicial categories and how they arise from 2-categories as well as the notion of equivalence of simplicial categories.

Section 4 is the technical heart of the article, where we investigate categories of Tarski style logics. We first introduce, in Sect. 4.1, the two different notions of homotopy equivalence and weak equivalence and see how they relate differently in the strict and flexible settings. We next, in Sect. 4.2, give a short preview on work to appear that addresses the so-called the hammock localization of the category of Hilbert systems. It is in this setting that we can particularly well handle the combination of logics along flexible translations, which we hinted at above. In particular, we get preservation results for homotopy colimits parallel to those for fibring, which was what motivated our discussion of question 4. Our treatment of this, however, uses a particular kind of presentation of an $(\infty, 1)$-category, which we felt was too much to expose here in detail.

Section 4.3 exploits the fact that the set of translations between two logics carries a natural equivalence relation: Two translations $f, g \colon L \to L'$ can be called equivalent if for every formula φ of L the images $f(\varphi)$ and $g(\varphi)$ are mutually derivable from each other in L'. Equivalence relations are a special kind of groupoid and thus categories of logics can be seen as 2-categories and come with a natural notion of equivalence, which under mild hypotheses coincides with the ones of Sect. 4.1. The resulting $(\infty, 1)$-categories, which we call the 2-categorical localizations, are homotopy theoretically very simple; their mapping spaces are homotopy discrete. This means that they are equivalent to the quotient categories with respect to the above equivalence relations. These quotient categories have been studied by Mariano and Mendes in [41] and [42], where they show, among other things, that the quotient category of congruential Hilbert systems is complete and cocomplete. In Sect. 4.3.2, we show how these results of Mariano and Mendes, can be cast into the language of $(\infty, 1)$-categories, here embodied by categories enriched in simplicial sets. On the one hand, this is because, in our view, the simplicial categories are the natural objects one would want to study, and that this boils down to the study of their homotopy categories could be seen as merely a technical convenience. On the other hand, this is to offer the reader an easy entry point to get acquainted with the language—it is the language that will be needed for the more refined categories of logics of Sect. 5.3.

It is in Sect. 4.3 that the main technical results of the article appear. These are: Theorem 4.26 (crucially relying on work of Mariano and Mendes), which asserts that in the world of $(\infty, 1)$-categories the category of logics and flexible morphisms is a reflexive subcategory of that of strict morphisms, Theorem 4.39, which asserts that the category of logics and flexible morphisms has all homotopy limits (contrary to the 1-categorical case) and the discussion of Sect. 4.3.3, which asserts the existence of all homotopy colimits. We chose to construct homotopy limits in a pedestrian way to give a feeling of how one can handle single logics homotopically, and to sketch a proof of the existence of homotopy colimits by abstract results, to give a different sample of homotopy theoretical methods.

In the remaining Sect. 5, we gather questions and prospects for further developments suggested by the homotopy theoretical viewpoint on logic.

Acknowledgements and Birthday Wishes: I thank Eyjafjallajökull for letting me give a "volcanic lecture" on these, then still very vague, thoughts at the Unilog 2010. I am greatly indebted to Hugo Mariano for helpful conversations, for sharing the work [42] with me on which some of the key results here rely and which made me notice an error in an earlier version, and for his very stimulating interest in these ideas. I thank Markus Spitzweck for helpful conversations and Oliver Bräunling, Moritz Groth, Krzysztof Kapulkin, Caio Mendes, Darllan Pinto and Michael Völkl for their comments on an earlier version. And I give my thanks to Jean-Yves for the spirit of the Unilogs and the whole endeavor of Universal logic, where such ideas can flourish.

Happy birthday, Jean-Yves!!

2 Categories of Logics

2.1 Signatures

Definition 2.1 A *signature* S is a sequence of sets $(S_n, n \in \mathbb{N})$.

We think of the elements of S_n as the generating n-ary connectives of a formal language. We fix once and for all a set $\mathrm{Var} := \{x_n \mid n \in \mathbb{N}\}$ of variables and denote as usual by $\mathrm{Fm}(S)$ the absolutely free algebra with signature S generated by Var. We have a decomposition $\mathrm{Fm}(S) = \coprod_{n\in\mathbb{N}} \mathrm{Fm}(S)[n]$, where $\mathrm{Fm}(S)[n]$ denotes the set of formulas with n free variables. We also denote by $\mathrm{Fm}(S)[x_1, \dots, x_n]$ the set of formulas containing exactly the variables $x_1, \dots, x_n$.

Definition 2.2 A *strict morphism* $f : S \to S'$ of signatures is a sequence of maps $(f_n : S_n \to S'_n, n \in \mathbb{N})$.

A *flexible morphism*, or simply a *morphism*, $f : S \to S'$ of signatures is a sequence of maps $(f_n : S_n \to \mathrm{Fm}(S')[x_1, \dots, x_n], n \in \mathbb{N})$.

Thus a strict morphism is an arity-preserving map sending generating connectives to generating connectives while a flexible morphism can be seen as a map sending generating connectives to derived connectives. A strict morphism can be seen as a flexible morphism which happens to send generating connectives to generating connectives (where a generating connective $c \in S_n$ is seen as the formula $c(x_1, \dots, x_n) \in \mathrm{Fm}(S')[x_1, \dots, x_n]$), see Definition 2.5.1.

A morphism $f : S \to S'$, either strict or flexible, induces a map $f : \mathrm{Fm}(S) \to \mathrm{Fm}(S')$ which is inductively defined as usual.

Example 2.3 The usual double negation translation from the standard signature of classical propositional logic to the standard signature of intuitionistic logic is a flexible morphism which is not strict as it sends, for example, the binary connective $\wedge$ (or $(x_1 \wedge x_2)$) to the derived connective $(\neg\neg(x_1) \wedge \neg\neg(x_2))$. One can make the double negation translation into a strict morphism, if one chooses to present intuitionistic logic with extra connectives and axioms: To resolve, for example, the above obstacle to a strict translation, one could add a binary connective $\wedge^{\mathrm{class}}$ to the presentation of intuitionistic logic and add the axiom $(x_1 \wedge^{\mathrm{class}} x_2) \dashv\vdash (\neg\neg(x_1) \wedge \neg\neg(x_2))$. Then a translation from classical logic could be defined by sending $\wedge$ to $\wedge^{\mathrm{class}}$.

Definition 2.4 The category $\mathcal{S}ig^{\mathrm{strict}}$ is the category whose objects are signatures and whose morphisms are strict morphisms. The category $\mathcal{S}ig$ is the category whose objects are signatures and whose morphisms are flexible morphisms.

We note that the category $\mathcal{S}ig^{\mathrm{strict}}$ is equivalent to the category $\mathbf{Set}^{\mathbb{N}}$ of sequences of sets and morphisms, in particular it is complete and cocomplete.

Central to our main results in Sect. 4.3 is the following adjunction established by Mariano and Mendes in [42].

Definition 2.5 ([42, Proposition 1.5, Mariano/Mendes])

1. The functor $i\colon \mathcal{S}ig^{\text{strict}} \to \mathcal{S}ig$ is defined on objects by the identity and on morphisms by associating to $f = (f_n)_{n\in\mathbb{N}}\colon S \to S'$ the flexible morphism $i(f)$ with $i(f)_n\colon S_n \to \mathrm{Fm}(S')[x_1, \ldots, x_n]$, $c \mapsto (f_n(c))(x_1, \ldots, x_n)$.
2. The functor $Q\colon \mathcal{S}ig \to \mathcal{S}ig^{\text{strict}}$ is defined on objects by $S \mapsto Q(S)$ where $Q(S)_n := \mathrm{Fm}(S)[x_1, \ldots, x_n]$ and by sending a flexible morphism $f\colon S \to S'$, given by $(f_n\colon S_n \to \mathrm{Fm}(S')[x_1, \ldots, x_n])$, to the sequence of induced maps $\mathrm{Fm}(S)[x_1, \ldots, x_n] \to \mathrm{Fm}(S')[x_1, \ldots, x_n]$.

Theorem 2.6 ([42, Theorem 1.6, Mariano/Mendes]) *The functor i is left adjoint to Q.*

Proof The natural isomorphisms

$$\mathrm{Hom}_{\mathcal{S}ig}\big(i(S), S'\big) \cong \big\{\big(f_n\colon S_n \to \mathrm{Fm}\big(S'\big)[x_1, \ldots, x_n]\big)_{n\in\mathbb{N}}\big\} \cong \mathrm{Hom}_{\mathcal{S}ig^{\text{strict}}}\big(S, Q\big(S'\big)\big)$$

follow straight from the definitions of the morphisms of $\mathcal{S}ig$ (resp., $\mathcal{S}ig^{\text{strict}}$) and the functors i and Q. □

The unit $S \to i(Q(S))$ of the adjunction is given by the inclusions $S_n \to \mathrm{Fm}(S)[x_1, \ldots, x_n]$, $c \mapsto c(x_1, \ldots, x_n)$. The counit $Q(i(S)) \to S$ is the flexible morphism given by the identity maps $\mathrm{Fm}(S)[x_1, \ldots, x_n] \to \mathrm{Fm}(S)[x_1, \ldots, x_n]$.

In fact, by [42, Theorem 1.12, Mariano/Mendes], the category $\mathcal{S}ig$ is the Kleisli category of the above adjunction. Thus it is a category of free algebras and has much worse categorical properties than $\mathcal{S}ig^{\text{strict}}$. It is neither complete nor cocomplete. This is to be expected, as (co)limits of free algebras, formed within their (co)complete ambient category of all algebras, are not usually free again.

Example 2.7 The category $\mathcal{S}ig$ has no terminal object. Indeed, a terminal signature would have to have a generating connective of arity ≥ 2, since if there were only generating connectives of arities 0 and 1 the sets of n-ary formulas $\mathrm{Fm}(S)[x_1, \ldots, x_n]$ would be empty and there could be no morphism from a signature with n-ary connectives. But if c is an n-ary connective, then $\mathrm{Fm}(S)[x_1, \ldots, x_{2n-1}]$ contains the two different formulas $c(c(x_1, \ldots, x_n), x_{n+1}, \ldots, x_{2n-1})$ and $c(x_1, \ldots, x_{n-1}, c(x_n, \ldots, x_{2n-1}))$ and hence admits two different morphisms from the signature with just one generating $(2n-1)$-ary connective.

Some (co)limits do exist. Among these the colimits imported via the functor i (which is left adjoint, hence colimit preserving) from the cocomplete category $\mathcal{S}ig^{\text{strict}}$ are well behaved, but others are degenerate and do not express what we would like to achieve with them in logic.

Example 2.8 Consider the signature S generated by a single unary connective $\Box$. We have the two flexible morphisms $f, g\colon S \to S$ defined by $f(\Box) := \Box\Box x_1$, $g(\Box) := \Box\Box\Box x_1$, respectively. Any flexible morphism $h\colon S \to S'$ which satisfies $h \circ f = h \circ g$ (i.e. which "coequalizes" f and g) can only map $\Box$ to the variable $x_1 \in \mathrm{Fm}(S')[x_1]$: If $\Box$ is mapped to any other formula $\varphi(x_1)$ then the formulas $\Box\Box x_1$ and $\Box\Box\Box x_1$ will have the images $\varphi(\varphi(x_1))$ and $\varphi(\varphi(\varphi(x_1)))$, and these images will be different because the target is an absolutely free algebra. The coequalizer can then easily be seen to be empty signature

$\emptyset$ with no generating connectives: By the usual definition of formulas, the set of formulas is the smallest set containing all variables and closed under application of connectives, so there we have $x_1 \in \mathrm{Fm}(\emptyset)[x_1]$ and the uniqueness property of a colimit is satisfied since this signature is initial. However, this is not what one would like in practice. The coequalizer should remember that there was a connective $\Box$ and that "$\Box\Box = \Box\Box\Box$", not just forget it completely.

Remark 2.9 One could adopt a yet more flexible notion of morphism by considering the set $\mathrm{Fm}(S')\langle x_1, \ldots, x_n\rangle$ of formulas which contain no other variables than $x_1, \ldots x_n$ (but are allowed to contain less than these), i.e. $\mathrm{Fm}(S')\langle x_1, \ldots, x_n\rangle := \coprod_{i=0}^{n} \mathrm{Fm}(S)[x_1, \ldots, x_i]$, and then defining the set of morphisms as $\mathrm{Hom}(S, S') := \{(f_n\colon S_n \to \mathrm{Fm}(S')\langle x_1, \ldots, x_n\rangle)_{n\in\mathbb{N}}\}$. Such morphisms would no longer preserve the arity of formulas, and, for example, one could "delete" n-ary connectives by mapping them to the single variable x_1 (substitution into which would correspond to the identity operation). Much of what will be said in this article would carry over to this setting, as well as to many other variants, but the notions of morphism we chose to consider seem to be the ones of biggest interest in practice.

Definition 2.10 A *substitution* is a map $\sigma\colon \mathrm{Var} \to \mathrm{Fm}(S)$.

Again a substitution induces an inductively defined map $\sigma\colon \mathrm{Fm}(S) \to \mathrm{Fm}(S)$.

2.2 Logics

Definition 2.11 1. Let S be a signature. A *consequence relation* over S is a relation $\vdash \subseteq \mathscr{P}(\mathrm{Fm}(S)) \times \mathrm{Fm}(S)$ between subsets of $\mathrm{Fm}(S)$ and elements of $\mathrm{Fm}(S)$. As usual we write it in infix notation $\Gamma \vdash \varphi$.

2. A *logic* is a pair $L = (S, \vdash)$, where S is a signature and $\vdash$ a consequence relation on $\mathrm{Fm}(S)$.

Given a logic L, we will sometimes denote its underlying signature by S_L and its consequence relation by $\vdash_L$.

Definition 2.12 A consequence relation is *Tarskian* if the associated operation $\mathrm{Cn}\colon \mathscr{P}(\mathrm{Fm}(S)) \to \mathscr{P}(\mathrm{Fm}(S))$, $\Gamma \mapsto \{\varphi \mid \Gamma \vdash \varphi\}$ satisfies

1. (increasingness) $\Gamma \subseteq \mathrm{Cn}(\Gamma)$ for all $\Gamma \subseteq \mathrm{Fm}(S)$
2. (idempotence) $\mathrm{Cn}(\mathrm{Cn}(\Gamma)) \subseteq \mathrm{Cn}(\Gamma)$ for all $\Gamma \subseteq \mathrm{Fm}(S)$
3. (monotonicity) $\Gamma \subseteq \Gamma' \Rightarrow \mathrm{Cn}(\Gamma) \subseteq \mathrm{Cn}(\Gamma')$ for all $\Gamma, \Gamma' \subseteq \mathrm{Fm}(S)$

These conditions say exactly that Cn is a closure operator on $\mathscr{P}(\mathrm{Fm}(S))$.

Two common further additional conditions that one likes to impose on consequence relations are:

(Finitarity) If $\Gamma \vdash \varphi$ then there exists a finite subset $\Gamma' \subseteq \Gamma$ such that $\Gamma' \vdash \varphi$
(Substitution invariance) If $\Gamma \vdash \varphi$, then for any substitution σ we have $\sigma(\Gamma) \vdash \sigma(\varphi)$

Substitution invariance is also called *structurality*.

Finally, the role of the following notion for the study of categories of logics has been brought to light by Mariano and Mendes in [41, 42].

Definition 2.13 A logic $(S, \vdash)$ is *congruential* if for every two sequences of formulas $\varphi_1, \ldots \varphi_n$ and $\psi_1, \ldots, \psi_n$ with $\varphi_i \dashv\vdash \psi_i$ and every pair of formulas $\beta(x_{i_1}, \ldots, x_{i_n})$, $\gamma(x_{i_1}, \ldots, x_{i_n})$ with $\beta \dashv\vdash \gamma$ we have $\beta(\varphi_1, \ldots \varphi_n) \dashv\vdash \gamma(\psi_1, \ldots, \psi_n)$.

Congruentiality is a notion taking its place in the Leibniz hierarchy of degrees of algebraizability.

By the results of Łoś and Suszko in [38], the consequence relations that are finitary, substitution invariant and Tarskian are exactly the provability relations coming from a Hilbert style system, where consequence is given by finite derivations using axioms, rules and substitutions. We therefore call a logic $(S, \vdash)$ a *Hilbert system* if the consequence relation has these properties. On the semantical side, there is Wójcicki's result from [52], saying that a finitary and substitution invariant Tarskian logic is sound and complete for an appropriate finitary matrix semantics (see also [53, Theorem 3.1.6]).

The consequence relations over a fixed signature S can be ordered by setwise inclusion: $\mathrm{Cn}_1 \leq \mathrm{Cn}_2 :\Leftrightarrow \mathrm{Cn}_1(\Gamma) \subseteq \mathrm{Cn}_2(\Gamma)\ \forall \Gamma \subseteq \mathrm{Fm}(S)$. Obviously, consequence relations (without further conditions) form a complete lattice with respect to the above order. By [53, Theorem 1.5.4], also the subset of Tarskian consequence relations on $\mathrm{Fm}(S)$ forms a complete lattice with respect to this order; by [53, Theorems 1.5.5–1.5.6], the same is true for the subsets of finitary (resp., structural) Tarskian consequence relations; and finally, by [53, Theorem 1.5.7], the same is true for Hilbert systems.

About congruential Hilbert systems there is the following result by Mariano and Mendes:

Proposition 2.14 ([42, Proposition 2.18, Mariano/Mendes]) *The category of congruential Hilbert systems is a reflexive subcategory of the category of all Hilbert systems.*

By considering colimits of diagrams of congruential logics whose underlying signature morphisms are the identity, one can conclude that congruential Hilbert systems form a complete lattice of consequence relations as well. It is also easy to see that intersections of congruential consequence relations are congruential again.

In particular, for a signature S we have on $\mathrm{Fm}(S)$ a maximal and a minimal consequence relation of each of the types just listed.

The completeness of the considered lattices of consequence relations gives us the possibility of defining direct and inverse image logics, as done for Hilbert systems in [2, Definition 2.9]:

Definition 2.15 Given a logic $(S, \vdash)$ and a signature morphism $f : S \to S'$, we can view $\vdash$ as a subset of $\mathscr{P}(\mathrm{Fm}(S)) \times \mathrm{Fm}(S)$, take the set-theoretic image $f(\vdash) \subseteq \mathscr{P}(\mathrm{Fm}_{S'}) \times \mathrm{Fm}_{S'}$ and define the *direct image* $f_*(\vdash)$ as the infimum of all consequence relations of the given type (e.g. Tarskian (resp., finitary and/or structural Tarskian) consequence relations) containing $f(\vdash)$.

Likewise, given a logic $(S', \vdash')$ and a signature morphism $f : S \to S'$, the *inverse image* $f^*(\vdash')$ can be defined as the infimum of all consequence relations of the given type (i.e. Tarskian (resp., finitary and/or substitution invariant Tarskian) consequence relations) on $\mathrm{Fm}(S)$ containing $f^{-1}(\vdash')$, the set-theoretic pre-image of $\vdash'$.

Definition 2.16 A *translation* (resp., *strict translation*) $L = (S, \vdash) \to (S', \vdash') = L'$ of logics is a signature morphism (resp., *strict signature morphism*) $f : S \to S'$ such that $\Gamma \vdash \varphi \Rightarrow f(\Gamma) \vdash' f(\varphi)$.

Remark 2.17 As noted in [2, Fact 5 (p. 12)], given two logics $L = (S, \vdash)$, $L' = (S', \vdash')$, a signature morphism $f : S \to S'$ is a translation iff $\vdash \leq f^*(\vdash')$ iff $f_*(\vdash) \leq \vdash'$.

Definition 2.18 We denote by $\mathcal{LOG}$ the category of logics whose objects are logics and whose morphisms are translations. We denote by $\mathcal{L}og^{(\text{Tarsk})}$ (resp., $\mathcal{L}og^{(\text{fin,Tarsk})}$, $\mathcal{L}og^{(\text{subst,Tarsk})}$, $\mathcal{L}og^{(\text{subst,Tarsk,con})}$, etc.) the full subcategory of Tarskian (resp., of finitary Tarskian, substitution invariant Tarskian, substitution invariant congruential Tarskian, etc.) logics. Finally, we denote by $\mathcal{H}ilb$ the full subcategory of Hilbert systems and by $\mathcal{H}ilb^{(\text{con})}$ the full subcategory of congruential Hilbert systems.

Convention 2.19 For the remainder of the article, we denote by $\mathcal{L}og$ any of the following full subcategories of $\mathcal{LOG}$: $\mathcal{LOG}$, $\mathcal{L}og^{(\text{Tarsk})}$, $\mathcal{L}og^{(\text{con})}$, $\mathcal{L}og^{(\text{Tarsk,con})}$, $\mathcal{L}og^{(\text{fin,Tarsk})}$, $\mathcal{L}og^{(\text{subst,Tarsk})}$, $\mathcal{L}og^{(\text{subst,Tarsk,con})}$, $\mathcal{H}ilb$, $\mathcal{H}ilb^{(\text{con})}$. By the terms "logic" and "consequence relation" we will mean a logic (resp., a consequence relation) taken from this chosen category $\mathcal{L}og$. If we need to distinguish consequence relations from $\mathcal{LOG}$ and consequence relations defining objects of $\mathcal{L}og$, we will call the latter "admissible consequence relations". In parts of Sect. 4, we will need to assume additional properties of our logics and will then say so.

We invite the reader to read the article with a specific category of logics, such as $\mathcal{L}og^{(\text{Tarsk})}$ or $\mathcal{H}ilb$ in mind.

Remark 2.20 For much of what follows we could be rather flexible about what properties exactly we demand from our consequence relations. Much of the article can be read by fixing a set of properties that one wishes our consequence relations to have and that satisfy the following assumptions:

1. The lattice of consequence relations satisfying the properties is complete
2. The direct image maps satisfy $g_*(f_*(\vdash)) = (g \circ f)_*(\vdash)$

These assumptions are exactly what is needed for Proposition 2.24 below to hold, i.e. that one can construct (co)limits in the corresponding category of logics by constructing them in $\mathcal{S}ig$ (resp., $\mathcal{S}ig^{\text{strict}}$) and then endow the resulting signature with an appropriate consequence relation. Much of Sect. 4, however, needs the property of idempotence.

Remark 2.21 The completeness of the lattice of consequence relations of a chosen kind also makes it possible to define a consequence relation by giving generating rules. For example, given a logic $(S, \vdash)$ and $\varphi, \psi \in \text{Fm}(S)$, the Tarskian consequence relation generated by $\vdash$ and the rules $\{\varphi\} \vdash \psi$, $\{\psi\} \vdash \varphi$ is defined to be the infimum in the lattice of Tarskian consequence relations of all those consequence relations containing $\vdash$ and which satisfy $\{\varphi\} \vdash \psi$ and $\{\psi\} \vdash \varphi$. We will freely make use of such constructions and if we talk of the consequence relation generated by some given rules, we will always mean the consequence relation in our chosen category $\mathcal{L}og$.

Definition 2.22 Let $U\colon \mathcal{L}og \to \mathcal{S}ig$, $(S, \vdash) \mapsto S$ denote the obvious faithful forgetful functor forgetting the consequence relation. We will denote its restriction to the subcategories of strict morphisms $\mathcal{L}og^{\text{strict}} \to \mathcal{S}ig^{\text{strict}}$ by U as well.

Lemma 2.23 *The functor U has a left adjoint $Min\colon \mathcal{S}ig \to \mathcal{L}og$, $S \mapsto (S, \vdash_{\min})$ which endows a signature S with the minimal consequence relation on* $\mathrm{Fm}(S)$, *as well as a right adjoint $Max\colon \mathcal{S}ig \to \mathcal{L}og$, $S \mapsto (S, \vdash_{\max})$ placing the maximal consequence relation on* $\mathrm{Fm}(S)$. *These adjunctions restrict to the subcategories of strict morphisms.*

Proof The direct/inverse image characterization of translations of Remark 2.17 implies that, for a morphism of signatures $f\colon S \to S'$, the pair (f_*, f^*) is a pair of adjoint functors between the preorders of consequence relations, seen as categories. Here f_* is the left adjoint, hence it preserves colimits, i.e. suprema of consequence relations. In particular, it preserves the supremum of the empty family, i.e. the minimal consequence relation, i.e. $f_*(\vdash^S_{\min}) = \vdash^{S'}_{\min}$. Now we have natural bijections

$$\mathrm{Hom}_{\mathcal{L}og}\big(Min(S), L\big) = \big\{ f \in \mathrm{Hom}_{\mathcal{S}ig}(S, S_L) \;\big|\; f_*(\vdash_{\min}) \leq \vdash_L \big\} = \mathrm{Hom}_{\mathcal{S}ig}(S, S_L)$$

where the left equality is the direct image characterization of translations of Remark 2.17 and the right equality can be seen from the fact that $f_*(\vdash^S_{\min}) = \vdash^{S_L}_{\min}$ and thus the condition in the middle set is empty.

This shows that Min is left adjoint to U. The right adjointness of Max works by dualizing the proof, the restriction statement is clear. □

Note that we have $U \circ Min = U \circ Max = id_{\mathcal{S}ig}$.

Proposition 2.24 *The category $\mathcal{L}og$ has (co)limits of a given diagram shape if and only if the (co)limits of this shape exist in $\mathcal{S}ig$.*

Proof Since $U\colon \mathcal{L}og \to \mathcal{S}ig$ has a left and a right adjoint, it preserves colimits and limits. Thus if (co)limits of shape D exist in $\mathcal{L}og$, then, given a diagram of shape D in $\mathcal{S}ig$, we can lift it to $\mathcal{L}og$, e.g. via the functor Min, take the colimit in $\mathcal{L}og$ and apply U. This will yield the (co)limit of the diagram we started with.

Conversely, if colimits of shape D exist in $\mathcal{S}ig$, then given a diagram $F\colon D \to \mathcal{L}og$, we can take the colimit of the underlying diagram $U \circ F$ in $\mathcal{S}ig$ and then endow the resulting signature with the smallest consequence relation such that all the incoming signature morphisms from the original diagram become translations (this exists because of the completeness of the lattice of consequence relations). This construction of colimits is done in all details in [2, Proposition 2.11].

Likewise, limits in $\mathcal{L}og$ can be constructed by taking them in $\mathcal{S}ig$ and endowing the resulting signature with the maximal consequence relation such that all outgoing signature morphisms into the diagram become translations.

An inspection of the proofs in [2] (which are written for Hilbert systems) shows that the assumptions of Remark 2.20 are all that is needed. □

On the one hand, this shows that the category of $\mathcal{L}og^{\text{strict}}$ of logics and strict morphisms is complete and cocomplete, since the underlying category $\mathcal{S}ig^{\text{strict}} \simeq \mathbf{Set}^{\mathbb{N}}$ of signatures is complete and cocomplete. On the other hand, the category $\mathcal{S}ig$ of signatures and flexible translations is not (co)complete as seen in the last section and hence the same is true for $\mathcal{L}og$.

Clearly, we would like a category with the more flexible morphisms, in which we can perform constructions such as (co)limits and which has good categorical properties such local presentability. Proposition 2.24 seems to rule this out. It is one of the main points of this article to argue that, in fact, we do not just have a category $\mathcal{L}og$, but instead a category endowed with an extra structure: A distinguished class of morphisms which we want to see as "equivalences". This extra structure tells us that $\mathcal{L}og$ is most naturally seen not as a category but as a so-called $(\infty, 1)$-category. There is a theory of $(\infty, 1)$-categories, largely parallel to usual category theory, whose basic ideas we will sketch in the following section. It will then turn out in Sect. 4 that, seen as an $(\infty, 1)$-category, $\mathcal{L}og$ does not have the defects that it has as a 1-category.

3 Abstract Homotopy Theory

In this chapter, we review how the basic datum of a category with a distinguished class of morphisms gives rise to structures and notions of a homotopy theoretical flavor.

A *relative category* is a pair $(\mathcal{C}, W)$ where $\mathcal{C}$ is a category and $W \subseteq \mathrm{Mor}\,\mathcal{C}$ is a class of morphisms containing the identity morphisms. We will call the morphisms in W the *weak equivalences* and will try to find constructions and notions that are invariant under weak equivalences.

3.1 Simplicial Sets and Nerves

The archetypical example of a category with a distinguished class of weak equivalences is the category **Top** of topological spaces and continuous maps, where a weak equivalence $X \to Y$ is a map inducing isomorphisms $\pi_n(X) \to \pi_n(Y)$ between all homotopy groups π_n, $n \geq 1$ (with respect to all possible base points) and between the sets of connected components $\pi_0(X), \pi_0(Y)$. One important feature of abstract homotopy theory is that one can often replace a relative category by another one which is better behaved but contains the same information regarding "weak equivalence types". For the category of topological spaces the most common choice is the relative category of simplicial sets:

Definition 3.1 A *simplicial set* is a functor $\Delta^{\mathrm{op}} \to$ **Set**, where Δ denotes the category of finite linearly ordered sets and order preserving maps. A morphism of simplicial sets is a natural transformation. We denote the category of simplicial sets by **sSet**.

The category Δ (or a skeleton thereof) can be described by generators and relations: We can take as objects the standard linearly ordered sets $[n] := \{0 \leq \cdots \leq n\}$ and observe that any order preserving map arises as a composition of the basic maps $[n] \to [n-1]$ which identify two neighboring numbers and the maps $[n] \to [n+1]$ which leave a gap between two neighboring numbers. To give a functor from Δ (or Δ^{op}) to some category, it is then enough to say what it does on objects and on these basic maps and to ensure that the choice of values on the maps satisfies some relations. Thus a simplicial set can alternatively be described as a diagram of sets of the following shape

$$X_0 \;\substack{\longleftarrow \\ \longrightarrow \\ \longleftarrow}\; X_1 \;\substack{\longleftarrow \\ \longrightarrow \\ \longleftarrow \\ \longrightarrow \\ \longleftarrow}\; X_2 \;\substack{\longleftarrow \\ \longrightarrow \\ \longleftarrow \\ \longrightarrow \\ \longleftarrow \\ \longrightarrow \\ \longleftarrow}\; X_3 \;\substack{\longleftarrow \\ \longrightarrow \\ \longleftarrow \\ \longrightarrow \\ \longleftarrow \\ \longrightarrow \\ \longleftarrow \\ \longrightarrow \\ \longleftarrow}\; \cdots$$

in which the arrows satisfy certain relations (see, e.g. [22, p. 4]).

One may think about a simplicial set as follows: The set X_n is the set of n-simplices of some abstract simplicial complex, and the maps $X_n \to X_{n-1}$ identify the faces of these n-simplices with certain $(n-1)$-simplices from X_{n-1}. Thus two n-simplices can share a common face and the whole diagram can be seen as giving combinatorial data for pasting together the several collections of simplices (let us not care about the index increasing maps).

We record the following standard result from category theory (see, e.g. [33, 2.7.1]):

Lemma 3.2 *Let* $\mathbf{C}$ *be a small category,* $\mathcal{D}$ *a cocomplete category and* $F: \mathbf{C} \to \mathcal{D}$ *a functor. Then there is an adjunction* $\overline{F}: \mathbf{Set}^{\mathbf{C}^{op}} \rightleftarrows \mathcal{D}: \mathrm{Hom}_{\mathcal{D}}(F(\cdot), -)$.

Here the left adjoint $\overline{F}$ is given by left Kan extension of F along the Yoneda embedding, i.e. by expressing a presheaf as a colimit of representables, mapping the representables to $\mathcal{D}$ as prescribed by F and then taking the colimit there.

There is a cosimplicial object in **Top**, i.e. a functor $\Delta \to \mathbf{Top}$, which associates to the object $[n]$ the standard n-simplex $\Delta^n := \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} \mid \sum x_i = 1,\ x_i \geq 0\ \forall i\}$, which send increasing maps to inclusions of faces of the standard simplex and decreasing maps to continuous maps collapsing a simplex to one of its faces, see [22, Example 1.1]. By Lemma 3.2, this induces an adjunction $|\cdot|: \mathbf{sSet} \rightleftarrows \mathbf{Top}: Sing$. Here the value $|X|$ of the left adjoint at a simplicial set X, called the *geometric realization* of X, is given by taking a standard n-simplex in **Top** for each element of the set X_n and gluing these simplices together as suggested by the face maps. The right adjoint is defined on objects by $Sing(X)_n := \mathrm{Hom}_{\mathbf{Top}}(\Delta^n, X)$ and on maps by precomposition of face inclusions/retractions. Geometric realization can be seen as a formalization of the intuition about simplicial sets offered in the previous paragraph.

One says that a map of simplicial sets is a *weak equivalence* if its geometric realization is a weak equivalence of topological spaces. A first indication (but not the complete story) that for the purpose of studying spaces up to weak equivalence one can replace **Top** by **sSet** is the fact that the unit and counit of the above adjunction are weak equivalences at each object, i.e. one has weak equivalences $|Sing(X)| \to X$ and $Y \to Sing(|Y|)$ so that going back and forth between the two categories results in weakly equivalent objects. A more complete statement is that **Top** and **sSet** form equivalent $(\infty, 1)$-categories, see below. We therefore sometimes refer to simplicial sets as "spaces", for example, when we will talk about mapping "spaces" below.

Similarly, we can define the *nerve functor* $N: \mathbf{Cat} \to \mathbf{sSet}$ by applying Lemma 3.2: There is a cosimplicial object in the category **Cat** of (small) categories simply given by considering the linearly ordered sets of Δ as categories in the usual way, i.e. taking the numbers $0, \dots, n$ as objects and declaring that there is a unique morphism from i to j if $i \leq j$. Mapping out of this cosimplicial object into a fixed category $\mathcal{C}$ produces a simplicial set $N(\mathcal{C}) := \mathrm{Hom}_{\mathbf{Cat}}(\Delta, \mathcal{C})$ as above, and the nerve functor is the right adjoint of Lemma 3.2 in this special case.

More concretely, the nerve of a category $\mathcal{C}$ is the simplicial set with $N(\mathcal{C})_0 = Ob\,\mathcal{C}$, $N(\mathcal{C})_1 = \mathrm{Mor}\,\mathcal{C}$, $N(\mathcal{C})_n = \{$chains of n composable morphisms of $\mathcal{C}\}$ and whose structure maps are given by composing arrows (resp., inserting identity morphisms into a chain).

3.2 Localization of Categories and Homotopy (Co)limits

One thing we can do with a relative category $(\mathcal{C}, W)$ is to force the morphisms from W to become isomorphisms: That means we can construct a category $\mathcal{C}[W^{-1}]$ together with a functor $L\colon \mathcal{C} \to \mathcal{C}[W^{-1}]$ mapping morphisms from W to isomorphisms and with the property that any other functor $\mathcal{C} \to \mathcal{D}$ mapping morphisms from W to isomorphisms factorizes through L uniquely up to unique natural isomorphism.

3.2.1 Localizing Categories

Here is a concrete construction of $\mathcal{C}[W^{-1}]$: As objects we take the objects of $\mathcal{C}$. To define the morphisms from A to B, first say that a zig-zag from A to B is a sequence of morphisms of $\mathcal{C}$ with arrows pointing in either direction

$$A \leftarrow X_1 \rightarrow X_2 \leftarrow \cdots \leftarrow X_n \rightarrow B$$

and in which the arrows pointing from right to left are from W. Two zig-zags can be related if one arises from the other by (a) composing two consecutive arrows which point into the same directions, (b) deleting identity arrows, or (c) deleting two arrows which are equal in $\mathcal{C}$ and point into opposite directions in the zig-zag. Now consider the equivalence relation $\sim$ generated by the relations (a), (b) and (c) and define $\mathrm{Hom}_{\mathcal{C}[W^{-1}]}(A, B) := \{$zig-zags from A to $B\}/\sim$. Composition in $\mathcal{C}[W^{-1}]$ is induced by concatenation of zig-zags. We have an obvious functor $L\colon \mathcal{C} \to \mathcal{C}[W^{-1}]$ mapping the morphisms of $\mathcal{C}$ to equivalence classes of zig-zags of length one. The arrows pointing from right to left can be seen as newly added inverses to the arrows of W and with this in mind it is not hard to prove the desired universal property of the functor L. Note that since the zig-zags can range over all the objects of $\mathcal{C}$, the class $\mathrm{Hom}_{\mathcal{C}[W^{-1}]}(A, B)$ is not a set in general (or lives in a higher Grothendieck universe), but in many cases in practice it turns out to be a set again. One also writes $Ho(\mathcal{C}, W) := \mathcal{C}[W^{-1}]$ and calls this category the *homotopy category* of $\mathcal{C}$ with respect to W. If the class W is clear, one often suppresses it from the notation and simply writes $Ho(\mathcal{C})$.

3.2.2 Homotopy (Co)limits

Suppose that $(\mathcal{C}, W)$ is a relative category and $\mathcal{C}$ has colimits of some diagram shape $\mathbf{D}$. The diagram category $\mathcal{C}^{\mathbf{D}}$ is itself a relative category with weak equivalences given by objectwise weak equivalences of $\mathcal{C}$, hence we also have a homotopy category $Ho(\mathcal{C}^{\mathbf{D}})$. By the universal property of the localization, if the colimit functor preserves weak equivalences we can find a functor (represented by the dotted arrow below) making the following diagram commute:

$$\begin{array}{ccc} \mathcal{C}^{\mathbf{D}} & \xrightarrow{colim} & \mathcal{C} \\ \downarrow & & \downarrow \\ Ho(\mathcal{C}^{\mathbf{D}}) & \dashrightarrow & Ho(\mathcal{C}) \end{array}$$

However, the colimit functor has no reason to preserve weak equivalences. A standard example in the category of topological spaces is to consider the pushouts $S^1 = \mathrm{colim}([0, 1] \leftarrow \{0, 1\} \rightarrow [0, 1])$ (the circle, obtained by gluing two unit intervals along their end points) and $* = \mathrm{colim}(* \leftarrow \{0, 1\} \rightarrow *)$ (where $*$ denotes the 1-point space): The unit interval $[0, 1]$ is contractible, so the obvious transformation from the first to the second diagram is an equivalence. Yet the respective colimits of the two diagrams are not equivalent (since S^1 has a nontrivial fundamental group).

The next best thing that we can do then is to form the right Kan extension of $\mathcal{C}^{\mathbf{D}} \rightarrow \mathcal{C} \rightarrow Ho(\mathcal{C})$ along $\mathcal{C}^{\mathbf{D}} \rightarrow Ho(\mathcal{C}^{\mathbf{D}})$—this is the universal approximation of the colimit construction by a construction that preserves weak equivalences. It results in a diagram

$$\begin{array}{ccc} \mathcal{C}^{\mathbf{D}} & \xrightarrow{colim} & \mathcal{C} \\ \downarrow & \nearrow\!\!\!\!\nearrow & \downarrow \\ Ho(\mathcal{C}^{\mathbf{D}}) & \xrightarrow{hocolim} & Ho(\mathcal{C}) \end{array}$$

which is not commutative but instead filled in with a universal natural transformation. We emphasize that the homotopy colimit is usually not the colimit in the homotopy category; indeed, in general $Ho(\mathcal{C}^{\mathbf{D}})$ is not equivalent to $Ho(\mathcal{C})^{\mathbf{D}}$, so it does not even have the right domain category. The notion of *homotopy limit* is dual, with a left Kan extension instead of a right one.

In the category of simplicial sets, all homotopy (co)limits in the sense just defined exist, and many techniques have been developed for computing them. We will use this fact in the next section.

In [16], the reader can find an exposition of the basic notions and statements around relative categories in which the weak equivalences satisfy a very mild closure condition. A more powerful setting continuing this line of thought about homotopy (co)limits is the theory of derivators; see, for example, [26] and the references therein.

We note that these approaches to homotopy (co)limits allow us to stay in the framework of usual category theory. However, they have their limitations, and a richer setting with a well developed theory is that of $(\infty, 1)$-categories; see the next sections.

3.3 Simplicial Categories, Simplicial Localizations and Homotopy (Co)limits

Lots of experience indicates that the category $\mathcal{C}[W^{-1}]$ is in general too crude an object. If one wants to treat objects of $\mathcal{C}$ up to equivalence, the passage from $\mathcal{C}$ to $\mathcal{C}[W^{-1}]$ forgets too much; morphisms from $\mathcal{C}$ get identified uncontrollably and equivalence preserving constructions in $\mathcal{C}$ cannot be characterized or performed just in $\mathcal{C}[W^{-1}]$ alone. This has already been visible in our above definition of homotopy (co)limits: For this we needed both the original relative category $(\mathcal{C}, W)$ and its localization $Ho(\mathcal{C})$.

Also note that the above definition of homotopy (co)limit only makes sense if the (co)limits in question exist in the category $\mathcal{C}$, otherwise we have nothing along which we could take Kan extensions. Many categories of logics, however, lack (co)limits.

One can now pass to a refined variant of localization, called the simplicial localization, which results in a category enriched in simplicial sets, i.e. instead of sets of morphisms we get simplicial sets of morphisms, also called the *mapping spaces* of the simplicial category. Such a category is also called *simplicial category* (this should not be confused with simplicial objects in the category of categories). The basic intuition about a simplicial category is that the 0-simplices of the mapping spaces are morphisms, the 1-simplices are homotopies between morphisms, the higherdimensional simplices are homotopies between homotopies, and so on.

We now give one possible construction of a simplicial localization: The *hammock localization* introduced in [17].

3.3.1 The Hammock Localization

Given a relative category $(\mathcal{C}, W)$ we define its hammock localization $L^H(\mathcal{C}, W)$ to be the following category enriched in simplicial sets: Again we take as objects those of $\mathcal{C}$. For two objects A, B we now have to give a simplicial set $\mathrm{map}(A, B)$, also called the *mapping space* of A and B. We define the nth set of this simplicial set to be the "set" of *reduced hammocks of width n*, i.e. commutative diagrams of the shape

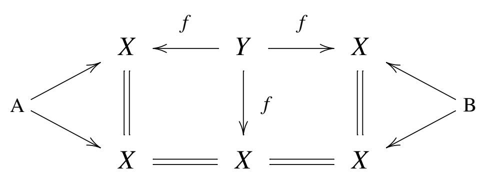
$$\begin{array}{ccccccccccccc}
 & & X_{01} & \longrightarrow & X_{02} & \longleftarrow & X_{03} & \longrightarrow & \cdots & \longrightarrow & X_{0k} & & \\
 & \nearrow & \downarrow & & \downarrow & & \downarrow & & & & \downarrow & \nwarrow & \\
A & \longrightarrow & X_{11} & \longrightarrow & X_{12} & \longleftarrow & X_{13} & \longrightarrow & \cdots & \longrightarrow & X_{1k} & \longleftarrow & B \\
 & \searrow & \vdots & & \vdots & & \vdots & & \ddots & & \vdots & \swarrow & \\
 & & X_{n1} & \longrightarrow & X_{n2} & \longleftarrow & X_{n3} & \longrightarrow & \cdots & \longrightarrow & X_{nk} & &
\end{array}$$

in which the vertical arrows go downwards and are from W, in each column the horizontal arrows all point into the same direction, the horizontal arrows going from right to left are from W, no column consists only of identity arrows and the maps in adjacent columns go into different directions. The structure maps for the simplicial sets are given by composing downward pointing arrows (resp., duplicating a row and inserting identity arrows).

The composition maps $\mathrm{map}(A, B) \times \mathrm{map}(B, C) \to \mathrm{map}(A, C)$ are given by concatenation of zig-zags, followed by reducing the resulting hammocks (i.e. composing adjacent columns pointing into the same direction and deleting identity columns). For more on the hammock localization see [16].

3.3.2 Homotopy Category of a Simplicial Category

From a simplicial category $\mathcal{C}$ one can obtain an ordinary category $Ho(\mathcal{C})$ by passing to the set of connected components of the Hom-simplicial sets, i.e. by keeping the same objects and defining $\mathrm{Hom}_{Ho(\mathcal{C})}(A, B) := \pi_0(\mathrm{map}_{\mathcal{C}}(A, B))$. The resulting category is called the *homotopy category* of the simplicial category. If the simplicial category is the hammock localization $L^H(\mathcal{C}, W)$ of a category with weak equivalences, the homotopy category is exactly the localization from above: $Ho(L^H(\mathcal{C}, W)) \simeq \mathcal{C}[W^{-1}]$. Indeed, the objects are the same, fullness is clear, and the main point for faithfulness is that a zig-zag with two consecutive arrows from W pointing in opposite directions is in the same connected component of $\mathrm{map}(A, B)$ as the zig-zag with these arrows canceled:



The details can be found in [17, Proposition 3.1].

3.3.3 Homotopy (Co)limits Revisited

A simplicial category $\mathcal{C}$ also has an underlying usual category $U(\mathcal{C})$ given by just remembering the 0-simplices of the mapping spaces, i.e. $\mathrm{Hom}_{U(\mathcal{C})}(A, B) := \mathrm{map}_{\mathcal{C}}(A, B)_0$. There is a natural class W of weak equivalences on $U(\mathcal{C})$ given by those morphisms which become isomorphisms in $Ho(\mathcal{C})$ (alternatively one could take as weak equivalences morphisms $f \in \mathrm{map}(A, B)_0$ such that for all objects X the induced map $\mathrm{map}(X, A) \to \mathrm{map}(X, B)$ is a weak equivalence of simplicial sets; the two definitions coincide in good cases).

For those types of diagrams which have (co)limits in $U(\mathcal{C})$, we thus have the notion of homotopy (co)limit introduced above by Kan extensions. However, in a simplicial category one can also speak of homotopy (co)limits without supposing that the corresponding (co)limits are present in the underlying category $U(\mathcal{C})$, by defining them through the mapping spaces. The idea is that in ordinary category theory one can define the limit of a diagram $\mathbf{D}$ as a representing object for the functor $X \mapsto \lim_{d\in\mathbf{D}} \mathrm{Hom}_{\mathcal{C}}(X, d)$, i.e. by asking for a natural isomorphism of **Set**-valued functors $\lim_{d\in\mathbf{D}} \mathrm{Hom}_{\mathcal{C}}(-, d) \simeq \mathrm{Hom}_{\mathcal{C}}(-, \lim_{d\in\mathbf{D}} d)$. Now in a simplicial category one can instead ask for a natural weak equivalence of **sSet**-valued functors $\mathrm{holim}_{d\in\mathbf{D}} \mathrm{map}_{\mathcal{C}}(-, d) \cong \mathrm{map}_{\mathcal{C}}(-, \mathrm{holim}_{d\in\mathbf{D}} d)$, where the homotopy limit on the left hand side is taken in simplicial sets where we already know what it means by the definition through Kan extensions. For more details see [40, A.3.3.13].

A different formulation which easily relates to classical category theory is to ask for a final object in the simplicial category of "homotopy coherent" cones over the given diagram, but we will not go into this further and instead refer the reader to the exposition in [48].

3.3.4 Simplicial Categories from 2-Categories

Besides the hammock localization of a relative category there is a further source of simplicial categories relevant for us:

A 2-category is a category enriched in the category **Cat** of categories, i.e. for any two objects A, B there is a category $\underline{\mathrm{Hom}}(A, B)$, together with composition functors satisfying the usual axioms. In such a category, there is a natural class of weak equivalences given by those morphisms $f: A \to B$ for which there exist a $g: B \to A$ and isomorphisms $f \circ g \simeq id_B \in \underline{\mathrm{Hom}}(B, B)$, $g \circ f \simeq id_A \in \underline{\mathrm{Hom}}(A, A)$. One way to get a simplicial category from this is to form the hammock localization with respect to the class of weak equivalences.

Another way is, for each pair of objects A, B, to take the maximal subgroupoid of the category $\underline{\mathrm{Hom}}(A, B)$, i.e. the subcategory of all objects and isomorphisms between them, and apply the nerve-functor to each of them (we could also apply the nerve functor to the whole category $\underline{\mathrm{Hom}}(A, B)$, but this would not capture the same class of weak equivalences).

These two constructions, though both very natural, do in general yield non-equivalent simplicial categories in the sense of Definition 3.3 below. The reason for this is that the hammock localization cannot distinguish between different automorphisms of the $\underline{\mathrm{Hom}}$-categories, while the nerve construction clearly captures them.

3.4 Equivalences of Simplicial Categories and $(\infty, 1)$*-Categories*

Simplicial categories are objects of usual enriched category theory, as exposed in [34], where one considers categories which have, instead of sets of morphisms $\mathrm{Hom}(A, B)$, objects of morphisms $\underline{\mathrm{Hom}}(A, B)$ which live in some monoidal category $\mathcal{M}$ (here the category of simplicial sets with the monoidal structure given by the product). An $\mathcal{M}$-enriched functor $F: \mathcal{C} \to \mathcal{D}$ is a mapping of objects $Ob(\mathcal{C}) \to Ob(\mathcal{D})$ and, for each pair A, B of objects a morphism $\underline{\mathrm{Hom}}_{\mathcal{C}}(A, B) \to \underline{\mathrm{Hom}}_{\mathcal{D}}(FA, FB)$ in $\mathcal{M}$, compatible with composition and identities—for example, a functor $F: (\mathcal{C}, W) \to (\mathcal{D}, W')$ of relative categories such that $F(W) \subseteq W'$ induces a simplicial functor $L^H(F): L^H(\mathcal{C}, W) \to L^H(\mathcal{D}, W')$ between the hammock localizations as one sees easily from the definition of hammock localization.

In enriched category theory, an $\mathcal{M}$-enriched functor is an equivalence if it is essentially surjective and fully faithful in the enriched sense, i.e. $\underline{\mathrm{Hom}}_{\mathcal{C}}(A, B) \to \underline{\mathrm{Hom}}_{\mathcal{D}}(A, B)$ is required to be an isomorphism in $\mathcal{M}$ for all A, B. However, the role played by the simplicial enrichment in the example of the hammock localization of a relative category suggests that one is not interested in the Hom-simplicial-sets themselves, but only in the homotopy types of spaces they represent. Hence one defines:

Definition 3.3 A simplicial functor $F: \mathcal{C} \to \mathcal{D}$ of simplicial categories is an equivalence if it is

1. *Essentially surjective*, i.e. every object in $\mathcal{D}$ is *equivalent*, i.e. isomorphic in $Ho(\mathcal{D})$, to an object in the image of F and
2. *Fully faithful*, i.e. the maps $\mathrm{map}_{\mathcal{C}}(A, B) \to \mathrm{map}_{\mathcal{D}}(A, B)$ are *weak equivalences* (instead of isomorphisms) of simplicial sets.

Note that in particular an equivalence $F: \mathcal{C} \to \mathcal{D}$ of simplicial categories induces an equivalence $Ho(F): Ho(\mathcal{C}) \to Ho(\mathcal{D})$ of usual categories, but the condition of being an equivalence is stronger than that: For having an equivalence of homotopy categories it would be enough to demand that F induces isomorphisms $\pi_0 \,\mathrm{map}_{\mathcal{C}}(A, B) \to \pi_0 \,\mathrm{map}_{\mathcal{D}}(FA, FB)$ between the sets of connected components $\pi_0 \,\mathrm{map}_{\mathcal{C}}(A, B)$ (resp., $\pi_0 \,\mathrm{map}_{\mathcal{D}}(FA, FB)$) of the mapping spaces, while for an equivalence of simplicial categories one asks for isomorphisms induced on π_0 and on all homotopy groups.

Definition 3.4 An $(\infty, 1)$-category is a simplicial category. A morphism of $(\infty, 1)$-categories is a simplicially enriched functor.

This definition allows keeping things simple and suffices for the purposes of this article. We insist however that with this definition the category of simplicial categories has to be seen as a relative category itself (with the class of equivalences just defined)—different simplicial categories may define equivalent $(\infty, 1)$-categories and it is only the simplicial categories up to equivalence that we are interested in. Roughly, we could also have defined an $(\infty, 1)$-category as an equivalence class of simplicial categories, with respect to the equivalence relation generated by the above notion of equivalence of simplicial categories. For more discussion on this see Sect. 3.5.1.

The name can be explained as follows: One thinks of the 0-simplices in the mapping spaces of a simplicial category as the morphisms (or “1-morphisms”), of the 1-simplices as homotopies between morphisms (or “2-morphisms”), of 2-simplices as homotopies between homotopies (or “3-morphisms”) and so on, so that one has n-morphisms for every $n \in \mathbb{N}$. This explains the “∞” in the term “$(\infty, 1)$-category”. Now homotopies between functions can always be inverted (if there is a homotopy from a function f to a function g, then there also is a homotopy from g to f and the composition of the two homotopies is homotopic to the identity), so that all n-morphisms for $n \geq 2$ are invertible—only the morphisms up to level 1 are actual directed morphisms, while the higher morphisms are (witnesses of) equivalences. This explains the “1” in the term “$(\infty, 1)$-category”. More generally and (n, k)-category is a category which has higher morphisms up to level n, all of which are invertible from level $k + 1$ onwards. An $(\infty, 1)$-category is also sometimes called “a homotopy theory”.

3.5 Models and Computation of Homotopy (Co)limits

There can be very different looking but equivalent simplicial categories. One sees such simplicial categories as *presentations* of $(\infty, 1)$-categories, just like one can define a group by a presentation via generators and relations. A(n isomorphism class of a) group can be defined by very different looking presentations and it can in practice be very difficult to determine whether the groups given by two presentations are isomorphic. In our context, one uses the term *model*, rather than "presentation".

Likewise one can see a relative category as a model (or presentation) of an $(\infty, 1)$-category, since it gives rise to a simplicial category via the hammock localization. This is the way in which $(\infty, 1)$-categories arise most commonly from usual mathematics. We have already seen an example of two different presentations of the same $(\infty, 1)$-category: The categories of simplicial sets and of topological spaces, with their respective classes of weak equivalences, are presentations of the same $(\infty, 1)$-category, commonly called "the $(\infty, 1)$-category of spaces" (we gave the functors inducing this equivalence, but only hinted at the essential surjectivity part).

Now while the datum of a relative category allows formulating the notions of homotopy (co)limits (as well as that of derived functors) it does not help in constructing them in concrete cases or determining whether they exist. To this end, one frequently employs models which are not just categories with a class of weak equivalences, but which are endowed with an additional auxiliary structure.

Probably the best kind of model one can ask for is a *model category*: A model category is a tuple $(\mathcal{C}, W, Fib, Cof)$ consisting of a category $\mathcal{C}$ with finite limits and colimits and three classes of morphisms, W (weak equivalences), *Fib* ("fibrations") and *Cof* ("cofibrations"), which are related by several axioms (often one asks for additional structure, such as "factorization functors"). This extra structure allows concrete constructions of homotopy limits and colimits and also the construction of adjunctions of $(\infty, 1)$-categories by the means of usual category theory.

A gentle introduction to model categories is [19], a more complete one is the book [29]. One downside of model categories is that it is hard to establish the existence of a model structure on a category. The category of logics and flexible morphisms of Sect. 2.2, for example, has no chance of bearing the structure of a model category simply because it lacks (co)limits.

A less demanding kind of model is given by the notion of *cofibration category*, see [46]. This is, roughly, half a model category structure, having only a class of cofibrations and weak equivalences, and only requiring the existence of special colimits. In a cofibration category, homotopy colimits can be constructed explicitly and, by the results of Szumiło in [49], every $(\infty, 1)$-category with all homotopy colimits has a presentation by a cofibration category. The category of logics and flexible morphisms of Sect. 2.2 carries such a structure.

There are several other kinds of models, such as Baues cofibration categories and semi-model categories, which meet different purposes and one has to see in each particular situation whether such a kind of model exists and is useful.

3.5.1 The Homotopy Theory of Homotopy Theories

Above we endowed the category of simplicial categories and simplicial functors with a class of weak equivalences. Thus the category of simplicial categories becomes itself a relative category. We also have a notion of weak equivalence of relative categories: This a functor $F : (\mathcal{C}, W) \to (\mathcal{D}, W')$ satisfying $F(W) \subseteq W'$ and inducing an equivalence $L^H(\mathcal{C}, W) \to L^H(\mathcal{D}, W')$ of simplicial categories on hammock localizations.

The resulting two relative categories, that of simplicial categories and that of relative categories, are again equivalent as relative categories. In fact, on both the category of simplicial categories and the category of relative categories there are model structures, and the equivalence can be given in a highly structured way (see [4, 5]).

There are several other ways of encoding $(\infty, 1)$-categories, where the emphasis is not on extra structure allowing constructions internal to an $(\infty, 1)$-category, but rather on relating $(\infty, 1)$-categories to each other, constructing new $(\infty, 1)$-categories from old ones, and formulating and recognizing properties of $(\infty, 1)$-categories. For a survey of some such settings see [7]. The setting with the best developed theory is that of *quasicategories*, featuring, for example, $(\infty, 1)$-categorical notions of and theorems about fibrations, accessible and locally presentable categories, toposes, sketches and algebraic theories. A gentle introduction is [25], a good further introduction is the first chapter of [32], and the main references are the books [40] and [39].

4 $(\infty, 1)$-Categories of Logics

In this chapter, we consider the categories of logics from Sect. 2.2 as $(\infty,1)$-categories. We will do this in the two ways given in Sects. 3.3.1 and 3.3.4. For both we need to fix a notion of weak equivalences of logics.

4.1 Weak Equivalences

We place ourselves in a category $\mathcal{L}og$ of logics as in Sect. 2.2.

Definition 4.1

1. If L is a logic and $\varphi, \psi \in \mathrm{Fm}(L)$ are formulas satisfying $\{\varphi\} \vdash_L \psi$ and $\{\psi\} \vdash_L \varphi$, we write $\varphi \dashv\vdash_L \psi$ and call the formulas *logically equivalent*.
2. A morphism of logics $f: L = (S_L, \vdash_L) \to (S_{L'}, \vdash_{L'}) = L'$ is called a *homotopy equivalence* if there exists a morphism $g: L' \to L$ (a “homotopy inverse”) such that for all $\varphi \in \mathrm{Fm}(L)$ we have $\varphi \dashv\vdash_L (g \circ f)(\varphi)$ and for all $\psi \in \mathrm{Fm}(L')$ we have $\psi \dashv\vdash_{L'} (f \circ g)(\psi)$.
3. A morphism of logics $f: L \to L'$ is called a *weak equivalence*, if $\Gamma \vdash_L \varphi \Leftrightarrow f(\Gamma) \vdash_{L'} f(\varphi)$ (i.e. it is a “conservative translation”) and if for every formula φ in the target there exists a formula in the image of f with the same arity as φ which is logically equivalent to φ (it has “dense image”).

Proposition 4.2

1. *In any category of logics with idempotent consequence relations, homotopy equivalences are weak equivalences.*
2. *In any category of substitution invariant logics over $\mathcal{S}ig$ (i.e. with flexible morphisms), weak equivalences are homotopy equivalences.*

Proof 1. Let $f: L \to L'$ be a homotopy equivalence. Choose a homotopy inverse g. Any formula $\psi \in \mathrm{Fm}(L')$ is logically equivalent to $(f \circ g)(\psi)$, which is in the image, hence f has dense image. If $\Gamma \vdash \phi$ for $\Gamma \cup \{\varphi\} \subseteq \mathrm{Fm}(L)$, then, since f is a translation, we have $\Gamma \vdash \varphi \Rightarrow f(\Gamma) \vdash f(\varphi)$. Conversely, since g is a translation we have $f(\Gamma) \vdash f(\varphi) \Rightarrow g(f(\Gamma)) \vdash g(f(\varphi))$. For every $\gamma \in \Gamma$ we have $\gamma \vdash g(f(\gamma))$ and also $g(f(\varphi)) \vdash \varphi$. Thus we have $\Gamma \vdash g(f(\Gamma)) \vdash g(f(\varphi)) \vdash \varphi$. Hence, by idempotence, $\Gamma \vdash \varphi$.

2. Let $f: L = (S_L, \vdash_L) \to (S_{L'}, \vdash_{L'}) = L'$ be a weak equivalence. To construct a homotopy inverse $g: L' \to L$, choose for every n-ary generating connective $c(x_1, \ldots, x_n)$ of L' a formula $\varphi \in \mathrm{Fm}(L)$ with $c(x_1, \ldots, x_n) \dashv\vdash_{L'} f(\varphi)$ (which exists since f has dense image). This defines a morphism of signatures $g: S_{L'} \to S_L$, which by construction satisfies $f(g(c(x_1, \ldots, x_n))) \dashv\vdash_{L'} c(x_1, \ldots, x_n)$ for all generating connectives. By substitution invariance, it follows that $f(g(\psi)) \dashv\vdash_{L'} \psi$ for all formulas $\psi \in \mathrm{Fm}(L')$. In particular, this holds for formulas of the form $\psi = f(\varphi)$ for any $\varphi \in \mathrm{Fm}(L)$, i.e. we have $f(g(f(\varphi))) \dashv\vdash_{L'} f(\varphi)$. By conservativity of f, we conclude $g(f(\varphi)) \dashv\vdash_L \varphi$. □

Corollary 4.3 *In a category $\mathcal{L}og$ of idempotent, substitution invariant logics the classes of weak equivalences and homotopy equivalences coincide.*

Remark 4.4 For Tarskian logics the homotopy equivalences have been characterized in [10, Proposition 4.3] as those (flexible) morphisms $f: L \to L'$ for which there exists a morphism $g: L' \to L$, such that f, g induce mutual inverse morphisms between the lattices of theories of L, L'.

Remark 4.5 The comparison of notions of weak equivalences and homotopy equivalences (those which have a morphism in the opposite direction that becomes an inverse in the homotopy category) is a standard theme in abstract homotopy theory. The coincidence of the two classes can be phrased as saying that every object is “fibrant” and “cofibrant”. Indeed, the category $\mathcal{H}ilb$ bears the structure of a so-called cofibration category and it is true that every object of $\mathcal{H}ilb$ is fibrant and cofibrant in the sense of cofibration categories (see [46]).

Remark 4.6 Note that the two classes of weak equivalences and of homotopy equivalences no longer coincide, even for idempotent and substitution invariant logics, if we consider them on the category $\mathcal{L}og^{\mathrm{strict}}$ of logics and strict morphisms. The reason is that to define the homotopy inverse in the proof of Proposition 4.2.2 we needed to map primitive connectives to derived connectives. Indeed, consider the two presentations of classical propositional logic $CPL_1 := \langle \wedge, \neg \mid \text{rules} \ldots \rangle$ and $CPL_2 := \langle \wedge, \neg, \vee, \to \mid \text{rules} \ldots \rangle$ of the introduction. Clearly, we have a conservative inclusion $CPL_1 \to CPL_2$ which has moreover dense image, since every formula of classical propositional logic is equivalent to one built from just the connectives $\wedge$ and $\neg$. So this inclusion is a weak equivalence. A homotopy inverse, or indeed any translation from CPL_2 to CPL_1, in the category $\mathcal{L}og^{\mathrm{strict}}$ would have to map the connective $\vee$ to $\wedge$, since the latter is the only primitive connective of CPL_1 of arity 2, but this is impossible since there are rules satisfied by $\vee$ but not by $\wedge$. We will come back to the relation of $\mathcal{L}og^{\mathrm{strict}}$ and $\mathcal{L}og$ in Sect. 4.3.1.

4.2 The Hammock Localization of $\mathcal{H}ilb$

With Definition 4.1 we have made the category $\mathcal{H}ilb$ of Hilbert systems into a relative category, and hence gained an $(\infty, 1)$-category via hammock localization, which we will denote by $\mathcal{H}ilb_{\text{hamm}}$. We actually have more than just a relative category:

Theorem 4.7 ([1]) *In the category $\mathcal{H}ilb$ of Hilbert systems, denote by W the class of weak equivalences of Definition 4.1 and by Cof the class of translations whose underlying signature morphism maps generating connectives injectively to generating connectives. Then the triple $(\mathcal{H}ilb, W, Cof)$ satisfies the axioms of a cofibration category in the sense of* [46].

The proof of this theorem is out of the scope of this article. We just give a hint of the kind of things one has to do: One of the axioms requires that every morphism factors as a cofibration followed by a weak equivalence. The proof of this for the category $\mathcal{H}ilb$ proceeds in close analogy to that for the category of topological spaces, namely by constructing "mapping cylinders":

Sketch of a Proof of the Factorization Axiom. To factorize a translation $f : L = (S_L, \vdash_L) \to (S_{L'}, \vdash_{L'}) = L'$, define an intermediate logic $\widetilde{L}$ with the signature $S_{\widetilde{L}} := S_L \coprod S_{L'}$, so that the formulas of $\widetilde{L}$ are mixed from the connectives of L and L'. In particular, we have the linguistic fragments $\text{Fm}(L) \subseteq \text{Fm}(\widetilde{L}) \supseteq \text{Fm}(L')$. The consequence relation on $\text{Fm}(\widetilde{L})$ is generated by the rules of $\vdash_L$ on $\text{Fm}(L)$, the rules of $\vdash_{L'}$ on $\text{Fm}(L')$ and by the rules $\varphi \dashv\vdash_{\widetilde{L}} f(\varphi)$ for every $\varphi \in \text{Fm}(L)$. This makes the linguistic fragment $\text{Fm}(L) \subseteq \text{Fm}(\widetilde{L})$ equivalent to its image under f in $\text{Fm}(L')$.

Now we have an obvious cofibration $L \to \widetilde{L}$ which is just the inclusion $\text{Fm}(L) \subseteq \text{Fm}(\widetilde{L})$ and a translation $\widetilde{L} \to L'$ given by mapping the connectives from $\text{Fm}(L') \subseteq \text{Fm}(\widetilde{L})$ to themselves and those of $\text{Fm}(L) \subseteq \text{Fm}(\widetilde{L})$ to their image under f. The latter map is a homotopy equivalence, with a homotopy inverse given by the inclusion $\text{Fm}(L') \subseteq \text{Fm}(\widetilde{L})$.

Corollary 4.8 *The $(\infty, 1)$-category $\mathcal{H}ilb_{\text{hamm}}$ has all homotopy colimits of small diagrams.*

Proof An $(\infty, 1)$-category for which there exists a presentation by a cofibration category, has all homotopy colimits, see [46, 49]. □

A proof of Theorem 4.7 will appear in [1], together with applications to the combination of logics via homotopy colimits: The extra structure of cofibrations gives an easy construction of homotopy colimits by means of those 1-categorical colimits which do exist in $\mathcal{H}ilb$. Also the concrete choice of cofibrations allows transferring the usual preservation results for properties of logics under fibring (like existence of implicit connectives, position in the Leibniz hierarchy, etc.) to the combination of logics through homotopy colimits.

4.3 The 2-Categorical Localization of $\mathcal{L}og$

Our categories of logics from Sect. 2.2 are naturally enriched in preorders: We can define a preorder on $\mathrm{Hom}_{\mathcal{L}og}(L, L')$ by

$$f \leq g \quad :\Leftrightarrow \quad \forall \varphi \in \mathrm{Fm}(L): \quad f(\varphi) \vdash g(\varphi).$$

Since preorders can be regarded as categories, this makes every category enriched in preorders into a 2-category.

Recall that we defined the equivalences of a 2-category to be those 1-morphisms $f : L \to L'$ for which there exists a 1-morphism $g : L' \to L$ and 2-isomorphisms $f \circ g \simeq id_{L'}$, $g \circ f \simeq id_L$. In our context, the existence of these 2-isomorphisms simply means that $f(g(\psi)) \dashv\vdash_{L'} \psi$ and $g(f(\varphi)) \dashv\vdash_L \varphi$. Thus the notion of homotopy equivalence from Definition 4.1 is exactly the notion of equivalence coming from the structure of 2-category.

We can now perform the construction of a simplicial category from Sect. 3.3.4 with the 2-category just defined: Pass to the maximal subgroupoids of the hom-categories and then take their nerves. The maximal subgroupoids are simply the categories having the set of translations as objects and having a unique isomorphism between two translations f and g whenever $\forall \varphi \in \mathrm{Fm}(L) : f(\varphi) \dashv\vdash g(\varphi)$ holds. In this case, we also say that f and g are *homotopic*. We will apply this to the categories $\mathcal{L}og$ and $\mathcal{L}og^{\mathrm{strict}}$ and call the resulting simplicial categories $\mathcal{L}og_{2\text{-cat}}$ and $\mathcal{L}og^{\mathrm{strict}}_{2\text{-cat}}$.

Proposition 4.9 *The simplicial categories $\mathcal{L}og_{2\text{-cat}}$ and $\mathcal{L}og^{\mathrm{strict}}_{2\text{-cat}}$ are equivalent, in the sense of Sect. 3.4, to their respective homotopy categories.*

Proof Since the hom-categories are preorders, the maximal subgroupoids are equivalent to discrete categories: Any two objects are either uniquely isomorphic or live in distinct connected components. Since the nerve functor sends equivalences of categories to weak equivalences of simplicial sets, we can replace the hom-categories by actual discrete categories, namely the set of connected components of the groupoids. By definition, the nerve functor sends discrete categories to discrete simplicial sets; therefore, $\mathcal{L}og_{2\text{-cat}}$ is equivalent to a simplicial category with discrete mapping spaces. Its homotopy category is constructed by taking π_0 (the set of connected components) of each mapping space. Since the mapping spaces are already weakly equivalent to discrete spaces, they are equivalent to their sets of connected components, i.e. the functor $\mathcal{L}og_{2\text{-cat}} \to Ho(\mathcal{L}og_{2\text{-cat}})$ is fully faithful in the sense of simplicial categories and hence an equivalence.

The same reasoning goes through for $\mathcal{L}og^{\mathrm{strict}}_{2\text{-cat}}$. □

As we emphasized, homotopy (co)limits are almost never (co)limits in the homotopy category, but here this is the case.

Corollary 4.10 *Homotopy (co)limits in $\mathcal{L}og_{2\text{-cat}}$ (resp., $\mathcal{L}og^{\mathrm{strict}}_{2\text{-cat}}$) are (co)limits in $Ho(\mathcal{L}og_{2\text{-cat}})$ (resp., $Ho(\mathcal{L}og^{\mathrm{strict}}_{2\text{-cat}})$).*

Proof Homotopy limits of discrete spaces, i.e. sets, are discrete again and are their limits in the category of sets: The inclusion of the sub-$(\infty, 1)$-category of discrete spaces into the

$(\infty, 1)$-category of all spaces is right adjoint to the functor π_0 which takes a space to its set of connected components—therefore, it preserves homotopy limits, see [40, 5.5.6.5].

Given a diagram $\mathbf{D}$ in $\mathcal{L}og_{2\text{-cat}}$, its homotopy limit was defined through the weak equivalence of mapping spaces $\operatorname{holim}_{\mathbf{D}} \operatorname{map}_{Hilb_{2\text{-cat}}}(-, d) \simeq \operatorname{map}_{\mathcal{L}og_{2\text{-cat}}}(-, \operatorname{holim}_{\mathbf{D}} d)$ where the left hand homotopy limit was taken in spaces. Now we have the chain of equivalences

$$
\begin{aligned}
\lim_{\mathbf{D}} \operatorname{Hom}_{Ho(\mathcal{L}og_{2\text{-cat}})}(-, d) &\simeq \operatorname{holim}_{\mathbf{D}} \operatorname{Hom}_{Ho(\mathcal{L}og_{2\text{-cat}})}(-, d) \\
&\simeq \operatorname{holim}_{\mathbf{D}} \pi_0 \operatorname{map}_{\mathcal{L}og_{2\text{-cat}}}(-, d) \\
&\simeq \operatorname{holim}_{\mathbf{D}} \operatorname{map}_{\mathcal{L}og_{2\text{-cat}}}(-, d) \\
&\simeq \operatorname{map}_{\mathcal{L}og_{2\text{-cat}}}(-, \operatorname{holim}_{\mathbf{D}} d) \\
&\simeq \pi_0 \operatorname{map}_{\mathcal{L}og_{2\text{-cat}}}(-, \operatorname{holim}_{\mathbf{D}} d) \\
&= \operatorname{Hom}_{Ho(\mathcal{L}og_{2\text{-cat}})}(-, \operatorname{holim}_{\mathbf{D}} d).
\end{aligned}
$$

Since this is a weak equivalence of discrete spaces, it is a bijection of sets and hence identifies $\operatorname{holim}_{\mathbf{D}} d$ as the limit of the diagram in $Ho(\mathcal{L}og_{2\text{-cat}})$.

The proof for homotopy colimits is completely analogous (note that homotopy *colimits* in $\mathcal{L}og_{2\text{-cat}}$ turn to homotopy *limits* of mapping spaces when mapping out of them, so the same remarks about homotopy limits of discrete spaces apply). □

In the following two statements, we notice that homotopy equivalences in $\mathcal{L}og$ and $\mathcal{L}og^{\text{strict}}$ are characterized by the fact that they induce equivalences of mapping spaces.

Lemma 4.11 *A homotopy equivalence $z\colon X \to Y$ in $\mathcal{L}og$ (resp., $\mathcal{L}og^{\text{strict}}$) induces an equivalence of mapping spaces $z_*\colon \operatorname{map}(A, X) \to \operatorname{map}(A, Y)$*

Proof The homotopy equivalence z has a homotopy inverse $z'\colon Y \to X$ with $z \circ z' \dashv\vdash id$ and $z' \circ z \dashv\vdash id$; this means literally that z'_* becomes an inverse after applying π_0. Since the mapping spaces of $\mathcal{L}og_{2\text{-cat}}$ (resp., $\mathcal{L}og^{\text{strict}}$) are homotopy discrete, this already means that the map z_* is a weak equivalence. □

Proposition 4.12 *Let $f\colon L \to L'$ be a morphism of logics such that for all logics H the induced map $f_* := (f \circ -)\colon \operatorname{map}(H, L) \to \operatorname{map}(H, L')$ is a weak equivalence of mapping spaces. Then f is a homotopy equivalence of logics.*

Proof From the hypothesis in particular, we get a weak equivalence $f_*\colon \operatorname{map}(L', L) \to \operatorname{map}(L', L')$, hence the identity morphism $id_{L'}$ is in the connected component of some morphism in the image, i.e. there is a morphism $g\colon L' \to L$ such that $f \circ g\colon L' \to L \to L'$ satisfies $(f \circ g) \dashv\vdash id_{L'}$, i.e. we get a left inverse up to homotopy g. We will show that we also have $(g \circ f) \dashv\vdash id_L$ and hence f is a homotopy equivalence.

For every logic H there is the diagram

$$\begin{array}{ccc} \mathrm{map}\big(H, L'\big) & \xrightarrow{g_*} & \mathrm{map}(H, L) \\ & \underset{(f\circ g)_*}{\overset{\sim}{\searrow}} & \sim \big\downarrow f_* \\ & & \mathrm{map}\big(H, L'\big) \end{array}$$

Since two of the three arrows are weak equivalences, so is the third, hence $g\colon L' \to L$ satisfies the hypothesis of the proposition and, by what we have already proved, we get a left inverse $h\colon L \to L'$ to g with $g \circ h \dashv\vdash id_L$.

Now we know $f = f \circ id_L \dashv\vdash f \circ g \circ h \dashv\vdash id_{L'} \circ h = h$, and hence f is a left and right homotopy inverse. □

For a category $\mathcal{L}og$ of idempotent, substitution invariant logics, we know that weak equivalences and homotopy equivalences coincide, and that hence weak equivalences can be detected on mapping spaces. In homotopy theoretical terms, this can be seen as another incarnation of the fact that all objects of such a category $\mathcal{L}og$ are fibrant and cofibrant.

Remark 4.13 On the category $\mathcal{L}og^{\mathrm{strict}}$ the 2-categorical notion of equivalence is that of homotopy equivalence. We also have the notion of weak equivalence and there are strictly more weak equivalences than homotopy equivalences. Weak equivalences in general do not induce equivalences of mapping spaces as we can once again see from the example of the two presentations of classical propositional logic CPL_1 with underlying signature $\{\wedge, \neg\}$ and CPL_2 with underlying signature $\{\wedge, \neg, \vee, \to\}$: $\mathrm{map}^{\mathrm{strict}}(CPL_2, CPL_1) \to \mathrm{map}(CPL_2, CPL_1)$ is not surjective on connected components, since there is no strict translation equivalent to the flexible translations that are equivalences $CPL_2 \to CPL_1$. However, every logic L is weakly equivalent in $\mathcal{L}og^{\mathrm{strict}}$ to a logic $Q(L)$ such that weak equivalences into $Q(L)$ induce equivalences of mapping spaces, see Lemma 4.19 below.

4.3.1 $\mathcal{L}og^{\mathrm{strict}}_{\mathrm{2\text{-}cat}}$ Versus $\mathcal{L}og_{\mathrm{2\text{-}cat}}$

Convention 4.14 From now on we will suppose that the logics of $\mathcal{L}og$ have the property of idempotence.

We will relate the simplicial categories $\mathcal{L}og^{\mathrm{strict}}_{\mathrm{2\text{-}cat}}$ and $\mathcal{L}og_{\mathrm{2\text{-}cat}}$ via the adjunction of Proposition 2.6. First, we need to extend these functors from signatures to the corresponding categories of logics.

Let $L = (S_L, \vdash_L)$ be a logic. Recall from Definition 2.5 that the signature $Q(S_L)$ is defined by $Q(S_L)_n := \{c_\varphi \mid \varphi \in \mathrm{Fm}(S_L)[x_1, \ldots, x_n]\} \cong \mathrm{Fm}(S_L)[x_1, \ldots, x_n]$. We have an inclusion of signatures $s\colon S_L \to Q(S_L)$, $g \mapsto g(x_1, \ldots x_n)$ given by considering the old n-ary generating connectives g of S_L as formulas $g(x_1, \ldots, x_n)$ in $\mathrm{Fm}(S_L)[x_1, \ldots, x_n]$. This signature morphism induces an inclusion of sets of formulas $s\colon \mathrm{Fm}(S_L) \to \mathrm{Fm}(Q(S_L))$.

Definition 4.15 Let $L = (S_L, \vdash_L)$ be a logic. We define $Q(L)$ to be the logic over the signature $Q(S_L)$ with the consequence relation generated by $s_*(\vdash_L)$ (i.e. the rules of L

imported via s) and the rules $\{\psi \dashv\vdash \varphi\}$ for every pair of formulas which arise from each other by replacing connectives of the form c_φ with the corresponding formulas φ or vice versa.

Remark 4.16 If the logics in our category $\mathcal{L}og$ are substitution invariant and consequential, then it is enough to take the consequence relation generated by the rules $s_*(\vdash_L)$ and $\{c_\varphi(x_1, \ldots, x_n) \dashv\vdash \varphi(x_1, \ldots, x_n)\}$. The rules for more complex formulas from Definition 4.15 then become derivable by substitution invariance and congruentiality.

Lemma 4.17 *In $Q(L)$, every formula $\varphi(x_1, \ldots, x_n)$ is equivalent to a formula $c(x_1, \ldots, x_n)$ where c is a generating connective.*

Proof By definition of the consequence relation on $Q(L)$, every formula of $Q(L)$ is logically equivalent to a formula φ of $s(\mathrm{Fm}(L)) \subseteq \mathrm{Fm}(Q(S_L))$, obtained by replacing all occurrences of the new connectives c_ψ with ψ. But this formula $\varphi \in \mathrm{Fm}(L)$ is itself logically equivalent to the generating connective c_φ of $Q(L)$. □

Lemma 4.18 *There are flexible homotopy equivalences $r : Q(L) \leftrightarrows L : s$.*

Proof The inclusion $s\colon L \to Q(L)$ is a homotopy equivalence with homotopy inverse $r\colon Q(L) \to L$ given by sending the connectives of S_L to themselves and the connective c_φ to φ. Thus r takes a formula and replaces every occurrence of a connective c_φ by the corresponding formula φ. Clearly, s respects the consequence relation. To see the same for r, note that, for $\Gamma \cup \{\varphi\} \subseteq \mathrm{Fm}(Q(L))$ satisfying $\Gamma \vdash \gamma$, we have $r(\Gamma) \dashv\vdash \Gamma \vdash \varphi \dashv\vdash r(\varphi)$, hence by idempotence $r(\Gamma) \vdash r(\varphi)$. The composition $r \circ s$ is the identity and $(s \circ r)(\varphi)$ is logically equivalent to φ by definition of the consequence relation on $Q(L)$. □

Lemma 4.19 *Let H, L be logics. The inclusion $\mathrm{map}^{\mathrm{strict}}(H, Q(L)) \to \mathrm{map}(H, Q(L))$ is an equivalence of simplicial sets.*

Proof By Lemma 4.17, every n-ary formula φ of $Q(L)$ is equivalent to the formula $c_\varphi(x_1, \ldots, x_n)$. Hence a (flexible) morphism $f\colon H \to L$ is homotopic to the morphism $\tilde{f}\colon H \to L$ defined on generating connectives by $c \mapsto c_{f(c)}(x_1, \ldots, x_n)$, which is a strict morphism. Therefore, the map in question becomes a bijection after applying π_0. This is enough since the mapping spaces are homotopy discrete. □

Our aim in this subsection is to show Theorem 4.26, saying that logics and flexible morphisms form a reflexive sub-$(\infty, 1)$-category of logics and strict morphisms. A crucial step is the following theorem of Mariano and Mendes:

Theorem 4.20 ([42, Theorem 2.12, Mariano/Mendes]) *The adjunction of signatures of Theorem 2.6 lifts to an adjunction $i : \mathcal{L}og^{\mathrm{strict}} \rightleftarrows \mathcal{L}og : Q$ of logics.*

Since this adjunction respects homotopies between logics, it lifts further to an adjunction of $(\infty, 1)$-categories:

Proposition 4.21 *The functors $i : \mathcal{L}og^{\mathrm{strict}}_{2\text{-cat}} \rightleftarrows \mathcal{L}og_{2\text{-cat}} : Q$ form an adjunction of $(\infty, 1)$-categories.*

Proof It is clear that the maps of sets $\mathrm{Hom}_{\mathcal{S}ig^{\mathrm{strict}}}(S, S') \to \mathrm{Hom}_{\mathcal{S}ig}(iS, iS')$ and $\mathrm{Hom}_{\mathcal{S}ig}(S, S') \to \mathrm{Hom}_{\mathcal{S}ig^{\mathrm{strict}}}(Q(S), Q(S'))$ obtained from the functors i and Q of Definition 2.5 map translations to translations.

From Theorem 2.6, we have an adjunction on the level of signatures and from this the natural isomorphism $\mathrm{Hom}_{\mathcal{S}ig}(i(S), S') \to \mathrm{Hom}_{\mathcal{S}ig^{\mathrm{strict}}}(S, Q(S'))$ which sends a flexible morphism $(f_n : S_n \to \mathrm{Fm}(S')[x_1, \ldots, x_n])_{n \in \mathbb{N}}$ to itself (but becoming a strict morphism to $Q(S')$), in other words it is given by postcomposition with the map $S' \to Q(S')$.

By Theorem 4.20, this lifts to an adjunction $i : \mathcal{L}og^{\mathrm{strict}} \rightleftarrows \mathcal{L}og : Q$ of logics, i.e. the isomorphism restricts to an isomorphism $\mathrm{Hom}_{\mathcal{L}og}(i(L), L') \to \mathrm{Hom}_{\mathcal{L}og^{\mathrm{strict}}}(L, Q(L'))$ between the sets of translations: Indeed, by Lemma 4.18, the morphism of logics $L \to Q(L)$ is a homotopy equivalence in $\mathcal{L}og$, hence by idempotence and Lemma 4.2 a weak equivalence and in particular conservative, and so a morphism of signatures $iS \to S'$ is a translation if and only if $iS \to S' \to Q(S')$ is.

Since homotopic translations get mapped to homotopic translations, this extends to a morphism $\mathrm{map}(iL, L') \to \mathrm{map}^{\mathrm{strict}}(L, Q(L'))$ of mapping spaces which can be seen to be an equivalence from the diagram

$$\begin{array}{ccc}
\mathrm{map}(iL, L') & \xrightarrow{s_*} & \mathrm{map}^{\mathrm{strict}}(L, Q(L')) \\
\Vert & & \downarrow \sim \\
\mathrm{map}(L, L') & \underset{s_*}{\xrightarrow{\sim}} & \mathrm{map}(L, Q(L'))
\end{array}$$

Here the right vertical arrow is the equivalence of Lemma 4.19 and the lower horizontal arrow is an equivalence since the equivalence $s : L \to Q(L)$ of Lemma 4.18 induces an equivalence on mapping spaces by Lemma 4.11. Therefore, the upper horizontal arrow must be an equivalence, too. □

Lemma 4.22 *The map of simplicial sets* $\mathrm{map}(Q(H), Q(L)) \to \mathrm{map}(H, Q(L))$, $f \mapsto f \circ (H \to Q(H))$ *is an equivalence.*

Proof Again it is enough to show that applying π_0 induces a bijection. (Injectivity) A morphism $Q(H) \to Q(L)$ is, up to equivalence, determined by its restriction to $H \subseteq Q(H)$, since every formula of $Q(H)$ is equivalent to one of H. (Surjectivity) Any morphism $f : H \to Q(L)$ can be extended to $Q(H)$ by sending the additional generating connectives c_φ ($\varphi \in \mathrm{Fm}(H)$) to $f(\varphi)$. □

We denote by $Q(\mathcal{L}og)$ the image of the functor Q in $\mathcal{L}og^{\mathrm{strict}}$ and we will for the rest of the subsection suppress the subscript "2-cat".

Lemma 4.23 *The functor* $Q \circ i\,|_{Q(\mathcal{L}og)} : Q(\mathcal{L}og) \to Q(\mathcal{L}og)$ *is an equivalence of* $(\infty, 1)$*-categories.*

Proof (Essential surjectivity) We have to show that any object in the target category $Q(\mathcal{L}og)$ is equivalent to one in the image, i.e. one of the form $Q(Q(L))$. This is the case because the unit map $Q(L) \to Q(Q(L))$ is a weak equivalence of logics in $\mathcal{L}og^{\text{strict}}$ by Lemma 4.18, idempotence and Lemma 4.2.

(Full faithfulness) We have to show that the morphism of simplicial sets $\text{map}^{\text{strict}}(Q(H), Q(L)) \to \text{map}^{\text{strict}}(QQ(H), QQ(L))$ is an equivalence. We know from Lemma 4.19 that this is equivalent to showing that we have an equivalence of flexible mapping spaces $\text{map}(Q(H), Q(L)) \to \text{map}(QQ(H), QQ(L))$. We have a commutative diagram

$$\begin{array}{ccc} \text{map}\big(Q(H), Q(L)\big) & \longrightarrow & \text{map}\big(QQ(H), QQ(L)\big) \\ & \underset{\sim}{\searrow} & \downarrow{\scriptstyle \sim} \\ & & \text{map}\big(Q(H), QQ(L)\big) \end{array}$$

where the vertical arrow is the equivalence of Lemma 4.22 and the diagonal arrow is induced by the homotopy equivalence s of Lemma 4.18 and hence an equivalence by Lemma 4.11. Since two of the three arrows are equivalences, so is the third. □

Proposition 4.24 *The inclusion functor $i\,|_{Q(\mathcal{L}og)}\colon Q(\mathcal{L}og) \to \mathcal{L}og$ is an equivalence of simplicial categories.*

Proof (Essential surjectivity) This is the fact (Lemma 4.18) that $L \to Q(L)$ is an equivalence of logics. (Full faithfulness) This is Lemma 4.19. □

Lemma 4.25 *The category $Q(\mathcal{L}og)$ is a reflective sub-$(\infty, 1)$-category of $\mathcal{L}og^{\text{strict}}$.*

Proof We know from Proposition 4.21 combined with Proposition 4.24 that the inclusion of $Q(\mathcal{L}og)$ into $\mathcal{L}og^{\text{strict}}$ has a left adjoint, and from Lemma 4.23 that their composition is an idempotent functor. This is what defines a reflexive subcategory. □

Theorem 4.26 *$\mathcal{L}og$ is (equivalent to) a reflective sub-$(\infty, 1)$-category of $\mathcal{L}og^{\text{strict}}$ via the reflection functor Q.*

Proof This is Lemma 4.25 combined with Proposition 4.24. □

4.3.2 Homotopy Limits in $\mathcal{L}og_{\text{2-cat}}$

We will now show how homotopical thinking can lead one to the construction of homotopy limits in $\mathcal{L}og_{\text{2-cat}}$. By Corollary 4.10, homotopy limits are limits in $Ho(\mathcal{L}og_{\text{2-cat}})$ and the existence of these (for congruential Hilbert systems) has been established by Mariano and Mendes in [42, Theorem 2.33].

Here we will give direct constructions of several kinds of homotopy limits to show some of the homotopy theoretical flavor. For example, in our construction of homotopy equalizers, there occur logics resembling the path spaces of the corresponding constructions in topology. This shows how signatures can be tailored to fulfill the needs of particular constructions while keeping them as small as possible. A more general approach would be to use always the construction $Q(-)$ of Definition 2.5.

We start with the easiest case.

Homotopy Terminal Objects. We have seen in Example 2.7 that the category $\mathcal{S}ig$ has no terminal object, hence by Proposition 2.24 neither does $\mathcal{L}og_{2\text{-cat}}$. In the homotopical world, things look better:

Proposition 4.27 *The category $\mathcal{L}og_{2\text{-cat}}$ has homotopy terminal objects.*

Proof By Corollary 4.10, it is enough to show that there is a terminal object in $Ho(\mathcal{L}og_{2\text{-cat}})$. For this take the signature S which has one generating connective of each arity (or any other signature which produces formulas of any arity) and endow it with the maximal consequence relation. Clearly, for any logic L there is a translation $L \to (S, \vdash_{\max})$ given on signatures by mapping each n-ary generating connective to the n-ary generating connective of S. Since in $(S, \vdash_{\max})$ any two formulas are logically equivalent, all morphisms into $(S, \vdash_{\max})$ are equal in the homotopy category. □

The reader who wishes to test this statement for another category of logics satisfying the assumptions of Remark 2.20, should make sure that for the maximal consequence relation actually all formulas are equivalent. This is certainly true for those of Convention 2.19.

Homotopy Equalizers. Given two parallel arrows $f, g\colon L \rightrightarrows M$ in $\mathcal{L}og_{2\text{-cat}}$, their homotopy equalizer should be a morphism $e\colon E \to L$ such that $(f \circ e)(\varphi) \dashv\vdash_M (g \circ e)(\varphi)$ (i.e. instead of asking for equality, we ask for pointwise logical equivalence) and such that each $h\colon H \to L$ with $(f \circ h)(\varphi) \dashv\vdash_M (g \circ h)(\varphi)$ factorizes uniquely up to homotopy through e (this is usually not equivalent to the homotopy limit condition expressed through mapping spaces, but here it is since the mapping spaces are discrete).

Thus we would like to simply take as equalizer the set of formulas $\{\varphi \in \mathrm{Fm}(L) \mid f(\varphi) \dashv\vdash_M g(\varphi)\}$ and endow it with the consequence relation of L restricted to this subset. But this set of formulas is in general not a free algebra over some signature. On the other hand, if one tries to take as generating connectives of the homotopy equalizer just those c which satisfy $f(c(x_1, \dots, x_n)) \dashv\vdash_M g(c(x_1, \dots, x_n))$ then one might miss formulas φ with $f(\varphi) \dashv\vdash_M g(\varphi)$. However, since homotopy equalizers are invariant under equivalences, we can substitute L by an equivalent logic for which the second solution works.

Definition 4.28 Let $L = (S_L, \vdash_L)$ be a logic, $f, g\colon L \rightrightarrows M$ be two morphisms in $\mathcal{L}og_{2\text{-cat}}$.

1. The signature $S^{(f,g)}$ is defined by

$$S_n^{(f,g)} := (S_L)_n \coprod \{c_\varphi \mid \varphi \in \mathrm{Fm}(L),\ f(\varphi) \dashv\vdash_M g(\varphi)\},$$

   i.e. by taking as generating connectives those of S_L plus one extra n-ary connective c_φ for every $\varphi \in \mathrm{Fm}(L)[n]$ whose images under f and g are logically equivalent.
2. The logic $L^{(f,g)}$ is the logic over the signature $S^{(f,g)}$, defined by endowing $\mathrm{Fm}(S^{(f,g)})$ with the consequence relation generated by $\vdash_L$ (for the formulas of the linguistic fragment generated by S_L) and by the rules $\{\psi \dashv\vdash \varphi\}$ for every pair of formulas which

arise from each other by replacing connectives of the form c_φ with the corresponding formulas φ, or vice versa.

Thus the logic $L^{(f,g)}$ contains a copy of L as well as lots of new generating connectives which are equivalent to formulas of L which become logically equivalent under f (resp., g). As before, for substitution invariant, congruential logics it is enough to demand as generating rules the rules $\{c_\varphi \dashv\vdash \varphi \mid \varphi \in \mathrm{Fm}(S_L)\}$.

Lemma 4.29 *The inclusion $i\colon L \to L^{(f,g)}$ is a homotopy equivalence with homotopy inverse $r\colon L^{(f,g)} \to L$ given by sending the connectives of S_L to themselves and the connective c_φ to φ.*

Proof This is precisely parallel to Lemma 4.18 where the inclusion $L \to Q(L)$ was shown to be a homotopy equivalence. Both morphisms respect the consequence relation by definition of the consequence relation on $L^{(f,g)}$. The composition $r \circ i$ is the identity and $i \circ r$ is the identity on $\mathrm{Fm}(L) \subseteq \mathrm{Fm}(L^{(f,g)})$ and maps the formula $c_\varphi(x_1, \ldots, x_n)$ to φ, which is logically equivalent. □

Definition 4.30 Define the signature S_E as the subsignature of $S^{(f,g)}$ given by the connectives $\{c_\varphi \mid \varphi \in \mathrm{Fm}(L),\ f(\varphi) \dashv\vdash_M g(\varphi)\}$. Then the logic $E^{(f,g)}$ is the logic over the signature S_E endowed with the strongest consequence relation such that the inclusion $S_E \to S^{(f,g)}$ becomes a translation. We denote the inclusion by $e\colon E^{(f,g)} \to L^{(f,g)}$.

Proposition 4.31 *The diagram*

$$E^{(f,g)} \xrightarrow{r \circ e} L \underset{g}{\overset{f}{\rightrightarrows}} M$$

is a homotopy equalizer diagram in $\mathcal{L}og_{2\text{-cat}}$.

Proof We will directly show the criterion for mapping spaces, i.e. for any logic H the map of (homotopy discrete) simplicial sets

$$\mathrm{map}(H, E^{(f,g)}) \xrightarrow{(r \circ e)_*} hoequ(\mathrm{map}(H, L) \underset{g_*}{\overset{f_*}{\rightrightarrows}} \mathrm{map}(H, M))$$

is a weak equivalence.

Since, by Lemma 4.11, a homotopy equivalence $z\colon X \to Y$ in $\mathcal{L}og$ induces an equivalence of mapping spaces, by Lemma 4.29, we have a weak equivalence $\mathrm{map}(H, L) \leftrightarrows \mathrm{map}(H, L^{(f,g)})$. Also, by Proposition 4.9, we have equivalences $\mathrm{map}(H, L^{(f,g)}) \to$

$\pi_0\,\mathrm{map}(H, L^{(f,g)})$, and hence the following equivalences of diagrams of the shape $\bullet \rightrightarrows \bullet$:

$$\begin{array}{ccc}
\mathrm{map}(H, L) & \overset{f_*}{\underset{g_*}{\rightrightarrows}} & \mathrm{map}(H, M) \\
\simeq \downarrow & & \downarrow \simeq \\
\mathrm{map}\big(H, L^{(f,g)}\big) & \overset{f_*}{\underset{g_*}{\rightrightarrows}} & \mathrm{map}(H, M) \\
\simeq \downarrow & & \downarrow \simeq \\
\pi_0\,\mathrm{map}\big(H, L^{(f,g)}\big) & \overset{f_*}{\underset{g_*}{\rightrightarrows}} & \pi_0\,\mathrm{map}(H, M)
\end{array}$$

Since homotopy equalizers do not change up to equivalence upon replacing a diagram by an equivalent one, and the homotopy equalizer of a diagram of sets is just the equalizer, we have the following equivalences:

$$\begin{aligned}
&\mathit{hoequ}\Big(\mathrm{map}(H, L) \overset{f_*}{\underset{g_*}{\rightrightarrows}} \mathrm{map}(H, M)\Big) \\
&\simeq \mathit{equ}\Big(\pi_0\,\mathrm{map}(H, L^{(f,g)}) \overset{f_*}{\underset{g_*}{\rightrightarrows}} \pi_0\,\mathrm{map}(H, M)\Big) \\
&\simeq \Big\{\overline{h} \in \pi_0\,\mathrm{map}\big(H, L^{(f,g)}\big) \mid \overline{f} \circ \overline{h} = \overline{f} \circ \overline{h}\Big\} \\
&\simeq \Big\{\overline{h} \in \pi_0\,\mathrm{map}\big(H, L^{(f,g)}\big) \mid f \circ h \dashv\vdash g \circ h\Big\}
\end{aligned}$$

where $\overline{(-)}$ denotes the equivalence class of a morphism in the quotient $\pi_0\,\mathrm{map}(-, -)$.

The map $\mathrm{map}(H, E^{(f,g)}) \to \{\overline{h} \in \pi_0\,\mathrm{map}(H, L^{(f,g)}) \mid f \circ h \dashv\vdash g \circ h\}$, since it goes into a set, factors through $\pi_0\,\mathrm{map}(H, E^{(f,g)})$, and we have to show that $\pi_0\,\mathrm{map}(H, E^{(f,g)}) \to \{\overline{h} \in \pi_0\,\mathrm{map}(H, L^{(f,g)}) \mid f \circ h \dashv\vdash g \circ h\}$ is a bijection.

(Surjectivity) Every $h \in \mathrm{Hom}_{\mathcal{L}og}(H, L^{(f,g)})$ such that $f \circ h \dashv\vdash g \circ h$ is homotopic to a morphism going into the linguistic fragment of E: For each connective $c \in S_H$ we have $(f \circ h)(c(x_1, \ldots, x_n)) \dashv\vdash (g \circ h)(c(x_1, \ldots, x_n))$, and thus h is homotopic to $h' \colon H \to L^{(f,g)}$ defined by $c \mapsto c_{h(c)} \in S_{E^{(f,g)}} \subseteq S_{L^{(f,g)}}$.

(Injectivity) If $h, h' \colon H \to E^{(f,g)}$ go after composition with e to the same morphism $(H \to E^{(f,g)} \to L^{(f,g)}) \in \pi_0\,\mathrm{map}(H, L^{(f,g)})$, then $(e \circ h)(\varphi) \dashv\vdash_{L^{(f,g)}} (e \circ h')(\varphi)$ for all $\varphi \in \mathrm{Fm}(H)$. The fact that $\vdash_{E^{(f,g)}}$ is the restriction of $\vdash_{L^{(f,g)}}$ to $\mathrm{Fm}(S_{E^{(f,g)}})$ means that e is conservative, so $h(\varphi) \dashv\vdash_{E^{(f,g)}} h'(\varphi)$ for all $\varphi \in \mathrm{Fm}(H)$, hence $\overline{h} = \overline{h'} \in \pi_0\,\mathrm{map}(H, E^{(f,g)})$. ☐

We still note that the homotopy equalizer can be characterized by $\mathrm{map}(H, E^{(f,g)}) \simeq \pi_0\,\mathrm{map}(H, E^{(f,g)}) \cong \{\overline{h} \in \pi_0\,\mathrm{map}(H, L^{(f,g)}) \mid f \circ h \dashv\vdash g \circ h\} \simeq \{\overline{h} \in \pi_0\,\mathrm{map}(H, L) \mid f \circ h \dashv\vdash g \circ h\}$, where the latter bijection comes from the fact that every translation $H \to L^{(f,g)}$ is equivalent to one going into L, seen as a sublogic of $L^{(f,g)}$.

In the construction of homotopy equalizers, we could have used $Q(L)$ instead of $L^{(f,g)}$ with almost the same proofs, but we thought it to be instructive to show how one can construct a logic tailored to the problem at hand. If we demand more properties from the logics of the category $\mathcal{L}og$, these specialized logics can become much smaller than the universal solution via $Q(L)$.

One can construct homotopy pullbacks in an entirely similar way to the construction of homotopy equalizers given here. We leave this to reader, but the existence of homotopy pullbacks follows from Theorem 4.39 below which asserts the homotopy cocompleteness of $\mathcal{L}og_{2\text{-cat}}$.

Homotopy Products. Given a family of logics $(L_i = (S_i, \vdash_i) \mid i \in I)$, a first tentative construction of the product logic might be as follows: Take as signature $S := \prod S_i$, so that the generating connectives are tuples of generating connectives from the S_i. This signature has projection maps $pr_i \colon S \to S_i$, defined on generating connectives by $(c_i)_{i \in I} \mapsto c_i$. Then define a logic $\prod_i L_i$ by endowing $\mathrm{Fm}(S)$ with the strongest consequence relation such that all these projection maps are translations—this is the consequence relation given by $(\Gamma_i) \vdash (\varphi_i) \Leftrightarrow \forall i \in I \; \Gamma_i \vdash_{L_i} \varphi_i$.

We would now have to show that $\langle pr_i \mid i \in I \rangle \colon \pi_0 \,\mathrm{map}(H, \prod_i L_i) \to \prod_i \pi_0 \,\mathrm{map}(H, L_i)$ is a bijection, but surjectivity in general fails, as seen by the following example:

Example 4.32 Take all the L_i to be the same logic $L = (S, \vdash)$ with signature S generated by a single unary connective $\Box$ and the rule $\Box x \vdash x$. Then $\mathrm{Fm}(L) = \{\Box^i x_k \mid i, k \in \mathbb{N}\}$, where $\Box^i := \Box\Box \cdots \Box$ (i times). This logic has the feature that no two formulas are equivalent, unless one can be obtained from the other by a substitution of variables with other variables. The product $\prod_{n \in \mathbb{N}} \pi_0 \,\mathrm{map}(L, L)$ contains the family (f_n) with $f_n \colon L \to L$ given by mapping $\Box \mapsto \Box^n$. Now if $\langle pr_i \mid i \in I \rangle \colon \pi_0 \,\mathrm{map}(H, \prod_i L_i) \to \prod_i \pi_0 \,\mathrm{map}(H, L_i)$ were surjective, there would have to be (up to equivalence) a formula in $\prod_{n \in \mathbb{N}} L$ of the form $(\Box^n x)_{n \in \mathbb{N}}$. This can not be the case since our tentatively defined product signature is generated by the single connective $(\Box)_{n \in \mathbb{N}}$. Even if we allow a unary "identity connective" which does not change a formula when it is applied, our product signature would be generated by $\{(\triangle_n)_{n \in \mathbb{N}} \mid \triangle_n \in \{id, \Box\}\}$, i.e. by tuples which in each place have either the identity connective or $\Box$. Formulas of this signature would, as usual, be finite combinations of the generating connectives and any such connective would have a highest n such that a $\Box^n$ occurs in some place of the tuple.

The reason behind this failure is that products of free algebras are not in general free again, as already remarked in Sect. 2. The solution to the problem, much as in the case of homotopy equalizers, is to substitute the logics L_i whose homotopy product we want to form, by the equivalent logics $Q(L_i)$ which have the feature that every formula $\varphi(x_1, \ldots, x_n)$ is equivalent to a formula $c(x_1, \ldots, x_n)$ where c is a generating connective.

Definition 4.33 Let $L_i = (S_{L_i}, \vdash_{L_i}), i \in I$ be a family of logics. We define $\prod S_{L_i}$ to be the signature with $(\prod S_{L_i})_n := \{(c_i)_{i \in I} \mid c_i \in (S_{L_i})_n\}$, i.e. the signature whose n-ary generating connectives are tuples of n-ary generating connectives of the S_{L_i}. There are obvious projection maps $pr_i \colon \prod S_{L_i} \to S_{L_i}$. We define the consequence relation $\vdash_{\prod L_i}$ over $\prod S_{L_i}$ to be the biggest consequence relation such that all projection morphisms become translations. Finally we define the product logic to be $\prod L_i = (\prod S_{L_i}, \vdash_{\prod L_i})$.

Remark 4.34 The logic $\prod L_i$ is the product of the logics L_i in the category $\mathcal{L}og^{\text{strict}}$, by the construction recipe given in Proposition 2.24.

Temporary Convention 4.35 In the next two statements, we establish the existence of homotopy products in special categories $\mathcal{L}og$. We make the distinction between the following two cases:

(A) $\mathcal{L}og$ is a category of logics where infima of consequence relations are given by intersection. This means: If $\vdash_i$ are consequence relations over $\mathrm{Fm}(S)$ which are admissible for $\mathcal{L}og$ (i.e. $(S, \vdash_i)$ are objects of $\mathcal{L}og$), then their intersection $(\bigcap_i \vdash_i) \in \mathscr{P}(\mathscr{P}(\mathrm{Fm}(S)) \times \mathrm{Fm}(S))$ is also a consequence relation admissible for $\mathcal{L}og$.

(B) $\mathcal{L}og$ is the full subcategory of one of the categories from A, given by logics which additionally are finitary.

Thus categories $\mathcal{L}og$ of type A include, for example, $\mathcal{L}og^{(\text{Tarsk})}$ and $\mathcal{L}og^{(\text{subst,Tarsk})}$ and categories $\mathcal{L}og$ of type B include $\mathcal{L}og^{(\text{fin,Tarsk})}$ and $\mathcal{H}ilb$. We insert this digression for special categories, because this admits a nice concrete construction of homotopy products (only finite ones in case B).

The impatient reader can skip to Proposition 4.38 which guarantees the existence of homotopy products for general categories $\mathcal{L}og$, using results from abstract homotopy theory which will be sketched in Sect. 4.3.3.

Lemma 4.36 *Suppose that either the objects of $\mathcal{L}og$ are finitary logics (case B) and the family L_i of Definition 4.33 is finite, or that finitariness is not required for the objects of $\mathcal{L}og$ (case A). Then the consequence relation $\vdash_{\prod L_i}$ is given by $\Gamma \vdash_{\prod L_i} \varphi \;:\Leftrightarrow \forall i\, pr_i(\Gamma) \vdash_{L_i} pr_i(\varphi)$.*

Proof In the case (B) of a category of finitary logics $\mathcal{L}og$, the consequence relation was defined as the infimum in the complete lattice of finitary (and possibly substitution invariant, Tarskian, etc.) consequence relations of the inverse images $pr_i^{-1}(\vdash_{L_i})$. The inverse image relations are given by $\Gamma\, pr_i^{-1}(\vdash_{L_i})\varphi \Leftrightarrow pr_i(\Gamma) \vdash pr_i(\varphi)$.

By the proof of [2, Fact 4], this finitary infimum $[\mathit{inf}(\vdash_i)]$ of a family of consequence relations $\vdash_i$ is given by $\Gamma\, [\mathit{inf}(\vdash_i)]\, \varphi \Leftrightarrow \exists\, \Gamma' \subseteq_{\text{finite}} \Gamma$ such that $\forall i : \Gamma' \vdash_i \varphi$.

Since all the $\vdash_i$ are finitary, there exist $\Gamma_i \subseteq_{\text{finite}} \Gamma$ such that $\Gamma_i \vdash_i \varphi$. If the family is finite, then the union of these Γ_i is still finite, so that the existence of the $\Gamma' \subseteq_{\text{finite}} \Gamma$ is always ensured. This gives the claimed description of the product consequence relation.

In the case (A) where finitariness is not demanded from the objects of $\mathcal{L}og$, the infimum of a family of consequence relations is given by $\Gamma\, [\mathit{inf}(\vdash_i)]\, \varphi \Leftrightarrow \forall i : \Gamma' \vdash_i \varphi$. ☐

Proposition 4.37 *The category $\mathcal{L}og_{2\text{-cat}}$ has finite homotopy products, if the objects are demanded to be finitary (case B) and all homotopy products if the objects are not demanded to be finitary (case A).*

Proof Let $L_i = (S_{L_i}, \vdash_{L_i}), i \in I$ be a family of logics, which we suppose to be finite in the first case. We claim that the strict product $\prod Q(L_i)$ of the replaced logics $Q(L_i)$ is a homotopy product of this family. For this we need to show that for any H the map

$$\pi_0\, \mathrm{map}\Big(H, \prod Q(L_i)\Big) \to \prod_i \pi_0\, \mathrm{map}(H, L_i), \qquad \overline{f} \mapsto (\overline{pr_i \circ f})_{i \in I}$$

is a bijection.

(Surjectivity) From Lemma 4.18, we know that $\pi_0\,\mathrm{map}(H, L_i) \cong \pi_0\,\mathrm{map}(H, Q(L_i))$, so we can replace the target by $\pi_0\,\mathrm{map}(H, Q(L_i))$. Given a family $(\overline{f_i})_{i\in I} \in \prod_i \pi_0\,\mathrm{map}(H, Q(L_i))$, we know from Lemma 4.19 that the $\overline{f_i}$ have representatives f_i given by strict morphisms. Now we can define a preimage f of the family (f_i): If f_i sends a generating connective c of H to a generating connective c_i of $Q(L_i)$, then define f by $f(c) := (c_i)_{i\in I} \in \prod S_{L_i}$. This is clearly a translation and a preimage of the family (f_i).

(Injectivity) Suppose we have $\overline{f}, \overline{f'}$ such that $\overline{pr_i \circ f} = \overline{pr_i \circ f'}$. The latter condition means that $(pr_i \circ f)(\varphi) \dashv\vdash (pr_i \circ f')(\varphi)\ \forall\varphi$. By Lemma 4.36, this implies $f(\varphi) \dashv\vdash f'(\varphi)\ \forall\varphi$, i.e. $\overline{f} = \overline{f'}$. □

End of Temporary Convention. Now we return to our convention of denoting by $\mathcal{L}og$ any category of idempotent logics satisfying the assumptions of Remark 2.20.

Proposition 4.38 *The category $\mathcal{L}og_{2\text{-cat}}$ has all homotopy products.*

Proof This follows from the fact that $\mathcal{L}og_{2\text{-cat}}$ is a reflective $(\infty, 1)$-subcategory of $\mathcal{L}og^{\text{strict}}_{2\text{-cat}}$ and that $\mathcal{L}og^{\text{strict}}_{2\text{-cat}}$ has all homotopy products—the latter will be established in the next section by means of a model structure on $\mathcal{L}og^{\text{strict}}$. Indeed, our adjoint functors i and Q induce an adjunction on the homotopy categories and, by Corollary 4.10, it is enough to establish the existence of products in the homotopy category $Ho(\mathcal{L}og_{2\text{-cat}})$. Now it is an exercise in usual category theory that a reflective subcategory of a category with products has products itself. These products are given by forming the product in the ambient category and then applying the reflection functor. Thus the homotopy product of a family L_i is given by $Q(\prod_i L_i)$. □

Theorem 4.39 *The simplicial category $\mathcal{L}og_{2\text{-cat}}$ has all homotopy limits.*

Proof We have constructed homotopy products (including a homotopy terminal object) and homotopy equalizers. By the dual of [40, Proposition 4.4.3.2], one can build homotopy limits from homotopy equalizers and homotopy products, analogously to the non-homotopical statement in classical category theory. That the limits of the cited proposition really correspond to homotopy limits as explained in Sect. 3 is the content of [40, Proposition 4.2.4.1]. □

4.3.3 Homotopy Colimits in $\mathcal{L}og_{2\text{-cat}}$

It would be possible to construct homotopy colimits by hand as we did for homotopy limits. Instead, we sketch a proof of the existence of homotopy colimits by different means, to give a feeling for how homotopy theoretical machinery can be brought into play for solving such questions. For this we use Theorem 4.26, saying that $\mathcal{L}og_{2\text{-cat}}$ is a reflective sub-$(\infty, 1)$-category of $\mathcal{L}og^{\text{strict}}_{2\text{-cat}}$, and show that $\mathcal{L}og^{\text{strict}}_{2\text{-cat}}$ has all homotopy colimits. First, we invoke the following theorem:

Theorem 4.40 ([35, Theorem 3.3, S. Lack]) *Let $\mathcal{C}$ be a finitely complete and finitely cocomplete category enriched in the category* **Cat** *of categories. Then there is a model structure on $\mathcal{C}$, such that the weak equivalences are precisely the 2-categorical equivalences in the sense of Sect. 3.3.4. This model structure has the feature that every object is fibrant and cofibrant and is compatible with the enrichment.*

We can apply this to the category $\mathcal{L}og^{\text{strict}}$ enriched in the maximal subgroupoids of the usual Hom-preorders. This is not entirely trivial, as the (co)completeness condition of the theorem is to be understood in the enriched sense: Apart from the completeness and cocompleteness of $\mathcal{L}og^{\text{strict}}$, which we know from Proposition 2.24 (or [2, Proposition 2.11]) one has to show that $\mathcal{L}og^{\text{strict}}$ is tensored and cotensored, in the sense of [34], over the category **Cat** of categories. This can be done by techniques similar to those we used in the construction of homotopy limits in $\mathcal{L}og^{\text{strict}}_{\text{2-cat}}$.

From the theorem, we then get a **Cat**-enriched model category in the sense of [35, Sect. 2.2]. Recall that the $(\infty, 1)$-category corresponding to a model category is given by taking the subcategory of fibrant and cofibrant objects and applying the hammock localization. Here, since all objects are fibrant and cofibrant this is simply the hammock localization of the whole category $\mathcal{L}og^{\text{strict}}$ with respect to the homotopy equivalences. As the $(\infty, 1)$-category corresponding to a model category has all homotopy limits and colimits, we have established the homotopy (co)completeness of this hammock localization.

However, we need to know that the *2-categorical* localization $\mathcal{L}og^{\text{strict}}_{\text{2-cat}}$ is cocomplete, so we still have to relate this to the hammock localization. The notion of **Cat**-enriched model category is such that we get, when we apply the nerve functor to the hom-groupoids (and here it is important that they are groupoids), a *simplicial model category* in the sense of Quillen, see [29, Definition 4.2.18]. Now by [18, Proposition 4.8] for a simplicial model category, the simplicial subcategory of fibrant and cofibrant objects—which here is exactly $\mathcal{L}og^{\text{strict}}_{\text{2-cat}}$—is equivalent to the hammock localization considered before. Hence we conclude that the $(\infty, 1)$-category $\mathcal{L}og^{\text{strict}}_{\text{2-cat}}$ has all homotopy limits and colimits.

Finally, now we can now use Theorem 4.26, saying that $\mathcal{L}og_{\text{2-cat}}$ is a reflective sub-$(\infty, 1)$-category of $\mathcal{L}og^{\text{strict}}_{\text{2-cat}}$: The homotopy colimit of a diagram in $\mathcal{L}og_{\text{2-cat}}$ can be constructed by seeing it, via the functor Q, as a diagram in $\mathcal{L}og^{\text{strict}}_{\text{2-cat}}$, forming its homotopy colimit there, and then applying the reflection functor, which as a left adjoint preserves colimits.

A detailed elaboration of these arguments, and further exploration of the Lack model structure on $\mathcal{L}og^{\text{strict}}_{\text{2-cat}}$ will be given in a future work.

5 Vista

5.1 Further Studies of Categories of Logics

In the previous chapter, we gave two natural constructions of $(\infty, 1)$-categories of logics, and showed how to explore some of their properties with the examples of Hilbert systems. Many natural questions about $\mathcal{L}og_{\text{2-cat}}$ and $\mathcal{L}og_{\text{hamm}}$, and their strict versions, remain to be pursued. Most of these questions should not be hard to tackle, as we find ourselves in a rather easy region of the realm of abstract homotopy theory. The most immediate question is the following:

Question 5.1 Are the simplicial categories $\mathcal{L}og_{2\text{-cat}}$ and $\mathcal{L}og_{\text{hamm}}$ equivalent?

One approach to proving an equivalence is to simply write down an enriched functor between the two simplicial categories and prove it to be an equivalence. This would involve an analysis of the mapping spaces of the hammock localization, which would be interesting in its own right, as it might reveal criteria for determining whether two translations are homotopic. Other approaches could proceed by constructing appropriate models of the two $(\infty, 1)$-categories, which are more easily comparable than the simplicial categories. For example, one could try to find model categories presenting both $(\infty, 1)$-categories and produce a Quillen equivalence between them. This allows staying in the realm of usual categories. Candidates for such models appear below.

On the category $\mathcal{L}og^{\text{strict}}$, we have the two different notions of homotopy equivalence and weak equivalence and we can form the hammock localizations with respect to both of these notions of equivalence. From Lemmas 4.18 and 4.19, we know that every logic is weakly equivalent to one such that every translation into it is homotopic to a strict translation. So it is natural to ask:

Question 5.2 Are $L^H(\mathcal{L}og^{\text{strict}}, W \cap \mathcal{L}og^{\text{strict}})$ and $L^H(\mathcal{L}og, W) = \mathcal{L}og_{\text{hamm}}$ equivalent $(\infty, 1)$-categories (where W denotes the class of weak equivalences)?

Again one can either directly try to construct an equivalence of simplicial categories or approach the question via models. The second approach is related to the next question.

Recall from Sect. 4.3.3 that there is a model structure on $\mathcal{L}og^{\text{strict}}$ whose equivalences are the homotopy equivalences. One may ask if this model structure admits a *Bousfield localization*: A Bousfield localization of a model category is a new model structure on the same underlying category which has additional weak equivalences and fewer fibrations, see [40, A.3.7]. The $(\infty, 1)$-category presented by the Bousfield localization is a reflexive $(\infty, 1)$-subcategory of the $(\infty, 1)$-category presented by the original model structure.

Question 5.3 Is there a Bousfield localization of the Lack model structure on the category $\mathcal{H}ilb^{\text{strict}}$ whose weak equivalences are the weak equivalences of logics from Definition 4.1.3?

The construction of Bousfield localizations is, for example, available for the so-called combinatorial model categories by a general theorem of J. Smith, see [3]. These are model categories which are "cofibrantly generated" (see [40, Definition A.2.6.1(2–3)]) and whose underlying category is locally presentable. The latter is the case for $\mathcal{H}ilb^{\text{strict}}$ by [2, Theorem 2.16].

Here is a further candidate category for a model for the $(\infty, 1)$-category $\mathcal{L}og_{\text{hamm}}$. In several places in the literature, proposals have appeared to consider logics not just on absolutely free algebras, but to instead consider consequence relations on arbitrary algebras—note that the definition of consequence relation, Definition 2.11, does not have to be altered for this to make sense. These gadgets have been called *abstract logics* in [9]. An advantage of allowing non-free algebras is that one has a better behaved, e.g. complete and cocomplete, category, which is, for example, the reason that they appeared in the context of fibring of institutions in [11]. A disadvantage is that one introduces objects which one would not commonly perceive as logics.

Question 5.4 Is there a model structure on abstract logics presenting the $(\infty, 1)$-category $\mathcal{L}og_{\mathrm{hamm}}$?

The (co)completeness makes it possible in the first place to hope for such a model structure. The weak equivalences would have to be chosen such that every general logic would be weakly equivalent to a logic in the traditional sense and this would make the disadvantage of unusual objects disappear (up to equivalence). One approach is the following: Every algebra is a quotient of an absolutely free algebra, and for an algebra with consequence relation $(A, \vdash)$ one can choose such an absolutely free algebra $F \twoheadrightarrow A$ and endow it with the biggest consequence relation such that the quotient map becomes a translation (then terms which become equal in A are logically equivalent in F). A different candidate for the underlying category of a model category presenting $\mathcal{L}og_{\mathrm{hamm}}$ would be the category of operads proposed in [2, Sect. 5].

We know from [2, Theorem 2.16] that the 1-category $\mathcal{H}ilb^{\mathrm{strict}}$ of Hilbert systems and strict translations is locally finitely presentable. A locally finitely presentable category is a complete category with a small subcategory of *finitely presentable* objects (i.e. the functors corepresented by them commute with filtered colimits) such that every object is a filtered colimit of these. By [2, Proposition 2.15], finitely presentable objects in the case of $\mathcal{H}ilb^{\mathrm{strict}}$ are the logics with finitely many generating connectives whose consequence relation is generated by finitely many rules. There is a corresponding notion of presentability for $(\infty, 1)$-categories, see [40, Definition 5.5.0.1].

Conjecture 5.5 The $(\infty, 1)$-category $\mathcal{H}ilb_{\mathrm{hamm}}$ is presentable.

From [1] and [2], one can deduce that every logic is a filtered homotopy colimit of the finitely presentable objects in the sense defined above. It would remain to show that these finitely presentable objects are also finitely presentable in the sense of $(\infty, 1)$-categories, i.e. that the mapping space functors corepresented by them commute with filtered homotopy colimits, see [40, Definition 5.3.4.5]. We note that the corresponding conjecture about $\mathcal{H}ilb^{(\mathrm{con})}_{2\text{-cat}}$ is true by work of Mariano and Mendes: In [42, Theorem 2.33], they show that the homotopy category of this category is locally finitely presentable.

All these questions are still centered around our guiding examples of Tarski style abstract logic. Similar investigations to those carried out in the previous sections and proposed in the above questions make sense and would be interesting in other settings of Abstract Logic.

There is, for example, the category $\mathcal{A}lg$ of algebraizable logics in the sense of Blok–Pigozzi [8]. Morphisms of algebraizable logics are translations preserving the so-called algebraizing pairs that come with algebraizable logics, so this is a non-full subcategory of $\mathcal{H}ilb$. Jánossy, Kurucz and Eiben in [31, Definition 3.1.3] define an equivalence relation on the set $\mathrm{Hom}_{\mathcal{A}lg}(A, B)$ of morphisms of algebraizable logics in terms of algebraizing pairs. As for $\mathcal{H}ilb_{2\text{-cat}}$, this equivalence relation gives rise to a simplicial category $\mathcal{A}lg_{2\text{-cat}}$ with homotopy discrete mapping spaces. By Corollary 4.10, homotopy (co)limits in this simplicial category are precisely (co)limits in the homotopy category. The authors investigate this homotopy category, defined (without mention of a simplicial category) in [31, Definition 3.3]. They show that it is equivalent to a certain category of quasivarieties and that this category has non-empty colimits (the restriction to non-empty diagrams may be necessary; by [2, Remark 3.10] one has to be careful with initial objects). Thus we know that we have homotopy colimits of non-empty diagrams in $\mathcal{A}lg_{2\text{-cat}}$. Of course, there is also a hammock localization $\mathcal{A}lg_{\text{hamm}}$ and one may ask about the relationship between the two. In [2, Theorem 3.12], it is shown that the category $\mathcal{A}lg$ is finitely accessible, and one may ask if the same is true in the $(\infty, 1)$-categorical sense [40, Definition 5.4.2.1] for $\mathcal{A}lg_{2\text{-cat}}$.

In a similar vein, one can explore the categories corresponding to the various levels of the Leibniz hierarchy, see [14] for some of these. The results of [20] on preservation of the position in the Leibniz hierarchy under the formation of strict colimits should prove useful here. A pioneering work in this direction is Mariano and Mendes' study of the category of congruential Hilbert systems [42].

A close variant of the logics treated in Sects. 2 and 4, for which the proofs should go through with almost no modifications is obtained by admitting typed signatures. This allows, for example, a natural treatment of first order logic by allowing a type of propositions, as we had before, and types of terms, as well as operations like "=" going from pairs of terms to propositions. Note however that first order logics can also be encoded into propositional logics as done in the appendix of [8].

A variant that still studies Tarski style logics, but with a possibly coarser notion of equivalence, is the program of Mariano and Pinto of representation theory for logics [43]. Indeed, [43, Theorem 3.5] shows that their notion of left Morita equivalence is coarser than the notion of weak equivalence studied here and one can expect an interesting relationship between the two corresponding homotopical categories.

Many other notions of logic, translation and equivalence have been proposed, like those for metafibring [13], those of institutions [15] and π-institutions [21], model-theoretic abstract logics [37], logical spaces [23], type theories [30], and many more, and they all give rise to homotopical categories to be explored and compared.

5.2 Invariants of Logics

Given the fact that logics live naturally in homotopy-theoretical universes, we can ask which other ideas of homotopy theory might apply. Classical homotopy theory studies homotopy invariants of topological spaces. One use of invariants is to discover if two topological spaces are *not* weakly equivalent. One often also tries to compute invariants because of specific information that they contain about a space or a map, not just to merely distinguish them.

Many of these invariants are given by mapping into, or out of, some test object. Let us review the example of singular homology: In Sect. 3.1, we gave a cosimplicial object $\Delta^\bullet$ in the category **Top** of topological spaces and defined the simplicial set $Sing(X)$ by $Sing(X)_n := \mathrm{Hom}_{\mathbf{Top}}(\Delta^n, X)$. Applying the "free abelian group" functor to each of the sets $Sing(X)_n$, one gets an abelian group object in simplicial sets. The homotopy groups of this new simplicial set are the singular homology groups $H_n(X; \mathbb{Z})$. Alternatively, one can build a chain complex from a simplicial abelian group by forming alternating sums of the face maps and take the homology of this chain complex.

Simplicial Sets. The process of the formation of the simplicial set $Sing(X)$ works for any cosimplicial object in a category, but nothing guarantees that this construction has the good properties of singular homology, like the long exact sequences which make computations feasible. But we do indeed have a natural similar construction for logics.

Definition 5.6

1. The category of *general logics*, $\mathcal{G}enLog$, is the category whose objects are pairs $(X, \vdash)$, where X is a set and $\vdash$ a consequence relation on it, and whose morphisms are consequence preserving maps of sets.
2. The notions of *homotopy equivalence* and *weak equivalence* of general logics are defined to be morphisms of general logics satisfying the conditions of Definition 4.1. Note that Proposition 4.2 holds here as well.
3. The *homotopy category of general logics* $Ho(\mathcal{G}enLog)$ is defined to be the homotopy category of the 2-categorical localization (with respect to the obvious preorder enrichment, parallel to the one from the beginning of Sect. 4.3), of $\mathcal{G}enLog$.
4. Denote by $U : \mathcal{L}og^{\text{Tarsk}} \to \mathcal{G}enLog$ the forgetful functor from Tarskian logics to general logics given by $(S, \vdash) \mapsto (\text{Fm}(S), \vdash)$

Note that the functor U induces a functor between the homotopy categories which we will also denote by $U : Ho(\mathcal{L}og^{\text{Tarsk}}) \to Ho(\mathcal{G}enLog)$. We introduced the category $\mathcal{G}enLog$ because it contains a natural cosimplicial object:

Definition 5.7 Let D_n $(n \in \mathbb{N}_0)$ be the general logic whose underlying set is the n-element set $\{\varphi_0, \dots, \varphi_n\}$ and whose consequence relation is the idempotent, increasing consequence relation generated by the rules $\varphi_i \vdash \varphi_{i+1}$.

Since the consequence relations are required to be idempotent and increasing, the general logics D_n clearly form a cosimplicial object, i.e. the obvious maps of ordered sets which leave out or duplicate a proposition become morphisms of general logics:

Definition 5.8 The cosimplicial object $D^\bullet : \Delta \to \mathcal{G}enLog$ is defined on objects by $[n] \mapsto D_n$ and by sending a morphism $f : [n] \to [k]$ of ordered sets to the map $D_n \to D_k, \varphi_i \mapsto \varphi_{f(i)}$.

Mapping out of a cosimplicial object produces a simplicial set. We use this to define a tentative invariant of logics:

Definition 5.9 The simplicial set of inferences $Inf_\bullet(L)$ of a Tarskian logic L is defined to be the simplicial set $\text{Hom}_{\mathcal{G}enLog}(D^\bullet, U(L))$. This defines a functor $Inf_\bullet(L) : \mathcal{L}og^{\text{Tarsk}} \to$ **sSet**.

This simplicial set encodes the implication relations between the formulas of L. Indeed, we have

$$Inf_\bullet(L)_0 \cong \text{Fm}(S_L),$$

$$Inf_\bullet(L)_1 \cong \{(\varphi_0, \varphi_1) \in \mathrm{Fm}(S_L) \times \mathrm{Fm}(S_L) \mid \varphi_0 \vdash \varphi_1\},$$

$$Inf_\bullet(L)_2 \cong \{(\varphi_0, \varphi_1, \varphi_2) \in \mathrm{Fm}(S_L)^3 \mid \varphi_0 \vdash \varphi_1 \vdash \varphi_2\},$$

and so on. One might now have the idea of applying homotopy theoretic invariants like homotopy and homology groups to this simplicial set, but this would contain very limited information about the logic: A simplicial set $X_\bullet$ remembers the direction of the edges; the two structure maps $X_1 \rightrightarrows X_0$ can be seen as source and target maps (and indeed this is what they are if X is the nerve of a category) and they are not interchangeable.

The topological invariants of simplicial sets, however, do not distinguish the direction of the edges, for example, the nerve of a category and the nerve of its opposite are weakly equivalent simplicial sets. Thus applying topological invariants to $Inf_\bullet(L)$ would amount to only remembering whether formulas are connected by some inference, without distinguishing into which direction it goes. Of course, it could still be possible to distinguish non-equivalent logics by these data but a lot of information would be lost.

Instead, one should see the simplicial set $Inf_\bullet(L)$ as an object of *directed homotopy theory*. That is, one should see it not as an object of the standard model category of simplicial sets, where simplicial sets are taken to model topological spaces and in which paths have no preferred direction, but rather as an object of the model category of quasicategories where the objects are representing $(\infty, 1)$-categories.

In general, directed homotopy theory is a field where one studies spaces (for example, modeled by simplicial sets) with directed paths, which cannot be gone backwards. For such a directed space, instead of the fundamental groupoid, one has a fundamental category, instead of homology groups one has preordered homology groups. These seem to be more promising invariants of logics.[2] A good introductory survey of these ideas is [24].

Dendroidal Sets. The simplicial set $Inf_\bullet(L)$ encodes implication relations between single formulas. For finitary logics with a conjunction, this should give fairly complete information. In general, however, the consequence relation is between sets of formulas and single formulas and such inferences with many hypotheses are not captured by the invariant $Inf_\bullet(L)$. For finitary logics we can produce a *dendroidal set* capturing such inferences. We will consider the definition of dendroidal sets via broad posets, as in [50] and [51].

Definition 5.10 A *commutative broad poset* is an idempotent, increasing general logic with the following properties:

1. If $\Gamma \vdash \varphi$, then Γ is finite.
2. If $\gamma \in \Gamma$, $\varphi \in \Phi$, $\Gamma \vdash \varphi$ and $\Phi \vdash \gamma$, then $\varphi = \gamma$.

A commutative broad poset is *finite*, if the relation $\vdash \subseteq \mathscr{P}(A) \times A$ is a finite set. Note that by increasingness this implies that A is finite.

Definition 5.11 Let $(X, \vdash)$ be a commutative broad poset.

An element $x \in X$ is called a *root*, if there do not exist any $y, x_1, \ldots, x_n \in A$, $y \neq x$ such that $\{x, x_1 \ldots, x_n\} \vdash y$.

An element $x \in X$ is called a *leaf*, if there is no $\Gamma \neq \{x\}$ such that $\Gamma \vdash x$.

[2] Cubical sets are more common than simplicial sets in directed algebraic topology, but it is also easy to write down a cocubical object in $\mathcal{G}en\mathcal{L}og$ to produce a cubical variant of $Inf_\bullet(L)$.

Definition 5.12 A commutative broad poset $(X, \vdash)$ is a *dendroidally ordered set*, if

1. It is finite;
2. It has a root; and
3. If $x \in X$ is not a leaf, then there exists a unique $\Gamma \subseteq X$ such that $\Gamma \vdash x$ and with the following property: There exists no $\Gamma' = \{\gamma_1, \ldots, \gamma_n\} \subseteq X$ such that there is a partition $\Gamma = \coprod_i \Gamma_i$ with $\Gamma_i \vdash \gamma_i \, \forall i$.

The full subcategory of general logics whose objects are the dendroidally ordered sets is called Ω.

Definition 5.13 A dendroidal set is a functor $\Omega^{\mathrm{op}} \to \mathbf{Set}$.

Remark 5.14 Our definition of Ω is a reformulation of [50, Definition 4.1.1] and [51, Definition 3.2] (where no logics are mentioned). The definition of Ω in the literature on dendroidal sets is not usually given in terms of broad posets, but rather in terms of trees and operads; see, e.g. [44]. The equivalence between these two definitions is the content of [50, Theorem 4.1.15].

We admit that our definition of Ω is somewhat opaque, but the only thing that matters for now is that it is a subcategory of the category of general logics.

Definition 5.15 The dendroidal set of inferences $\mathit{Inf}_\Omega(L)$ of a Tarskian logic L is defined to be the dendroidal set $\mathrm{Hom}_{\mathcal{G}en\mathcal{L}og}(\Omega, U(L))$. This defines a functor $\mathit{Inf}_\Omega(L) \colon \mathcal{L}og^{\mathrm{Tarsk}} \to \mathbf{Set}^{\Omega^{\mathrm{op}}}$.

Like the category of simplicial sets, the category of dendroidal sets bears several model structures. There is the Cisinski–Moerdijk model structure [12], with the fewest weak equivalences, which, roughly, sees dendroidal sets as encoding families of n-ary operations, closed under composition (operads). This will be the viewpoint that retains the most information about the logic. Further there are the covariant model structure of Heuts [28] which sees dendroidal sets as E_∞-spaces and the stable model structure of Bašić/Nikolaus [6] which sees dendroidal sets as connective spectra.

All of these model structures can be used to associate invariants to logics. In particular, the last one can be used to define the algebraic K-theory of a logic, see [45]. It remains to be seen how computable these invariants are, whether, e.g. there are long exact sequences induced by cofibrations of logics, and what exactly they capture about a logic. They are likely to get much more interesting, if one enriches the categories of logics via proof theory as in Sect. 5.3.

Galois Style Invariants. A further type of invariant of a logic L can be constructed by considering the category $L \downarrow \mathcal{L}og$ of logics receiving a translation from L and associating to it the group of automorphisms of the forgetful functor to $\mathcal{L}og$ or other categories as that of indexed frames, see Sect. 5.4. The homotopical viewpoint suggests, however, that one should take autoequivalences instead of automorphisms. This can be seen as a version of Galois theory for logics.

In a similar spirit, are the Morita style invariants of Mariano and Pinto [43] who, roughly, associate to a logic the categories of algebraizable logics over/under the logic.

5.3 Refined Categories of Logics from Proof Theory

The natural enrichment of the categories of logics in preorders lead us to the 2-categorical localizations of Sect. 4.3. As much as this was an improvement of the corresponding usual categories, the homotopy discrete mapping spaces are not very interesting objects. However, they can be seen as the shadows of richer structures. We just give some sketches and ideas here.

The analogy between proof theory and homotopy theory is as follows: A logic is a space, formulas are points, proofs are paths between the points, transformations of proofs are homotopies between the paths.

The somewhat degenerate situation of homotopy discrete mapping spaces in logic can be seen from this angle. We said that two translations f, g should be declared equivalent if for all φ of the domain, we have $f(\varphi) \dashv\vdash g(\varphi)$. This only asks for *provability* of $f(\varphi)$ from $g(\varphi)$, it does not distinguish different *proofs*. This is like asking, on the topological side, whether two points $f(\varphi)$ and $g(\varphi)$ are in the same path component, while ignoring the different paths. Distinguishing different proofs is a way to get to more interesting enrichments of the categories of logics.

Recall that categories (actually multicategories) of proofs were introduced by Lambek [36]. Given a Hilbert system, presented by a set of deduction rules, one can define a category whose objects are the formulas and whose morphisms are proofs. Let us say for the moment that a proof from a hypothesis φ to a conclusion ψ is a sequence of formulas such that the final formula is ψ and any intermediate formula is either φ or follows from the preceding ones through one of the deduction rules.

We wish to use this for an enrichment of the category of Hilbert systems. We could start by saying that $\mathrm{Hom}_{\mathcal{H}ilb}(L, L')$ is the category whose objects are translations from L to L' and where a morphism $f \to g$ is given by a collection of proofs $\{f(\varphi) \to g(\varphi) \mid \varphi \in \mathrm{Fm}(L)\}$, i.e. we could ask pointwise for morphisms in the proof category of L'. Categorical thinking would, however, demand some kind of coherence between the different morphisms. If we think of a morphism from f to g as something like a natural transformation between functors, then we would ask for the commutativity of a certain diagram: A proof $\varphi \to \psi$ in the proof category of L would be mapped to two proofs $f(\varphi) \to f(\psi)$ and $g(\varphi) \to g(\psi)$ and we could ask that the naturality diagram commutes:

$$\begin{array}{ccc} f(\varphi) & \longrightarrow & g(\varphi) \\ \downarrow & & \downarrow \\ f(\psi) & \longrightarrow & g(\psi) \end{array}$$

This means asking for an equality of proofs, which again seems quite restrictive. On a middle ground, we could ask that the two proofs are comparable in some sense, or transformable into each other.

This leaves us with lots of interesting options to choose from, all resulting in different categories. There are directed and symmetric versions of relations between proofs. Choosing the symmetric versions results in a (2, 1)-category of logics, while choosing the directed versions results in a (2, 2)-category of logics. Here are some examples:

1. Length of the proofs
   - (symmetric) equal length
   - (directed) one is longer than the other
2. Normal forms
   - (symmetric) both proofs have the same normal form of some kind
   - (directed) One is closer to normal form than the other (e.g. elimination rules occur before introduction rules)
3. Degree of generalizability (Lambek). Suppose we have a proof $p : A \to B$ and the formulas A and B arise by substitution of terms $t_1, \ldots, t_n$ into other formulas φ, ψ, i.e. $A = \varphi(t_1, \ldots, t_n)$, $B = \psi(t_1, \ldots, t_n)$. Then one can ask whether the proof carries over to the more general situation, i.e. whether there is a proof $\hat{p}(x_1, \ldots, x_n) : \varphi(x_1, \ldots, x_n) \to \psi(x_1, \ldots, x_n)$ such that $p = \hat{p}(t_1, \ldots, t_n)$.
   - (symmetric) p and q have the same degree of generalizability, i.e. for every "generalization" of A and B the proof p carries over if and only if q carries over
   - (directed) For every generalization to which p carries over, q also carries over
4. Required strength of the logic. One proof $p : A \to B$ might use, e.g. Modus Ponens, $\varphi \wedge \psi \vdash \varphi$ and $\neg\neg\phi \vdash \phi$ while another proof $q : A \to B$ only uses Modus Ponens. Or one proof might be constructive while the other is not.
   - (symmetric) Both proofs use the same logical strength
   - (directed) One uses less than the other

The list could easily go on, but the point is that there are many interesting relations between proofs that one might want to study and they all lead to different categories with more interesting enrichments than before.

Going further, one could try to get a 3-category of logics, i.e. an enrichment in 2-categories. For this consider the example of normal forms: Often a proof can be brought into some normal form by a sequence of elementary steps. These sequences of steps can be seen as 2-morphisms in the proof category. Again one has to ask when two sequences of steps can be considered equivalent and get a 2-category of translations by pointwise application.

In usual logic, nothing seems to be naturally coming after that. In Martin-Löf type theory, on the other hand, one has identity types and can iterate up to arbitrary levels, which is exactly what inspired the homotopy theoretical semantics used in homotopy type theory. This should lead to the richest, least truncated, higher categories of logics.

We still remark that one can adapt the invariants of logics from the previous section to the setting of these less truncated categories of logics involving some proof theory and that here they might get more of a homotopical flavor than their more truncated companions. We also remark that other natural enrichments can be found, for example, an enrichment in multicategories, via the provability relation between sets of formulas and single formulas.

5.4 Comparing Paradigms of Logic

Another question that gets an interesting twist, once we have turned our categories of logics into higher categories, is that of how the categories corresponding to different formalizations of abstract logic, like institution theory and Tarski style consequence relations, relate to each other. With the extra flexibility of $(\infty, 1)$-categories it seems easier to get adjunctions or equivalences between different such settings.

In another direction, there is the work [27] where the authors note that for several formalizations of the notion of logic the resulting category has a forgetful functor to the category of *indexed frames*. A typical way to assess a category with such a forgetful functor is to consider the automorphisms of this functor. In good situations, one can reconstruct the category from the base category and the knowledge of these automorphisms, but in any case one can associate in this way a group to a category. In the world of $(\infty, 1)$-categories, one should instead consider *autoequivalences*.

6 Conclusion

Our aim was to show that a homotopy-theoretical point of view is very natural and appropriate in logic. One is led to natural constructions and questions and this should be the main reason to adopt such a point of view. A good side effect is, as exemplified by the (co)completeness results of Sect. 4, that things look better than they might otherwise through the lens of usual category theory.

We believe that the rich homotopy-theoretical landscape of logics, of which we have unveiled a bit, gives ample confirmation of how fundamental and fruitful Jean-Yves' questions from the beginning of this article really are.

References

1. Arndt, P.: Homotopical fibring. To appear
2. Arndt, P., de Alvarenga Freire, R., Luciano, O., Mariano, H.: A global glance on categories in logic. Log. Univers. **1**(1), 3–39 (2007). doi:10.1007/s11787-006-0002-7. ISSN 1661-8297
3. Barwick, C.: On left and right model categories and left and right Bousfield localizations. Homol. Homotopy Appl. **12**(2), 245–320 (2010). http://projecteuclid.org/getRecord?id=euclid.hha/1296223884. ISSN 1532-0073
4. Barwick, C., Kan, D.: A characterization of simplicial localization functors and a discussion of DK equivalences. Indag. Math. **23**(1–2), 69–79 (2012). doi:10.1016/j.indag.2011.10.001. ISSN 0019-3577
5. Barwick, C., Kan, D.: Relative categories: another model for the homotopy theory of homotopy theories. Indag. Math. **23**(1–2), 42–68 (2012). doi:10.1016/j.indag.2011.10.002. ISSN 0019-3577
6. Bašić, M., Nikolaus, T.: Dendroidal sets as models for connective spectra. Preprint. arXiv:1203.6891
7. Bergner, J.E.: A survey of $(\infty, 1)$-categories. In: Towards Higher Categories. IMA Vol. Math. Appl., vol. 152, pp. 69–83. Springer, New York (2010). doi:10.1007/978-1-4419-1524-5_2
8. Blok, W., Pigozzi, D.: Algebraizable Logics. Mem. Am. Math. Soc., vol. 77(396) (1989). doi:10.1090/memo/0396. vi+78pp. ISSN 0065-9266.
9. Bloom, S., Brown, D., Suszko, R.: Some theorems on abstract logics. Algebra Log. **9**, 274–280 (1970). ISSN 0373-9252
10. Caleiro, C., Gonçalves, R.: Equipollent logical systems. In: Béziau, J.-Y. (ed.) Logica Universalis: Towards a General Theory of Logic. Birkhäuser, Basel (2007)
11. Caleiro, C., Mateus, P., Ramos, J., Sernadas, A.: Combining logics: parchments revisited. In: Recent Trends in Algebraic Development Techniques. Lecture Notes in Computer Science, vol. 2267, pp. 48–70 (2002)

12. Cisinski, D., Moerdijk, I.: Dendroidal sets and simplicial operads. J. Topol. **6**(3), 705–756 (2013). doi:10.1112/jtopol/jtt006. ISSN 1753-8416; 1753-8424/e
13. Coniglio, M.: Recovering a logic from its fragments by meta-fibring. Log. Univers. **1**(2), 377–416 (2007). doi:10.1007/s11787-007-0019-6. ISSN 1661-8297
14. Czelakowski, J.: Protoalgebraic Logics. Trends in Logic—Studia Logica Library, vol. 10. Kluwer Academic, Dordrecht (2001). xii+452pp. ISBN 0-7923-6940-8
15. Diaconescu, R.: Institution-Independent Model Theory. Studies in Universal Logic. Birkhäuser, Basel (2008). xii+376pp. ISBN 978-3-7643-8707-5
16. Dwyer, W., Hirschhorn, P., Kan, D., Smith, J.: Homotopy Limit Functors on Model Categories and Homotopical Categories. American Mathematical Society (AMS), Providence (2004). vii+181pp. ISBN 0-8218-3703-6
17. Dwyer, W., Kan, D.: Calculating simplicial localizations. J. Pure Appl. Algebra **18**(1), 17–35 (1980). doi:10.1016/0022-4049(80)90113-9. ISSN 0022-4049
18. Dwyer, W., Kan, D.: Function complexes in homotopical algebra. Topology **19**(4), 427–440 (1980). doi:10.1016/0040-9383(80)90025-7. ISSN 0040-9383
19. Dwyer, W., Spaliński, J.: Homotopy theories and model categories. In: Handbook of Algebraic Topology, pp. 73–126. North-Holland, Amsterdam (1995)
20. Fernández, V., Coniglio, M.: Fibring in the Leibniz hierarchy. Log. J. IGPL **15**(5–6), 475–501 (2007). doi:10.1093/jigpal/jzm036. ISSN 1367-0751
21. Fiadeiro, J., Sernadas, A.: Structuring theories on consequence. In: Sannella, D., Tarlecki, A. (eds.) Recent Trends in Data Type Specification. Lecture Notes in Computer Science, vol. 332, pp. 44–72. Springer, Berlin (1988). doi:10.1007/3-540-50325-0_3. ISBN 978-3-540-50325-5
22. Goerss, P., Jardine, J.F.: Simplicial Homotopy Theory. Birkhäuser, Basel (2009)
23. Gomi, K.: Theory of completeness for logical spaces. Log. Univers. **3**(2), 243–291 (2009). doi:10.1007/s11787-009-0008-z. ISSN 1661-8297
24. Grandis, M.: Directed algebraic topology, categories and higher categories. Appl. Categ. Struct. **15**(4), 341–353 (2007). doi:10.1007/s10485-007-9084-5. ISSN 0927-2852
25. Groth, M.: A short course on infinity-categories. Preprint. arXiv:1007.2925
26. Groth, M.: Derivators, pointed derivators and stable derivators. Algebr. Geom. Topol. **13**(1), 313–374 (2013). doi:10.2140/agt.2013.13.313. ISSN 1472-2747; 1472-2739/e
27. Haeusler, E., Martini, A., Wolter, U.: Towards a uniform presentation of logical systems by indexed categories and adjoint situations. J. Log. Comput. (2012). doi:10.1093/logcom/exs038
28. Heuts, G.: Algebras over infinity-operads. Preprint. arXiv:1110.1776
29. Hovey, M.: Model Categories. Mathematical Surveys and Monographs, vol. 63. American Mathematical Society, Providence (1999). xii+209pp. ISBN 0-8218-1359-5
30. Jacobs, B.: Categorical Logic and Type Theory. Studies in Logic and the Foundations of Mathematics, vol. 141. North-Holland, Amsterdam (1999). xviii+760pp. ISBN 0-444-50170-3
31. Jánossy, A., Kurucz, A., Eiben, A.: Combining algebraizable logics. Notre Dame J. Form. Log. **37**(2), 366–380 (1996). doi:10.1305/ndjfl/1040046092. Combining logics. ISSN 0029-4527
32. Joyal, A.: Notes on quasi-categories. Preprint. http://www.math.uchicago.edu/~may/IMA/Joyal.pdf
33. Kashiwara, M., Schapira, P.: Categories and Sheaves. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 332. Springer, Berlin (2006). x+497pp. ISBN 978-3-540-27949-5
34. Kelly, G.: Basic Concepts of Enriched Category Theory. Theory Appl. Categ., vol. 10 (2005). vi+137pp. Reprint of the 1982 original [Cambridge Univ. Press, Cambridge; MR0651714]
35. Lack, S.: Homotopy-theoretic aspects of 2-monads. J. Homotopy and Relat. Struct. (2007)
36. Lambek, J.: Deductive systems and categories II. Standard constructions and closed categories. In: Category Theory, Homology Theory and Their Applications I. Lecture Notes in Mathematics, vol. 86, pp. 76–122. Springer, Berlin (1969). doi:10.1007/BFb0079385. ISBN 978-3-540-04605-9
37. Lewitzka, S.: A topological approach to universal logic: model-theoretical abstract logics. In: Logica Universalis, pp. 35–63. Birkhäuser, Basel (2005). doi:10.1007/3-7643-7304-0_3
38. Łoś, J., Suszko, R.: Remarks on sentential logics. Proc. K. Ned. Akad. Wet., Ser. A, Indag. Math. **20**, 177–183 (1958)
39. Lurie, J.: Higher algebra. Preprint. http://www.math.harvard.edu/~lurie/papers/higheralgebra.pdf
40. Lurie, J.: Higher Topos Theory. Annals of Mathematics Studies, vol. 170. Princeton University Press, Princeton (2009). xviii+925pp. ISBN 978-0-691-14049-0

41. Mariano, H., Mendes, C.: Towards a good notion of categories of logics book of abstracts for: topology, algebra and categories in logic, Marseille (2011)
42. Mariano, H., Mendes, C.: Towards a good notion of categories of logics. Preprint. arXiv:1404.3780
43. Mariano, H., Pinto, D.: Representation theory of logics: a categorial approach. Preprint. arXiv:1405.2429
44. Moerdijk, I., Weiss, I.: On inner Kan complexes in the category of dendroidal sets. Adv. Math. **221**(2), 343–389 (2009)
45. Nikolaus, T.: Algebraic k-theory of infinity-operads. Preprint. arXiv:1303.2198
46. Radulescu-Banu, A.: Cofibrations in homotopy theory. arXiv preprint (2009). arXiv:math/0610009
47. Sernadas, A., Sernadas, C., Caleiro, C.C.: Fibring of logics as a categorial construction. J. Log. Comput. (1999)
48. Shulman, M.: Homotopy limits and colimits and enriched homotopy theory. Preprint
49. Szumiło, K.: Available at http://www.math.uni-bonn.de/people/szumilo/papers/cht.pdf
50. Trova, F.: On the geometric realization of dendroidal sets. Master's thesis, ALGANT, Leiden/Padova (2009)
51. Weiss, I.: Broad posets, trees and the dendroidally ordered category. Preprint. arXiv:1201.3987
52. Wójcicki, R.: Logical matrices strongly adequate for structural sentential calculi. Bull. Acad. Pol. Sci., Sér. Sci. Math. Astron. Phys. **17**, 333–335 (1969). ISSN 0001-4117
53. Wójcicki, R.: Basic theory of consequence operations. In: Theory of Logical Calculi. Synthese Library, vol. 199. Kluwer Academic, Dordrecht (1988). xviii+473pp. ISBN 90-277-2785-6

P. Arndt (✉)
Fakultät für Mathematik, Universität Regensburg, 93040 Regensburg, Germany
e-mail: peter.arndt@gmail.com

Semi-implication: A Chapter in Universal Logic

Arnon Avron

Abstract We introduce a general notion of semi-implication which generalizes both the implications used in classical, intuitionistic, and many other logics, as well as those used in relevance logics. It is mainly based on the relevant deduction property (RDP)—a weak form of the classical-intuitionistic deduction theorem which has motivated the design of the intensional fragments of the relevance logic **R**. However, $\mathbf{CL}_{\leftrightarrow}$, the pure equivalential fragment of classical logic, also enjoys the RDP with respect to $\leftrightarrow$. We show that in the language of $\rightarrow$ this is the only exception. This observation leads to an adequate definition of semi-implication, according to which a finitary logic **L** has a semi-implication $\rightarrow$ iff **L** has a strongly sound and complete Hilbert-type system which is an extension by axiom schemas of $HR_{\rightarrow}$ (the standard Hilbert-type system for the implicational fragment of **R**). We also show that in the presence of a conjunction, or a disjunction, or an implication, a connective $\rightarrow$ of a logic **L** is a semi-implication iff it is an implication (i.e. it satisfies the classical-intuitionistic deduction theorem), and the same is true if **L** is induced by a matrix which has a single designated value or a single non-designated value.

Keywords Implication · Semi-implication · Biconditional · Relevance logics · Classical logic · Deduction theorems

Mathematics Subject Classification (2000) Primary 03B20 · Secondary 03B47

1 Introduction

In general, we take a connective $\rightarrow$ of a logic **L** to be an implication for **L** if it reflects the underlying consequence relation of **L**. Usually, this means that it satisfies the classical-intuitionistic deduction theorem. In such a case, its availability makes it possible to directly reduce all inferences from finite sets of premises to theoremhood in a rather straightforward way: $\psi_1, \ldots, \psi_n \vdash_{\mathbf{L}} \varphi$ iff $\vdash_{\mathbf{L}} \psi_1 \rightarrow (\psi_2 \rightarrow (\cdots \rightarrow (\psi_n \rightarrow \varphi) \cdots))$. However, in relevance logics (see [1, 10]) it is reasonable to *demand* the presence of ψ_i in the above implication only if it is really relevant to the derivability of φ from $\{\psi_1, \ldots, \psi_n\}$. Our major goal in this paper is to introduce and investigate an adequate general notion of *semi-implication* which reflects this intuition, but does not depend on the somewhat problematic notion of "a use of a formula in a given proof" (like that used in [1, 10] to motivate some of the relevance systems), or even on the availability of any proof system for the logic. As we shall see, achieving this goal involves some complications, and the need to overcome them leads to a notion which is somewhat more complex than might be expected.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_3

2 Preliminaries

In the sequel, $\mathcal{L}$ denotes a propositional language. The set of well-formed formulas of $\mathcal{L}$ is denoted by $\mathcal{W}(\mathcal{L})$, and φ, ψ, θ vary over its elements. $\mathcal{T}, \mathcal{S}$ vary over theories of $\mathcal{L}$ (where by a 'theory' we simply mean here a subset of $\mathcal{W}(\mathcal{L})$). We denote by $\mathsf{Atoms}(\varphi)$ ($\mathsf{Atoms}(\mathcal{T})$) the set of atomic formulas that appear in φ (in the formulas of $\mathcal{T}$).

Definition 2.1 A (Tarskian) *consequence relation* (tcr) for a language $\mathcal{L}$ is a binary relation $\vdash$ between theories in $\mathcal{W}(\mathcal{L})$ and formulas in $\mathcal{W}(\mathcal{L})$, satisfying the following three conditions:

[R]	(*Reflexivity*)	$\psi \vdash \psi$ (i.e. $\{\psi\} \vdash \psi$);
[M]	(*Monotonicity*)	If $\mathcal{T} \vdash \psi$ and $\mathcal{T} \subseteq \mathcal{T}'$, then $\mathcal{T}' \vdash \psi$;
[C]	(*Cut (Transitivity)*)	If $\mathcal{T} \vdash \psi$ and $\mathcal{T}', \psi \vdash \varphi$ then $\mathcal{T} \cup \mathcal{T}' \vdash \varphi$.

Definition 2.2 Let $\vdash$ be a Tarskian consequence relation for $\mathcal{L}$.

- $\vdash$ is *structural*, if for every $\mathcal{L}$-substitution θ and every $\mathcal{T}$ and ψ, if $\mathcal{T} \vdash \psi$ then $\theta(\mathcal{T}) \vdash \theta(\psi)$.
- $\vdash$ is *non-trivial* if $p \nvdash q$ for distinct atoms $p, q \in \mathsf{Atoms}(\mathcal{L})$.
- $\vdash$ is *finitary* if, for every theory $\mathcal{T}$ and every formula ψ such that $\mathcal{T} \vdash \psi$, there is a *finite* theory $\Gamma \subseteq \mathcal{T}$ such that $\Gamma \vdash \psi$.

Now we define the notion of propositional logic which is used in this paper (which is identical to that used in [7]):

Definition 2.3

- A (propositional) *logic* is a pair $\mathbf{L} = \langle \mathcal{L}, \vdash_{\mathbf{L}} \rangle$, where $\mathcal{L}$ is a propositional language, and $\vdash_{\mathbf{L}}$ is a structural and non-trivial Tarskian consequence relation for $\mathcal{L}$.
- A logic $\mathbf{L} = \langle \mathcal{L}, \vdash_{\mathbf{L}} \rangle$ is *finitary* if so is $\vdash_{\mathbf{L}}$.

Next we present abstract characterizations of some basic connectives:

Definition 2.4 Let $\mathbf{L} = \langle \mathcal{L}, \vdash_{\mathbf{L}} \rangle$ be a propositional logic.

- A binary connective $\to$ of $\mathcal{L}$ is called an *implication for* $\mathbf{L}$ if the classical deduction theorem holds for $\to$ and $\vdash_{\mathbf{L}}$:

$$\mathcal{T}, \varphi \vdash_{\mathbf{L}} \psi \quad \text{iff } \mathcal{T} \vdash_{\mathbf{L}} \varphi \to \psi.$$

- A binary connective $\wedge$ of $\mathcal{L}$ is called a *conjunction for* $\mathbf{L}$ if it satisfies the following condition:

$$\mathcal{T} \vdash_{\mathbf{L}} \psi \wedge \varphi \quad \text{iff } \mathcal{T} \vdash_{\mathbf{L}} \psi \text{ and } \mathcal{T} \vdash_{\mathbf{L}} \varphi.$$

- A binary connective $\vee$ of $\mathcal{L}$ is called a *disjunction for* $\mathbf{L}$ if it satisfies the following condition:

$$\mathcal{T}, \psi \vee \varphi \vdash_{\mathbf{L}} \theta \quad \text{iff } \mathcal{T}, \psi \vdash_{\mathbf{L}} \theta \text{ and } \mathcal{T}, \varphi \vdash_{\mathbf{L}} \theta.$$

Note 2.5 It is easy to verify that $\wedge$ is a conjunction for $\mathbf{L} = \langle \mathcal{L}, \vdash_{\mathbf{L}} \rangle$ iff the following three conditions hold for every $\psi, \varphi \in \mathcal{W}(\mathcal{L})$:

1. $\psi \wedge \varphi \vdash_{\mathbf{L}} \psi$;
2. $\psi \wedge \varphi \vdash_{\mathbf{L}} \varphi$;
3. $\psi, \varphi \vdash_{\mathbf{L}} \psi \wedge \varphi$.

On the basis of this, it is not difficult to show that one may equivalently define a conjunction for $\mathbf{L}$ as a connective $\wedge$ such that

$$\mathcal{T}, \psi \wedge \varphi \vdash_{\mathbf{L}} \theta \quad \text{iff } \mathcal{T}, \psi, \varphi \vdash_{\mathbf{L}} \theta.$$

It follows that a conjunction connective for $\mathbf{L}$ assures that for every $\psi_1, \ldots, \psi_n, \varphi \in \mathcal{W}(\mathcal{L})$ it holds that $\psi_1, \ldots, \psi_n \vdash_{\mathbf{L}} \varphi$ iff $\psi_1 \wedge \cdots \wedge \psi_n \vdash_{\mathbf{L}} \varphi$.[1]

Note 2.6 It is well-known that a finitary logic $\mathbf{L}$ has an implication connective $\rightarrow$ iff $\mathbf{L}$ has a strongly sound and complete Hilbert-type system which is an extension by axiom schemas of $\mathbf{H}_{\rightarrow}$ (the standard Hilbert-type system for the pure implicational fragment of intuitionistic logic).

3 The Relevant Deduction Property (RDP)

As noted in the introduction, an implication connective $\rightarrow$ for a logic $\mathbf{L}$ reflects the underlying consequence relation of $\mathbf{L}$. However, such a connective necessarily validates non-intuitive formulas like $p \rightarrow (q \rightarrow p)$ and $p \rightarrow (q \rightarrow q)$. The latter formula also violates the variable sharing property (VSP)—the most basic criterion for validity of implication in relevance logics (see [1, 10]).[2] For logics which respect the principle that in valid entailments all the assumptions should be relevant to the conclusion, it is indeed reasonable to *demand* the derivability of $\varphi \rightarrow \psi$ from $\mathcal{T}$ only in case φ is absolutely necessary for the derivability of ψ from $\mathcal{T} \cup \{\varphi\}$. Hence it was noted in [7] that the entailment connective of these logics should satisfy a weaker condition:

Definition 3.1 Let $\mathbf{L} = \langle \mathcal{L}, \vdash_{\mathbf{L}} \rangle$ be a propositional logic, and let $\rightarrow$ be a (primitive or defined) connective of $\mathcal{L}$. $\rightarrow$ has in $\mathbf{L}$ (or $\mathbf{L}$ has with respect to $\rightarrow$) the *relevant deduction property* (RDP) if it satisfies the following condition:

$$\mathcal{T}, \varphi \vdash_{\mathbf{L}} \psi \quad \text{iff either } \mathcal{T} \vdash_{\mathbf{L}} \psi \text{ or } \mathcal{T} \vdash_{\mathbf{L}} \varphi \rightarrow \psi.$$

Note 3.2 As was noted in [7], the RDP does not depend on the notion of "a use of a formula in a given proof", and is indeed the way in which A. Church stated his deduction theorem in [8] (the paper in which the basic relevance logic $HR_{\rightarrow}$ described below was introduced).

[1]Here $\psi_1 \wedge \cdots \wedge \psi_n$ might be taken to stand, e.g. for $(\cdots((\psi_1 \wedge \psi_2) \wedge \psi_3)\cdots) \wedge \psi_n$.

[2]A connective $\rightarrow$ of a propositional logic $\mathbf{L}$ has the variable sharing property (VSP) if $\mathsf{Atoms}(\varphi) \cap \mathsf{Atoms}(\psi) \neq \emptyset$ whenever $\vdash_{\mathbf{L}} \varphi \rightarrow \psi$.

Axioms:

[Id]	$\varphi \to \varphi$	(Identity)
[Tr]	$(\varphi \to \psi) \to ((\psi \to \theta) \to (\varphi \to \theta))$	(Transitivity)
[Pe]	$(\varphi \to (\psi \to \theta)) \to (\psi \to (\varphi \to \theta))$	(Permutation)

Rule of inference:

[MP] $$\frac{\varphi \qquad \varphi \to \psi}{\psi}$$

Fig. 1 The proof system $HLL_{\to}$

We start our investigation of logics with the RDP by presenting a basic well-known logic whose theorems are valid whenever $\to$ has the RDP:

Definition 3.3 Let $\mathcal{L} = \{\to\}$. $\mathbf{LL}_{\to}$ is the logic induced by the Hilbert-type proof system $HLL_{\to}$ presented in Fig. 1.

Note 3.4 $\mathbf{LL}_{\to}$ is the pure implicational fragment of linear logic [11]. $HLL_{\to}$ is the axiomatization of this fragment given in [3].

Theorem 3.5 *Let* $\mathbf{L}$ *be a logic in a language which contains* $\to$, *and suppose that* $\to$ *has in* $\mathbf{L}$ *the* RDP. *Then* $\mathbf{L}$ *contains* $\mathbf{LL}_{\to}$.[3]

Proof Since $\mathbf{L}$ is a logic, and so structural, it suffices to show the validity of the axioms and inference rule of $HLL_{\to}$ for the case where φ, ψ and θ are different atomic formulas p, q, r (respectively).

- (Id) Since $p \vdash_{\mathbf{L}} p$, but $\nvdash_{\mathbf{L}} p$, the RDP implies that $\vdash_{\mathbf{L}} p \to p$.
- (MP) Since $p \to q \vdash_{\mathbf{L}} p \to q$, the RDP entails that $p, p \to q \vdash_{\mathbf{L}} q$.
- (Tr) From the validity of [MP] it follows that $p, p \to q, q \to r \vdash_{\mathbf{L}} r$. It is impossible that $p \to q, q \to r \vdash_{\mathbf{L}} r$, since otherwise we would get $\vdash_{\mathbf{L}} p$ by substituting p for q and r, and using the provability of [Id]. Hence $p \to q, q \to r \vdash_{\mathbf{L}} p \to r$ by the RDP. Now it is impossible that $p \to q \vdash_{\mathbf{L}} p \to r$, since otherwise by substituting p for q we would get $\vdash_{\mathbf{L}} p \to r$, although $p \nvdash_{\mathbf{L}} r$. Hence the RDP entails that $p \to q \vdash_{\mathbf{L}} (q \to r) \to (p \to r)$. Again it is impossible that $\vdash_{\mathbf{L}} (q \to r) \to (p \to r)$, since otherwise we would get $\vdash_{\mathbf{L}} (p \to q)$ by substituting q for r. Hence the RDP entails that $\vdash_{\mathbf{L}} (p \to q) \to ((q \to r) \to (p \to r))$.
- (Pe) The validity of [MP] entails that $p \to (q \to r), p, q \vdash_{\mathbf{L}} r$. On the other hand, $p \to (q \to r), q \nvdash_{\mathbf{L}} r$, because otherwise we would get $q \vdash_{\mathbf{L}} r$ by substituting $q \to r$ for p. Hence the RDP implies that $p \to (q \to r), q \vdash_{\mathbf{L}} p \to r$. Now $p \to (q \to r) \nvdash_{\mathbf{L}} p \to r$, because otherwise by substituting $q \to q$ for p and q for r, we would get $\vdash_{\mathbf{L}} q$ (using previous items). Hence the RDP implies that $p \to (q \to r) \vdash_{\mathbf{L}} q \to (p \to r)$. Finally, $\nvdash_{\mathbf{L}} q \to (p \to r)$, since otherwise we would get $\vdash_{\mathbf{L}} r$ by substituting $q \to q$ for p and q. Hence the RDP implies that $\vdash_{\mathbf{L}} (p \to (q \to r)) \to (q \to (p \to r))$. □

[3]This theorem has already been proved in [7]. We reproduce here the proof in order to make this paper self-contained.

Our next results show that in many cases a connective $\to$ which has the RDP is necessarily an implication.

Lemma 3.6 *Suppose $\to$ has in $\mathbf{L}$ the* RDP. *Then $\to$ is an implication for $\mathbf{L}$ iff $\vdash_{\mathbf{L}} \psi \to (\varphi \to \psi)$ for every φ and ψ, iff $q \vdash_{\mathbf{L}} p \to q$ for two distinct atomic formulas p, q.*

Proof Suppose $\to$ has in $\mathbf{L}$ the RDP. By Theorem 3.5, this implies that $\varphi, \varphi \to \psi \vdash_{\mathbf{L}} \psi$ for every φ and ψ, and so if $\mathcal{T} \vdash_{\mathbf{L}} \varphi \to \psi$ then $\mathcal{T}, \varphi \vdash_{\mathbf{L}} \psi$. It follows that $\to$ is an implication for $\mathbf{L}$ iff the converse holds as well, i.e. if $\mathcal{T}, \varphi \vdash_{\mathbf{L}} \psi$ implies that $\mathcal{T} \vdash_{\mathbf{L}} \varphi \to \psi$. Next we show that this is equivalent to $\psi \vdash_{\mathbf{L}} \varphi \to \psi$ for every φ and ψ. The latter is obviously a necessary condition for the former. To show that it is also sufficient, assume that $\mathcal{T}, \varphi \vdash_{\mathbf{L}} \psi$. By the RDP, either $\mathcal{T} \vdash_{\mathbf{L}} \varphi \to \psi$ or $\mathcal{T} \vdash_{\mathbf{L}} \psi$. In the first case, we are done, while the assumption that $\psi \vdash_{\mathbf{L}} \varphi \to \psi$ implies this in the second case, too.

Now since $\mathbf{L}$ is structural, that $\psi \vdash_{\mathbf{L}} \varphi \to \psi$ for every φ and ψ is equivalent to $q \vdash_{\mathbf{L}} p \to q$ for any distinct atomic formulas p, q. By the RDP, this in turn is equivalent to $\vdash_{\mathbf{L}} q \to (p \to q)$ (otherwise we would have had $\vdash_{\mathbf{L}} p \to q$. Hence $p \vdash_{\mathbf{L}} q$), and so to $\vdash_{\mathbf{L}} \psi \to (\varphi \to \psi)$ for every φ and ψ. □

Theorem 3.7 *If $\mathbf{L}$ has a conjunction or a disjunction or an implication, and $\to$ has in $\mathbf{L}$ the* RDP, *then $\to$ is an implication for $\mathbf{L}$.*

Proof Suppose that $\to$ has in $\mathbf{L}$ the RDP.

- Assume that $\wedge$ is a conjunction for $\mathbf{L}$. From the three properties of $\wedge$ listed in Note 2.5 it easily follows that if p and q are two distinct atomic formulas, then $p, q \vdash_{\mathbf{L}} p \wedge q$, while $p \nvdash_{\mathbf{L}} p \wedge q$ (otherwise we would have got that $p \vdash_{\mathbf{L}} q$). Since $\to$ has in $\mathbf{L}$ the RDP, this implies that $p \vdash_{\mathbf{L}} q \to (p \wedge q)$. Similarly, these properties of $\wedge$ implies that $\vdash_{\mathbf{L}} (p \wedge q) \to p$. Since, by Theorem 3.5, $\varphi \to \psi, \psi \to \theta \vdash_{\mathbf{L}} \varphi \to \theta$, it follows that $p \vdash_{\mathbf{L}} q \to p$, and so $\to$ is an implication for $\mathbf{L}$ by Lemma 3.6.
- Assume that $\vee$ is a disjunction for $\mathbf{L}$. Since $p \vee q \vdash_{\mathbf{L}} p \vee q$, we have (i) $p \vdash_{\mathbf{L}} p \vee q$ and (ii) $q \vdash_{\mathbf{L}} p \vee q$. Now Theorem 3.5 implies that $p, p \to q \vdash_{\mathbf{L}} q$. Obviously, also $q, p \to q \vdash_{\mathbf{L}} q$. As $\vee$ is a disjunction for $\mathbf{L}$, the last two facts imply that $p \vee q, p \to q \vdash_{\mathbf{L}} q$. Since $p \vee q \nvdash_{\mathbf{L}} q$ (otherwise we would get $p \vdash_{\mathbf{L}} q$ by (i)), this implies that $p \vee q \vdash_{\mathbf{L}} (p \to q) \to q$, and so by (ii) that $q \vdash_{\mathbf{L}} (p \to q) \to q$. By substituting $p \to q$ for p in the last fact, we get (*) $q \vdash_{\mathbf{L}} ((p \to q) \to q) \to q$. But it is easy to see that $\vdash_{\mathbf{LL}_{\to}} (((p \to q) \to q) \to q) \to (p \to q)$. Hence, Theorem 3.5 and (*) together imply that $q \vdash_{\mathbf{L}} p \to q$, and so $\to$ is an implication for $\mathbf{L}$ by Lemma 3.6.
- Assume that $\supset$ is an implication for $\mathbf{L}$. Then $p, p \supset q \vdash_{\mathbf{L}} q$ (because $p \supset q \vdash_{\mathbf{L}} p \supset q$). Now $p \supset q \nvdash_{\mathbf{L}} q$ (otherwise we would get $p \supset p \vdash_{\mathbf{L}} p$ as a special case, and so that $\vdash_{\mathbf{L}} p$). Therefore, the RDP for $\to$ implies that $p \supset q \vdash_{\mathbf{L}} p \to q$. Since $q \vdash_{\mathbf{L}} p \supset q$ (because $q, p \vdash_{\mathbf{L}} q$ and $\supset$ is an implication for $\mathbf{L}$) it follows that $q \vdash_{\mathbf{L}} p \to q$. Hence $\to$ is an implication for $\mathbf{L}$ by Lemma 3.6. □

Note 3.8 It is easy to see that $\vdash_{\mathbf{L}} p \to (q \to q)$ in case $\to$ is an implication for $\mathbf{L}$. Therefore, it follows from the last theorem that if $\to$ has in $\mathbf{L}$ both the RDP and the VSP (see footnote 2) then $\mathbf{L}$ has no conjunction, no disjunction, and no implication. This is the reason why the intensional fragments of the relevance logic $\mathbf{R}$ have no conjunction, disjunction, or implication, while the entailment connective $\to$ of $\mathbf{R}$ itself does not have the RDP.

4 The Minimal Logics with the RDP

Our next goal is to characterize the minimal logics which has a connectives $\to$ having the RDP:

Definition 4.1 Let $\mathcal{L} = \{\to\}$.

1. $HR_{\to}$ is the Hilbert-type system obtained from $HLL_{\to}$ by adding to it the following axiom:

$$[\text{Ct}] \quad \big(\varphi \to (\varphi \to \psi)\big) \to (\varphi \to \psi) \quad \text{(Contraction)}.$$

   Let $\mathbf{R}_{\to}$ be the logic induced by $HR_{\to}$.
2. $HCL_{\leftrightarrow}$ is the Hilbert-type system obtained from $HLL_{\to}$ by adding to it the following axiom:

$$[\text{Eq}] \quad \big(\varphi \to (\varphi \to \psi)\big) \to \psi \quad \text{(Equivalence)}.$$

   Let $\mathbf{CL}_{\leftrightarrow}$ be the logic induced by $HCL_{\leftrightarrow}$.

Theorem 4.2 *Let* $\mathbf{L}$ *be a logic in a language which contains* $\to$*, and suppose that* $\to$ *has in* $\mathbf{L}$ *the* RDP. *Then* $\mathbf{L}$ *is an extension of either* $\mathbf{R}_{\to}$ *or* $\mathbf{CL}_{\leftrightarrow}$.

Proof By Theorem 3.5, $\mathbf{L}$ contains $\mathbf{LL}_{\to}$. Now the validity in $\mathbf{L}$ of [MP] for $\to$ entails that $p \to (p \to q), p \vdash_{\mathbf{L}} q$. By the RDP, this in turn implies that either $p \to (p \to q) \vdash_{\mathbf{L}} p \to q$, or $p \to (p \to q) \vdash_{\mathbf{L}} q$. Since neither $\vdash_{\mathbf{L}} q$ nor $\vdash_{\mathbf{L}} p \to q$ (because $p \nvdash_{\mathbf{L}} q$ by our definition of a *logic*), by the RDP we get that either $\vdash_{\mathbf{L}} (p \to (p \to q)) \to (p \to q)$ or $\vdash_{\mathbf{L}} (p \to (p \to q)) \to q$. Hence either [Ct] or [Eq] is valid in $\mathbf{L}$, implying that it is an extension of either $\mathbf{R}_{\to}$ or $\mathbf{CL}_{\leftrightarrow}$. □

Note 4.3 From Theorem 5.5 below it easily follows that a logic cannot be an extension of both $\mathbf{R}_{\to}$ and $\mathbf{CL}_{\leftrightarrow}$.

Corollary 4.4 *Let* $\mathbf{L}$ *be a finitary logic, and suppose that* $\to$ *has in* $\mathbf{L}$ *the* RDP. *Then* $\mathbf{L}$ *has a strongly sound and complete Hilbert-type system which is an extension by axiom schemas of either* $HR_{\to}$ *or* $HCL_{\leftrightarrow}$.

Proof Let HL be the Hilbert-type which has [MP] for $\to$ as its sole rule of inference, and every theorem of $\mathbf{L}$ as an axiom. By Theorem 4.2 and its proof, HL is an extension by axiom schemas of either $HR_{\to}$ or $HCL_{\leftrightarrow}$, and is strongly sound for $\mathbf{L}$. To show that it is also strongly complete, assume that $\mathcal{T} \vdash_{\mathbf{L}} \varphi$. Since $\mathbf{L}$ is finitary, there is a finite subset $\Gamma \subseteq \mathcal{T}$ such that $\Gamma \vdash_{\mathbf{L}} \varphi$. Let $\Gamma = \{\psi_1, \ldots, \psi_n\}$ be a minimal subset of $\mathcal{T}$ with this property. Then the RDP, the minimality of Γ, and the validity of [MP] in $\mathbf{L}$ together entail that $\vdash_{\mathbf{L}} \psi_1 \to (\psi_2 \to (\cdots \to (\psi_n \to \varphi)\cdots))$. Hence this formula is an axiom of HL. Using [MP], this fact immediately implies that $\mathcal{T} \vdash_{HL} \varphi$. □

The converse of Corollary 4.4 also holds:

Theorem 4.5 *Let $\mathcal{H}$ be an extension by some axioms schemes of either $HR_{\rightarrow}$ or $HCL_{\leftrightarrow}$. Then the logic induced by $\mathcal{H}$ has the* RDP.[4]

Proof The "if" direction is trivial. The converse is proved by induction on the length of the proof of ψ from $\mathcal{T} \cup \{\varphi\}$. If ψ is an axiom of $\mathcal{H}$ or $\psi \in \mathcal{T}$ then $\mathcal{T} \vdash_{\mathcal{H}} \psi$. If $\psi = \varphi$ then $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow \psi$ by axiom [Id]. Finally, if ψ was inferred from θ and $\theta \rightarrow \psi$ then by the induction hypothesis there are four cases to consider:

1. If $\mathcal{T} \vdash_{\mathcal{H}} \theta$ and $\mathcal{T} \vdash_{\mathcal{H}} \theta \rightarrow \psi$ then $\mathcal{T} \vdash_{\mathcal{H}} \psi$.
2. If $\mathcal{T} \vdash_{\mathcal{H}} \theta$ and $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow (\theta \rightarrow \psi)$ then $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow \psi$ using axiom [Pe].
3. If $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow \theta$ and $\mathcal{T} \vdash_{\mathcal{H}} \theta \rightarrow \psi$ then $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow \psi$ using axiom [Tr].
4. Suppose $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow \theta$ and $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow (\theta \rightarrow \psi)$. Then, using axioms [Pe] and [Tr], we get that $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow (\varphi \rightarrow \psi)$. It follows that if $\mathcal{H}$ extends $HR_{\rightarrow}$ then $\mathcal{T} \vdash_{\mathcal{H}} \varphi \rightarrow \psi$ by axiom [Ct], while if $\mathcal{H}$ extends $HCL_{\leftrightarrow}$ then $\mathcal{T} \vdash_{\mathcal{H}} \psi$ by axiom [Eq]. □

The results of this section are summarized in the following theorem:

Theorem 4.6 *A logic* **L** *is finitary and has a connective $\rightarrow$ which has in* **L** *the* RDP *iff* **L** *has a strongly sound and complete Hilbert-type system which is an extension by axiom schemas of either $HR_{\rightarrow}$ or $HCL_{\leftrightarrow}$.*

Proof Immediate from Corollary 4.4 and Theorem 4.5. □

5 $\mathbf{CL}_{\leftrightarrow}$ and the Classical Equivalence

Our next goal is to show that the logic $\mathbf{CL}_{\leftrightarrow}$ is actually the logic of the classical equivalence connective. We start with the following lemma:

Lemma 5.1

1. $\vdash_{HCL_{\leftrightarrow}} (\varphi \rightarrow \psi) \rightarrow (\psi \rightarrow \varphi)$;
2. $\vdash_{HCL_{\leftrightarrow}} \varphi \rightarrow (\psi \rightarrow (\varphi \rightarrow \psi))$.

Proof

1. By substituting in axiom [Eq] $\psi \rightarrow \varphi$ for φ and $(\varphi \rightarrow \psi) \rightarrow (\psi \rightarrow \varphi)$ for ψ, we get that $\vdash_{HCL_{\leftrightarrow}} ((\psi \rightarrow \varphi) \rightarrow ((\psi \rightarrow \varphi) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\psi \rightarrow \varphi)))) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\psi \rightarrow \varphi))$. But it is easy to see that $\vdash_{\mathbf{LL}_{\rightarrow}} (\psi \rightarrow \varphi) \rightarrow ((\psi \rightarrow \varphi) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\psi \rightarrow \varphi)))$. The last two facts imply that $\vdash_{HCL_{\leftrightarrow}} (\varphi \rightarrow \psi) \rightarrow (\psi \rightarrow \varphi)$
2. By [Id] and [Pe], (i) $\vdash_{HCL_{\leftrightarrow}} \varphi \rightarrow ((\varphi \rightarrow \psi) \rightarrow \psi)$, while by the first item of this lemma $\vdash_{HCL_{\leftrightarrow}} ((\varphi \rightarrow \psi) \rightarrow \psi) \rightarrow (\psi \rightarrow (\varphi \rightarrow \psi))$. That $\vdash_{HCL_{\leftrightarrow}} \varphi \rightarrow (\psi \rightarrow (\varphi \rightarrow \psi))$ follows from (i) and (ii) using [Tr]. □

Theorem 5.2 *$HCL_{\leftrightarrow}$ is strongly sound and complete for the equivalence fragment of classical logic (i.e. $\mathcal{T} \vdash_{HCL_{\leftrightarrow}} \varphi$ iff by interpreting $\rightarrow$ as the classical biconditional $\leftrightarrow$, we get that every assignment that satisfies $\mathcal{T}$ also satisfies φ).*

[4] For the extensions of $HR_{\rightarrow}$ this theorem was first proved in [8].

Proof Obviously, [MP] is a valid rule of inference for the classical biconditional $\leftrightarrow$. It is also easy to check that every axiom of $HCL_{\leftrightarrow}$ becomes a classical tautology if $\to$ is interpreted as the classical biconditional.[5] Hence $HCL_{\leftrightarrow}$ is strongly sound for the equivalence fragment of classical logic.

To prove strong completeness, assume that $\mathcal{T} \nvdash_{HCL_{\leftrightarrow}} \theta$. We show that θ does not follow from $\mathcal{T}$ in the classical equivalence logic. For this, extend $\mathcal{T}$ to a maximal theory $\mathcal{T}^*$ such that $\mathcal{T}^* \nvdash_{HCL_{\leftrightarrow}} \theta$. Obviously, $\varphi \in \mathcal{T}^*$ iff $\mathcal{T}^* \vdash_{HCL_{\leftrightarrow}} \varphi$, and $\varphi \notin \mathcal{T}^*$ iff $\mathcal{T}^*, \varphi \vdash_{HCL_{\leftrightarrow}} \theta$. Therefore, the RDP implies that

$$(*) \quad \varphi \notin \mathcal{T}^* \quad \text{iff } \varphi \to \theta \in \mathcal{T}^*.$$

Now define a valuation v as follows:

$$v(\varphi) = \begin{cases} t & \text{if } \varphi \in \mathcal{T}^*, \\ f & \text{if } \varphi \notin \mathcal{T}^*. \end{cases}$$

We show if $\to$ is interpreted as the classical biconditional (i.e. $v(\varphi \to \psi) = t$ iff $v(\varphi) = v(\psi)$), then v is a legal classical valuation.

- Suppose $v(\varphi) = v(\psi) = t$. Then $\varphi \in \mathcal{T}^*$ and $\psi \in \mathcal{T}^*$. Therefore, it follows from the second item of Lemma 5.1 that $\mathcal{T}^* \vdash_{HCL_{\leftrightarrow}} \varphi \to \psi$. Hence $\varphi \to \psi \in \mathcal{T}^*$, and so $v(\varphi \to \psi) = t$.
- Suppose $v(\varphi) = t$ and $v(\psi) = f$. Then $\varphi \in \mathcal{T}^*$, while $\psi \notin \mathcal{T}^*$. Because of the presence of [MP], these facts immediately imply that $\varphi \to \psi \notin \mathcal{T}^*$, and so $v(\varphi \to \psi) = f$ in this case.
- Suppose $v(\varphi) = f$ and $v(\psi) = t$. By the previous item, this implies that $\psi \to \varphi \notin \mathcal{T}^*$. Therefore, the first item of Lemma 5.1 implies that $\varphi \to \psi \notin \mathcal{T}^*$, and so $v(\varphi \to \psi) = f$ in this case, too.
- Suppose $v(\varphi) = v(\psi) = f$. Then $\varphi \notin \mathcal{T}^*$ and $\psi \notin \mathcal{T}^*$. By (*) above it follows that $\varphi \to \theta \in \mathcal{T}^*$ and $\psi \to \theta \in \mathcal{T}^*$. By the first item of Lemma 5.1, the second fact implies that $\theta \to \psi \in \mathcal{T}^*$. Using [Tr] this last fact and the fact that $\varphi \to \theta \in \mathcal{T}^*$ together imply that $\varphi \to \psi \in \mathcal{T}^*$, and so $v(\varphi \to \psi) = t$ in this case.

Since $\mathcal{T} \subseteq \mathcal{T}^*$ and $\theta \notin \mathcal{T}^*$, it follows that if $\to$ is interpreted as the classical biconditional then v is a model of $\mathcal{T}$ which is not a model of θ. Hence θ does not follow from $\mathcal{T}$ in the classical equivalence logic. □

Corollary 5.3 *The pure $\leftrightarrow$-fragment of classical logic has the RDP for $\leftrightarrow$.*[6]

Proof This is immediate from Theorems 4.5 and 5.2. □

[5] This is particularly easy using Leśniewski famous criterion (see, e.g. Corollary 7.31.7 in [12]), according to which a formula built up using $\leftrightarrow$ as the sole connective is a classical tautology iff every propositional atom occurs in the formula an even number of times.

[6] This corollary was first proved by Surma (see [18]). A direct semantic proof of it can be found in 7.31.3 of [12].

Note 5.4

1. Axiom [Id] can easily be derived from the other axioms: it suffices to substitute in [Eq] $\varphi \to \varphi$ for both φ and ψ, and then use [Tr] and [MP]. It follows that [Id] can be dropped from the list of axioms of $HCL_{\leftrightarrow}$, leaving a system with just three axioms.

   Now in [18] several other axiomatizations of $\mathbf{CL}_{\leftrightarrow}$ are presented, with just two axioms and even a single axiom. However, none of them is a direct extension by axioms of $\mathbf{LL}_{\to}$.
2. Instead of [Eq], we could have adopted as an axiom the schema given in the first item of Lemma 5.1: $(\varphi \to \psi) \to (\psi \to \varphi)$. Indeed, [Eq] can easily be derived using [Tr] and [MP] from the following four schemas:

   (a) $(\varphi \to (\varphi \to \psi)) \to (\varphi \to (\psi \to \varphi))$;
   (b) $(\varphi \to (\psi \to \varphi)) \to (\psi \to (\varphi \to \varphi))$;
   (c) $(\psi \to (\varphi \to \varphi)) \to ((\varphi \to \varphi) \to \psi)$;
   (d) $((\varphi \to \varphi) \to \psi) \to \psi$.

   Now (b) and (d) are easily seen to be theorems of $\mathbf{LL}_{\to}$, (c) is an instance of the schema $(\varphi \to \psi) \to (\psi \to \varphi)$, while (a) easily follows from this schema in $\mathbf{LL}_{\to}$.

A property of $\mathbf{CL}_{\leftrightarrow}$ which is particularly important in the present context is given in the next theorem:

Theorem 5.5 $\mathbf{CL}_{\leftrightarrow}$ *is strongly Post-complete: it has no proper extension in its language.*[7]

Proof Suppose That $\mathbf{L}$ is a logic in the language of $\{\to\}$ which properly extends $\mathbf{CL}_{\leftrightarrow}$. Then there is a theory $\mathcal{T}$ and a sentence φ in this language such that $\mathcal{T} \vdash_{\mathbf{L}} \varphi$ but $\mathcal{T} \nvdash_{\mathbf{CL}_{\leftrightarrow}} \varphi$. The latter implies that there is a classical valuation v which assigns t to every element of $\mathcal{T}$, while $v(\varphi) = f$. Let p be some propositional atom of φ such that $v(p) = f$ (such p exists, since $v(\varphi) = f$). Obtain $\mathcal{T}'$ and φ' from $\mathcal{T}$ and φ (respectively) by substituting $p \to p$ for every atom q such that $v(q) = t$, and p for every atom q such that $v(q) = f$. Then $\mathcal{T}' \vdash_{\mathbf{L}} \varphi'$ (since $\vdash_{\mathbf{L}}$ is structural), while $\mathcal{T}' \nvdash_{\mathbf{CL}_{\leftrightarrow}} \varphi'$ (because any valuation u such that $u(p) = f$ is a model of $\mathcal{T}'$ which is not a model of φ'). Now if $\psi \in \mathcal{T}'$ then $\mathsf{Atoms}(\psi) = \{p\}$, and $u(\psi) = t$ in case $u(p) = f$. Since $u(\psi) = t$ also when $u(p) = t$ (because $\mathsf{Atoms}(\psi) = \{p\}$ and we interpret $\to$ as the classical equivalence), it follows that every $\psi \in \mathcal{T}'$ is a classical tautology, and so $\vdash_{\mathbf{CL}_{\leftrightarrow}} \psi$ by Theorem 5.2. Hence $\vdash_{\mathbf{L}} \psi$ for every $\psi \in \mathcal{T}'$, and so $\vdash_{\mathbf{L}} \varphi'$ as well. However, φ' is a formula such that $\mathsf{Atoms}(\varphi') = \{p\}$, and $u(\varphi') = f$ in case $u(p) = f$. It is easy to see that in this case $\varphi' \to p$ is a classical tautology (when $\to$ is interpreted as the classical equivalence), and so $\vdash_{\mathbf{CL}_{\leftrightarrow}} \varphi' \to p$ by Theorem 5.2. Hence $\vdash_{\mathbf{L}} \varphi' \to p$ as well, and since $\vdash_{\mathbf{L}} \varphi'$ too, it follows that $\vdash_{\mathbf{L}} p$ where p is atomic. A contradiction. □

[7] This is a slight generalization of a theorem of Prior, who showed in [14] (p. 307) that $\mathbf{CL}_{\leftrightarrow}$ is Post-complete in the sense that one cannot add any new axiom to it in its language.

6 The Notion of Semi-implication

Having the RDP seems at first to be an appropriate criterion for a notion of semi-implication that generalizes the classical-intuitionistic notion of implication, and is adequate also from the relevance point of view. However, the fact that $\rightarrow$ has this property in $\mathbf{CL}_{\leftrightarrow}$ means that this criterion is not sufficient: intuitively, $\mathbf{CL}_{\leftrightarrow}$ seems strange as a logic of an entailment connective, since we have that $\psi \rightarrow \varphi$ is valid in it whenever $\varphi \rightarrow \psi$ is. Now Theorems 4.6 and 5.5 imply that $\mathbf{CL}_{\leftrightarrow}$ is indeed a singular point among the logics in the language of $\{\rightarrow\}$ in which $\rightarrow$ has the RDP. All other such logics are extensions of $\mathbf{R}_{\rightarrow}$ (and unlike $\mathbf{CL}_{\leftrightarrow}$, there are infinitely many such extensions—see Note 6.8 below). Accordingly, the notion of semi-implication that we introduced next is designed to exclude this singular point by forcing one more, quite natural, condition:

Definition 6.1 (Semi-implication) Let $\mathbf{L} = \langle \mathcal{L}, \vdash_{\mathbf{L}} \rangle$ be a propositional logic. A (primitive or defined) binary connective $\rightarrow$ of $\mathcal{L}$ is called a *semi-implication for* $\mathbf{L}$ if $\rightarrow$ has in $\mathbf{L}$ the RDP, and in addition there are formulas φ and ψ such that $\vdash_{\mathbf{L}} \varphi \rightarrow \psi$ but $\nvdash_{\mathbf{L}} \psi \rightarrow \varphi$.

Note 6.2 Like implication, a semi-implication for $\mathbf{L}$ (or even just a connective which has in $\mathbf{L}$ the RDP) reflects the consequence relation of $\mathbf{L}$, but it does it in a more complicated way: $\psi_1, \ldots, \psi_n \vdash_{\mathbf{L}} \varphi$ iff for some subset $\{\theta_1, \ldots, \theta_k\}$ of $\{\psi_1, \ldots, \psi_n\}$ it holds that $\vdash_{\mathbf{L}} \theta_1 \rightarrow (\theta_2 \rightarrow (\cdots \rightarrow (\theta_k \rightarrow \varphi) \cdots))$.

Proposition 6.3 *Every implication for* $\mathbf{L}$ *is also a semi-implication for* $\mathbf{L}$.

Proof Let $\rightarrow$ be an implication for $\mathbf{L}$. Obviously, $\rightarrow$ has in $\mathbf{L}$ the RDP. To show that it satisfies also the second condition in the definition of a semi-implication, suppose for contradiction that $\vdash_{\mathbf{L}} \psi \rightarrow \varphi$ whenever $\vdash_{\mathbf{L}} \varphi \rightarrow \psi$. Since $\vdash_{\mathbf{L}} p \rightarrow (q \rightarrow q)$ (because $p, q \vdash_{\mathbf{L}} q$), this implies that $\vdash_{\mathbf{L}} (q \rightarrow q) \rightarrow p$. But $\vdash_{\mathbf{L}} q \rightarrow q$ as well (because $q \vdash_{\mathbf{L}} q$). It follows that $\vdash_{\mathbf{L}} p$, contradicting the fact that $\mathbf{L}$ is a logic. □

Our next result shows that $\mathbf{R}_{\rightarrow}$ has with respect to semi-implications the same role that $\mathbf{H}_{\rightarrow}$ has with respect to implications (see Note 2.6):

Theorem 6.4 *A logic* $\mathbf{L}$ *is finitary and has a semi-implication connective* $\rightarrow$ *iff* $\mathbf{L}$ *has a strongly sound and complete Hilbert-type system which is an extension by axiom schemas of* $HR_{\rightarrow}$.

Proof That if $\mathbf{L}$ is finitary and has a semi-implication then it has such a Hilbert-type system follows from Theorem 4.6 and Lemma 5.1. For the converse, assume that $\mathbf{L}$ has such a Hilbert-type system. Then it is finitary, and by Theorem 4.6 it also has the RDP. To end the proof, assume for contradiction that $\vdash_{\mathbf{L}} \psi \rightarrow \varphi$ whenever $\vdash_{\mathbf{L}} \varphi \rightarrow \psi$. Since $\vdash_{HR_{\rightarrow}} (p \rightarrow (p \rightarrow p)) \rightarrow (p \rightarrow p)$, this implies that $\vdash_{\mathbf{L}} (p \rightarrow p) \rightarrow (p \rightarrow (p \rightarrow p))$, and so that $\vdash_{\mathbf{L}} p \rightarrow (p \rightarrow p)$ (because $\vdash_{HR_{\rightarrow}} p \rightarrow p$). In turn, this fact and our assumption imply that $\vdash_{\mathbf{L}} (p \rightarrow p) \rightarrow p$, and so that $\vdash_{HR_{\rightarrow}} p$. A contradiction. □

Note 6.5 The last theorem means that $\mathbf{R}_{\rightarrow}$ is *the minimal* logic which has a semi-implication. It is interesting to note that in [6] it is shown that every extension of $\mathbf{R}_{\rightarrow}$

in its language is contained in $\mathbf{CL}_{\rightarrow}$, the implicational fragment of classical logic. Thus $\mathbf{CL}_{\rightarrow}$ is *the maximal* logic in the language of $\{\rightarrow\}$ in which $\rightarrow$ is a semi-implication.

Corollary 6.6 *If a connective $\rightarrow$ has in a finitary logic* $\mathbf{L}$ *both the VSP and the RDP, then it is a semi-implication for* $\mathbf{L}$.

Proof Since $\rightarrow$ has in $\mathbf{L}$ the RDP, it suffices by Theorems 4.6 and 6.4 to show that $\mathbf{L}$ is not an extension of $\mathbf{CL}_{\leftrightarrow}$. But from Theorem 5.2 it follows that $\vdash_{\mathbf{CL}_{\leftrightarrow}} (p \rightarrow p) \rightarrow (q \rightarrow q)$. Hence no extension of $\mathbf{CL}_{\leftrightarrow}$ can have the VSP. □

Note 6.7 From Proposition 6.3 and Corollary 6.6 it follows that the notion of a semi-implication is a common generalization of the notion of implication (as defined above) and the notion of strong relevant implication defined in [7].

Note 6.8 In [2], one can find an infinite family of logics in the language of $\{\rightarrow\}$, all of which has the VSP and are extensions of $HR_{\rightarrow}$ by axiom schemas. There are, of course, in the language of $\{\rightarrow\}$ also many extensions of $HR_{\rightarrow}$ by axiom schemas which do not have the VSP, like the implicational fragments of intuitionistic logic, classical logic, and any intermediate logic.

7 Semi-implication in Special Types of Matrices

Theorem 3.7 shows that in many cases a semi-implication is necessarily an implication. In this section, we describe some very common semantic circumstances in which the same phenomenon occurs. We start by recalling some basic definitions concerning multiple-valued logics.

Definition 7.1 Let $\mathcal{L}$ be a propositional language.

1. A *matrix* for $\mathcal{L}$ is a triple $\mathcal{M} = \langle \mathcal{V}, \mathcal{D}, \mathcal{O} \rangle$, where
   - $\mathcal{V}$ is a non-empty set of truth values;
   - $\mathcal{D}$ is a non-empty proper subset of $\mathcal{V}$ (whose elements are called the *designated* elements of $\mathcal{V}$);
   - $\mathcal{O}$ is a function that associates an n-ary function $\widetilde{\diamond}_{\mathcal{M}} : \mathcal{V}^n \rightarrow \mathcal{V}$ with every n-ary connective $\diamond$ of $\mathcal{L}$.
2. An $\mathcal{M}$-*valuation* for $\mathcal{L}$ is a function $\nu : \mathcal{W}(\mathcal{L}) \rightarrow \mathcal{V}$ such that for every $\psi_1, \ldots, \psi_n$: $\nu(\diamond(\psi_1, \ldots, \psi_n)) = \widetilde{\diamond}_{\mathcal{M}}(\nu(\psi_1), \ldots, \nu(\psi_n))$.
3. An $\mathcal{M}$-valuation ν is an $\mathcal{M}$-*model* of a formula ψ, or ν $\mathcal{M}$-*satisfies* ψ (notation, $\nu \models_{\mathcal{M}} \psi$), if $\nu(\psi) \in \mathcal{D}$. ν is an $\mathcal{M}$-*model* of a theory $\mathcal{T}$ (notation, $\nu \models_{\mathcal{M}} \mathcal{T}$), if it is an $\mathcal{M}$-model of every element of $\mathcal{T}$.
4. The consequence relation $\vdash_{\mathcal{M}}$ which is induced by (or is associated with) a matrix $\mathcal{M}$ is defined by: $\mathcal{T} \vdash_{\mathcal{M}} \psi$ if $\nu \models_{\mathcal{M}} \psi$ whenever ν is an $\mathcal{M}$-valuation such that $\nu \models_{\mathcal{M}} \mathcal{T}$.
5. If $\mathcal{M}$ is a matrix for $\mathcal{L}$ then $\mathbf{L}_{\mathcal{M}} = \langle \mathcal{L}, \vdash_{\mathcal{M}} \rangle$ is the *logic induced by* $\mathcal{M}$.

Next we introduced two very important special types of matrices.

Definition 7.2 Let $\mathcal{M} = \langle \mathcal{V}, \mathcal{D}, \mathcal{O} \rangle$ be a matrix for $\mathcal{L}$. $\mathcal{M}$ is called a *t-matrix* if $\mathcal{D}$ is a singleton, in which case we denote the single designated element by t. $\mathcal{M}$ is called an *f-matrix* if $\mathcal{V} - \mathcal{D}$ is a singleton, in which case we denote the single non-designated element by f.

Note 7.3 At the beginning of the study of many-valued logics all matrices that were considered (like Łuksiewicz' various matrices, the three-valued matrices of Kleene and Bochvar, Gödel matrices, etc.) were t-matrices, and these matrices are still the most studied class of matrices. In paraconsistent logics, on the other hand, the use of f-matrices is very common. A general extensive study of t-matrices and f-matrices has been done in [4, 5].

Theorem 7.4 *Let $\mathcal{M}$ be a matrix for $\mathcal{L}$ which is either a t-matrix or an f-matrix, and suppose that $\to$ is a semi-implication for $\mathbf{L}_{\mathcal{M}}$. Then $\to$ is an implication for $\mathbf{L}_{\mathcal{M}}$.*

Proof Let $\mathcal{M} = \langle \mathcal{V}, \mathcal{D}, \mathcal{O} \rangle$, and assume that $\to$ is a semi-implication for $\mathbf{L}_{\mathcal{M}}$. By Theorem 6.4, $\mathbf{L}_{\mathcal{M}}$ is an extension of $\mathbf{R}_{\to}$.[8] Therefore, for any two atomic formulas p and q we have:

(a) $p, p \to q \vdash_{\mathcal{M}} q$
(b) $\vdash_{\mathcal{M}} p \to p$
(c) $\vdash_{\mathcal{M}} p \to ((p \to q) \to q)$
(d) $\vdash_{\mathcal{M}} (p \to (p \to q)) \to (p \to q)$
(e) $\vdash_{\mathcal{M}} (p \to (q \to r)) \to (q \to (p \to r))$
(f) $\vdash_{\mathcal{M}} (p \to q) \to ((q \to r) \to (p \to r))$

Obviously, (a) implies that $(*)$ $b \in \mathcal{D}$ in case $a \in \mathcal{D}$ and $a \tilde{\to} b \in \mathcal{D}$. This and (f) together imply that $(**)$ $a \tilde{\to} c \in \mathcal{D}$ whenever $a \tilde{\to} b \in \mathcal{D}$ and $b \tilde{\to} c \in \mathcal{D}$.

In the rest of the proof, we separately deal with each of the two cases considered in the formulation of the theorem.

1. Assume that $\mathcal{M}$ is a t-matrix. Let $a \in \mathcal{V}$. Then (b) implies that $a \tilde{\to} a = t$, while (d) implies that $(a \tilde{\to} (a \tilde{\to} a)) \tilde{\to} (a \tilde{\to} a) = t$. Together these two facts imply that (i) $(a \tilde{\to} t) \tilde{\to} t = t$. On the other hand (c) implies that (ii) $a \tilde{\to} ((a \tilde{\to} t) \tilde{\to} t) = t$. Now from (i) and (ii) it follows that $a \tilde{\to} t = t$ for every $a \in \mathcal{V}$. Hence $q \vdash_{\mathcal{M}} p \to q$. By Lemma 3.6, this fact implies that $\to$ is an implication for $\mathbf{L}_{\mathcal{M}}$.
2. Assume that $\mathcal{M}$ is an f-matrix. In this case, (*) implies that (i) $a \tilde{\to} f = f$ for $a \in \mathcal{D}$, while (b) implies that (ii) $a \tilde{\to} a \in \mathcal{D}$ for every $a \in \mathcal{V}$. In turn, these two facts together imply that $(f \tilde{\to} f) \tilde{\to} f = f$. Therefore, had there been some $a \in \mathcal{V}$ such that $f \tilde{\to} a = f$ then we would have that $(f \tilde{\to} (f \tilde{\to} a)) \tilde{\to} (f \tilde{\to} a) = f$. However, this contradicts (d). It follows that (iii) $f \tilde{\to} a \in \mathcal{D}$ for every $a \in \mathcal{V}$. Now (iii) and (i) together imply that $(p \to q) \to p \vdash_{\mathcal{M}} p$. Since $\to$ is a semi-implication for $\mathbf{L}_{\mathcal{M}}$, this implies that $\vdash_{\mathcal{M}} ((p \to q) \to p) \to p$. It follows that $((b \tilde{\to} f) \tilde{\to} b) \tilde{\to} b \in \mathcal{D}$ for every $b \in \mathcal{V}$, and so by (i) we have: (iv) $(f \tilde{\to} b) \tilde{\to} b \in \mathcal{D}$ for every $b \in \mathcal{D}$. On the other hand,

[8]Theorem 6.4 is about finitary logics, but its proof shows that being an extension of $\mathbf{R}_{\to}$ is a property of *every* logic with a semi-implication. It is also worth noting that by a theorem of Shoesmith and Smiley (see [15, 16]), $\mathbf{L}_{\mathcal{M}}$ is finitary whenever $\mathcal{M}$ is finite.

(iii) implies that $f\tilde{\to}(a\tilde{\to}b) \in \mathcal{D}$ for every a and b, and so by (e) and (*) we have: (v) $a\tilde{\to}(f\tilde{\to}b) \in \mathcal{D}$. Now from (v), (iv), and (**) it follows that $a\tilde{\to}b \in \mathcal{D}$ whenever $b \in D$. Hence $q \vdash_{\mathcal{M}} p \to q$. Again by Lemma 3.6, this fact implies that $\to$ is an implication for $\mathbf{L}_{\mathcal{M}}$. □

We end with an example of a matrix $\mathcal{M}$ such that $\mathbf{L}_{\mathcal{M}}$ has a semi-implication which is not an implication. By the last theorem, this matrix has, of course, more than one designated element and more than one non-designated element.

Example 7.5 Let $S_{[0,1]}$ be the standard $\aleph$-Sugihara matrix for the language $\{\neg, \to\}$, i.e. $S_{[0,1]} = \langle [0,1], \{x \in [0,1] \mid x \geq 1/2\}, \{\tilde{\neg}, \tilde{\to}\}\rangle$, where $[0,1]$ is the unit interval, $\tilde{\neg}a = 1 - a$, and $a\tilde{\to}b = \max(1-a, b)$ if $a \leq b$, $\min(1-a, b)$ if $a > b$. It is well-known that $\mathbf{RM}_{\substack{\neg\\\to}}$, the implication-negation fragment of the semi-relevant logic **RM**, is strongly sound and complete for $S_{[0,1]}$, and that this fragment can be axiomatized by adding to $HR_{\to}$ the two axioms for negation of **R** (see [1, 10]), as well as the mingle axiom $\varphi \to (\varphi \to \varphi)$.[9] Therefore, it follows from Theorem 6.4 that $\to$ is a semi-implication for $\mathbf{L}_{S_{[0,1]}}$. Since $q \nvdash_{S_{[0,1]}} p \to q$ (take, e.g. $\nu(q) = 1/2$, $\nu(p) = 1$), by Lemma 3.6, $\to$ is not an implication for this logic. At this point, it is also worth noting that in [13] it was shown that $\mathbf{RM}_{\substack{\neg\\\to}}$ is *weakly* sound and complete for Sobociński three-valued logic (see [17]),[10] which is an f-matrix. This fact does not contradict, of course, Theorem 7.4 because $\mathbf{RM}_{\substack{\neg\\\to}}$ is *not* strongly complete for Sobociński's matrix.

An Open Problem. The matrix used in the above example is *infinite*, and so are all other examples I know of matrices which induce logics that have semi-implications which are not implications. Are there also finite matrices with this property?

References

1. Anderson, A.R., Belnap, N.D.: Entailment: The Logic of Relevance and Necessity, vol. I. Princeton University Press, Princeton (1975)
2. Avron, A.: Relevant entailment—semantics and formal systems. J. Symb. Log. **49**, 334–342 (1984)
3. Avron, A.: The semantics and proof theory of linear logic. Theor. Comput. Sci. **57**, 161–184 (1988)
4. Avron, A.: Multiplicative conjunction and an algebraic meaning of contraction and weakening. J. Symb. Log. **63**, 831–859 (1998)
5. Avron, A.: Implicational F-structures and implicational relevance logics. J. Symb. Log. **65**, 788–802 (2000)
6. Avron, A.: The classical constraint on relevance. Log. Univers. **8**, 1–15 (2014). doi:10.1007/s11787-013-0092-y
7. Avron, A.: What is relevance logic? Ann. Pure Appl. Log. **165**, 26–48 (2014). doi:10.1016/j.apal.2013.07.004
8. Church, A.: The weak theory of implication. In: Menne, A., Wilhelmi, A., Angsil, H. (eds.) Kontroliertes Denken: Untersuchungen zum Logikkalkül und zur Logik der Einzelwissenschften, pp. 22–37. Kommissioas-Verlag Karl Alber, Munich (1951)

[9] This is essentially proved in [9], which strengthened Meyer's result about the *weak* soundness and completeness of **RM** for the original $\aleph_0$-Sugihara matrix.

[10] This means that a sentence φ is valid in Sobociński's matrix iff $\vdash_{\mathbf{RM}_{\substack{\neg\\\to}}} \varphi$.

9. Dunn, J.M.: Algebraic completeness results for R-mingle and its extensions. J. Symb. Log. **35**, 1–13 (1970)
10. Dunn, J.M., Restall, G.: Relevance logic. In: Handbook of Philosophical Logic, 2nd edn., vol. 6, pp. 1–136. Kluwer, Dordrecht (2002)
11. Girard, J.E.: Linear logic. Theor. Comput. Sci. **50**, 1–102 (1987)
12. Humberstone, L.: The Connectives. The MIT Press, Cambridge (2011)
13. Parks, R.: A note on R-mingle and Sobociński three-valued logic. Notre Dame J. Form. Log. **13**, 227–228 (1972)
14. Prior, A.N.: Formal Logic, 2nd edn. Oxford (1962)
15. Shoesmith, D.J., Smiley, T.J.: Deducibility and many-valuedness. J. Symb. Log. **36**, 610–622 (1971)
16. Shoesmith, D.J., Smiley, T.J.: Multiple Conclusion Logic. Cambridge University Press, Cambridge (1978)
17. Sobociński, B.: Axiomatization of a partial system of three-value calculus of propositions. J. Comput. Syst. **1**, 23–55 (1952)
18. Surma, S.J.: A survey of the results and methods of investigations of the equivalential propositional calculus. In: Surma, S.J. (ed.) Studies in the History of Mathematical Logic, pp. 33–61. Polish Academy of Science, Institute of Philosophy and Sociology, Wrocław (1973)

A. Avron (✉)
School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel
e-mail: aa@tau.ac.il

A Formal Framework for Hypersequent Calculi and Their Fibring

Marcelo Esteban Coniglio and Martín Figallo

Abstract Hypersequents are a natural generalization of ordinary sequents which turn out to be a very suitable tool for presenting cut-free Gentzent-type formulations for diverse logics. In this paper, an alternative way of formulating hypersequent calculi (by introducing meta-variables for formulas, sequents and hypersequents in the object language) is presented. A suitable category of hypersequent calculi with their morphisms is defined and both types of fibring (constrained and unconstrained) are introduced. The introduced morphisms induce a novel notion of translation between logics which preserves metaproperties in a strong sense. Finally, some preservation features are explored.

Keywords Hypersequents · Fibring · Translation between logics

Mathematics Subject Classification (2000) Primary 03B22 · Secondary 03B62

1 Introduction

In recent years, the development of methods for combining logics has gained attention, and motivations came from different areas such as Philosophy and Computer Science (see, for instance, [4] and [5]). Logics presented in different ways require *ad hoc* combination techniques. In [11], the well-known method of fibring for combining modal logics was introduced. Categorical (a.k.a. algebraic) fibring, introduced in [16], is a wide and extremely useful tool for combining logics, allowing to combine a vast class of logic systems of different nature (consult [5]). In general, it is possible to define two different kinds of fibring: *unconstrained fibring*, in which there is no sharing of logic constructors from the combined logics (and so the resulting logic is a coproduct of the given logics), and *constrained fibring*, in which some constructors are shared. In categorical terms, the latter is obtained from the former by taking an appropriate quotient.

In [8], a novel category of formal sequent calculi was introduced, and both types of categorical fibring (called meta-fibring) were obtained. Two kinds of sequents were considered: commutative sequents, formed by pairs of sets of formulas (thus taking for

Dedicated to Jean-Yves Béziau on the occasion of his 50th birthday.

This paper constitutes a revised, extended and improved version from the preprint [9]. That preliminary version is described with full details in [10].

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_4

granted structural rules such as permutation and contraction) and non-commutative sequents, formed by pairs of sequences of formulas. The former are called *general assertions*, while the latter are called *general sequents*. A remarkable feature of the presentation of this approach is the use in the object language of variables for sets and sequences of formulas, respectively, besides the standard use of scheme variables for formulas. This peculiarity permits considering sequent rules with full generality, allowing to combine by fibring two calculi in a satisfactory way. Another novelty of the approach is the notion of morphism between sequent calculi proposed therein (called meta-translations), being stronger than the usual one in the sense that more meta-properties (in rough terms, sequent rules) of the consequence relation of the source logic are preserved. This is a key feature which allows the reconstruction of a given logic by means of the fibring of its fragments. An interesting possibility of generalization of the results obtained in [8] lies in the use of hypersequents instead of sequents, because of their subtleties.

Hypersequents (see [2, 15] among others) constitute a natural generalization of ordinary sequents and turn out to be a very suitable tool for presenting cut-free Gentzen-type formulations for several non-classical logics. In particular, hypersequents are well-suited for describing disjunctive properties by analytic means. The proof of cut-elimination in a (hyper)sequent calculus for a given logic is desirable, because of its important consequences, such as the consistency of the logic and interpolation properties.

This paper proposes a generalization of the work done in [8] to the richer framework of hypersequents. Thus, departing from a formal presentation of hypersequent in which meta-variables for contexts (i.e. sequents) are introduced in the object language, the fibring of such systems is defined within a suitable category of hypersequents. Besides, it is shown that fibring does not preserve, in general, rule-elimination properties such as the cut-elimination property. Finally, a brief conceptual discussion about the relevance of this approach concerning the theory of translations between logics is carried out.

2 The Category of Formal Hypersequent Calculi

This section presents the category of hypersequent calculi, generalizing the notion of *assertion calculi* introduced in [8]. An independent similar approach was developed in [7].

In what follows, we shall consider a denumerable set $\Xi = \{\xi_i : i \in \mathbb{N}\}$ of symbols called *variables of level 1* (or scheme variables); a denumerable set $\mathcal{X} = \{X_i : i \in \mathbb{N}\}$ whose elements are called *variables of level 2* (or context variables); and, finally, a denumerable set $\mathfrak{H} = \{H_i : i \in \mathbb{N}\}$ whose elements are called *variables of level 3* (or sequent variables) where these sets are pairwise disjoint.

A *propositional signature* is a denumerable family $C = \{C_n\}_{n\in\mathbb{N}}$ of pairwise disjoint denumerable sets; additionally, every C_n is disjoint with $\mathcal{H} \cup \mathcal{X} \cup \Xi$. Elements in C_n are called n-ary connectives (or *constructors*). The set of constructors of C is $|C| = \bigcup_{n\in\mathbb{N}} C_n$. The algebra of type C freely generated by Ξ is denoted by $L(C)$. Elements of $L(C)$ are called *formulas*. From now on, and for the sake of simplicity, in the examples we will refer to a signature C as the set $|C|$.

Generalizing to multisets the definition from [8] of *general assertion*, a *sequent* over a signature C is an expression of the form $(A; \Gamma \succ \Delta; B)$ where Γ and Δ are multisets of formulas in $L(C)$ and A, B are finite multisets of context variables such that $\Gamma \cup \Delta \cup A \cup B \neq \emptyset$. The set of sequents over C is denoted by $Seq(C)$.

It should be noted that a sequent (in our sense) is nothing more than an ordinary commutative sequent, enriched with both variables of type 2 for sets of formulas (describing the context of the sequent) and variables of type 1 for formulas. This formalism allows defining in a precise way sequent calculi and their fibring (cf. [8]). Now we shall introduce the notion of commutative hypersequent by using sequent variables, that is, variables of type 3.

Definition 2.1 A *commutative hypersequent* h over C is a pair $h = \langle \mathcal{H}; \mathcal{S} \rangle$ where $\mathcal{H}$ is a finite multiset of sequent variables and $\mathcal{S}$ is a finite multiset of sequents. The set of commutative hypersequents over C will be denoted by $HSeq(C)$.

Following the usual notation, a commutative hypersequent

$$h = \big\langle \{G_1, \ldots, G_n\}; \big\{(A_1; \Gamma_1 \succ \Delta_1; B_1), \ldots, (A_m; \Gamma_m \succ \Delta_m; B_m)\big\}\big\rangle$$

will be written, from now on, as

$$h = G_1|\cdots|G_n|A_1; \Gamma_1 \succ \Delta_1; B_1|\cdots|A_m; \Gamma_m \succ \Delta_m; B_m.$$

An empty component of a sequent will be simply omitted from the notation and so we will write, for instance, $(A; \Gamma \succ \Delta)$ and $(\succ \Delta; B)$ instead of $(A; \Gamma \succ \Delta; \emptyset)$ and $(\emptyset; \emptyset \succ \Delta; B)$, respectively. As usual, Γ, Γ' and Γ, φ will stand for $\Gamma \cup \Gamma'$ and $\Gamma \cup \{\varphi\}$, respectively. Besides, we shall write X and X, Y instead of $\{X\}$ and $\{X, Y\}$, respectively, for any variables X and Y. The same notation applies to sequent variables. Moreover, given a finite multisubset $\mathcal{H}$ of $\mathfrak{H}$, the hypersequent $\langle \mathcal{H}; \emptyset \rangle$ will be simply denoted by $\mathcal{H}$. Analogously, given a finite multisubset $\mathcal{S}$ of $Seq(C)$ the hypersequent $\langle \emptyset; \mathcal{S} \rangle$ will be simply denoted by $\mathcal{S}$.

Definition 2.2 Let C be a signature. A (n-ary) rule of inference of commutative hypersequents over C is a pair $r = \langle \{h_1, \ldots, h_n\}, h \rangle$ such that $h_i, h \in HSeq(C)$. If $n = 0$ then r is called *an axiom*. A *commutative hypersequent calculus* (chc) is a pair $\mathcal{A} = \langle C, R \rangle$ where C is a signature and R is a finite set of rules of inference of commutative hypersequents over C.

For simplicity, we shall denote pairs $\langle \{h_1, \ldots, h_n\}, h \rangle$ and $\langle \emptyset, h \rangle$ by

$$\frac{h_1 \quad \ldots \quad h_n}{h} \quad \text{and} \quad \frac{}{h}.$$

Example 2.3 The logical hypersequent rule $r_{\neg\succ}$ for negation which is usually represented by

$$r_{\neg\succ} = \frac{G|\,\Gamma \vdash \Delta, \alpha}{G|\,\neg\alpha, \Gamma \vdash \Delta}$$

is here represented by

$$\frac{\langle \{G\}; \{(\{X\}; \emptyset \succ \{\xi\}; \{Y\})\}\rangle}{\langle \{G\}; \{(\{X\}; \{\neg\xi\} \succ \emptyset; \{Y\})\}\rangle},$$

or simply by

$$\frac{G|X \succ \xi; Y}{G|X; \neg\xi \succ Y}$$

(see Remark 2.5 below).

From now on, we will denote by $\mathcal{MP}_{\text{fin}}(X)$ the set of all finite multisubsets of a set X.

Recall that a substitution over a signature C is a map $\sigma : \Xi \to L(C)$. We denote by $\hat{\sigma} : L(C) \to L(C)$ the unique homomorphic extension of σ to $L(C)$. Adapting [8], an instantiation over C is a map $\varrho : \mathcal{X} \to \mathcal{MP}_{\text{fin}}(L(C) \cup \mathcal{X})$. If ϱ is an instantiation over C and $A \in \mathcal{MP}_{\text{fin}}(\mathcal{X})$ is a finite multiset of variables of type 2, we define the following finite multisets:

$$A^{\varrho}_{\mathcal{X}} = \{Y \in \mathcal{X} : Y \in \varrho(X) \text{ for some } X \in A\} = \Big(\bigcup_{X \in A} \varrho(X)\Big) \cap \mathcal{X};$$

$$A^{\varrho}_{L(C)} = \{\varphi \in L(C) : \varphi \in \varrho(X) \text{ for some } X \in A\} = \Big(\bigcup_{X \in A} \varrho(X)\Big) \cap L(C).$$

Thus, given a substitution σ and an instantiation ϱ over C, respectively, the map $(\sigma, \varrho) : Seq(C) \to Seq(C)$ is defined in the following way: given a sequent $(A; \Gamma \succ \Delta; B)$, then

$$(\sigma, \varrho)(A; \Gamma \succ \Delta; B) = \big(A^{\varrho}_{\mathcal{X}}; \hat{\sigma}(\Gamma) \cup A^{\varrho}_{L(C)} \succ \hat{\sigma}(\Delta) \cup B^{\varrho}_{L(C)}; B^{\varrho}_{\mathcal{X}}\big).$$

In order to deal with hypersequents, we shall introduce a notion of substitution for variables of level 3.

Definition 2.4 We shall say that λ is a sequent instantiation over C if λ is a mapping from $\mathfrak{H}$ to the set $\mathcal{MP}_{\text{fin}}(\mathfrak{H} \cup Seq(C))$ of all finite multisubsets of $\mathfrak{H} \cup Seq(C)$.

Let λ be a sequent instantiation and $h = \langle \mathcal{H}; \mathcal{S} \rangle$ a hypersequent over C. Consider the following multisets:

$$\mathcal{H}^{\lambda}_{\mathfrak{H}} = \{G \in \mathfrak{H} : G \in \lambda(H) \text{ for some } H \in \mathcal{H}\} = \Big(\bigcup_{H \in \mathcal{H}} \lambda(H)\Big) \cap \mathfrak{H};$$

$$\mathcal{H}^{\lambda}_{Seq(C)} = \{s \in Seq(C) : s \in \lambda(H) \text{ for some } H \in \mathcal{H}\} = \Big(\bigcup_{H \in \mathcal{H}} \lambda(H)\Big) \cap Seq(C).$$

Then, given a substitution σ, an instantiation ϱ and a sequent instantiation λ over C, respectively, the map $(\sigma, \varrho, \lambda) : HSeq(C) \to HSeq(C)$ is defined as follows:

$$(\sigma, \varrho, \lambda)(h) = \big\langle \mathcal{H}^{\lambda}_{\mathfrak{H}}; (\sigma, \varrho)(\mathcal{S}) \cup \mathcal{H}^{\lambda}_{Seq(C)} \big\rangle.$$

Remark 2.5 (Extensional rules vs. Intensional rules) Recall the Example 2.3 above. Despite the apparent similarities between the traditional notation for hypersequent rules and our notation, there are deep differences between $r_{\neg\succ}$ and our representation. In the former, G denotes an arbitrary multiset of concrete sequents; in their turn, Γ and Δ denote arbitrary multisets of formulas; finally, α denotes an arbitrary concrete formula. That is,

$r_{\neg\succ}$ consists of *infinite* concrete rules obtained by instantiation of their *metavariables*, that is, variables in the metalanguage: it is an *extensional* approach to inference rules. On the other hand, our notation is extremely precise: G, X, Y and ξ are concrete variables of the formal language of hypersequents, and so the rule is represented by a single linguistic object instead of infinite ones represented by variables in the metalanguage, as done in the traditional approach to (hyper)sequents. In other words, we propose an *intensional* approach to inference rules. Despite the obvious advantages of this fact, there is a much more important advantage of our formal approach to (hyper)sequents. Since we are interested in combining different (hyper)sequent calculi, the use of formal variables instead of metavariables (that is, informal variables) is crucial. In fact, in the intensional approach, rules are *prepared* to be combined, since they are open to accept new connectives: ξ can be replaced by *any* formula, while X and Y can be replaced by *any* multisets of formulas, as well as G is open to be substituted by *any* multiset of sequents, and this holds for *any* language. That is, if we add new connectives to the language (as a consequence of a combination process), the meaning of the rule will be the same. On the other hand, in the traditional extensional approach to (hyper)sequent rules, this possibility is no longer allowed, and the rule must be extended in order to cope with the new language. This is the main novelty of the intensional approach to combination of hypersequents, first introduced in [8] in the setting of sequent calculi.

Now we are in conditions to define the notion of *derivation* in a commutative hypersequent calculus.

Definition 2.6 Let $\mathcal{A} = \langle C, R\rangle$ be a commutative hypersequent calculus and let $\Upsilon \cup \{h\} \subseteq HSeq(C)$ be a set of hypersequents. We shall say that h is derivable in $\mathcal{A}$ from Υ, and write $\Upsilon \vdash_{\mathcal{A}} h$, if there exists a finite sequence $h_1 \dots h_n$ of elements of $HSeq(C)$ such that $h_n = h$ and, for all $1 \leq i \leq n$, either $h_i \in \Upsilon$, or there exist an hypersequent rule $r = \langle\{h'_1, \dots, h'_k\}, h'\rangle$ in R, a substitution σ, an instantiation ϱ and a sequent instantiation λ over C such that $(\sigma, \varrho, \lambda)(h'_j) \in \{h_1, \dots, h_{i-1}\}$ (for $1 \leq j \leq k$) and $(\sigma, \varrho, \lambda)(h') = h_i$. If $\Upsilon = \emptyset$ we shall just say that h is provable in $\mathcal{A}$.

Remark 2.7 The expressive power of hypersequents, allied to the possibility of using variables either of type 1, 2 or 3, allows defining structural rules in many different ways. For instance, the contraction rule, as pointed out by A. Avron, admits two versions: an internal (inside a sequent) and an external (inside the context of the hypersequent). Thus, the internal version of contraction (using variables of type 1) and the external version (using variables of type 2) are as follows:

$$\frac{G|X; \xi, \xi \succ Y}{G|X; \xi \succ Y}, \qquad \frac{G|X \succ \xi, \xi; Y}{G|X \succ \xi; Y},$$

$$\frac{G|X \succ Y|X \succ Y}{G|X \succ Y}.$$

We can add one more possibility by using variables of level 3:

$$\frac{G|G|H}{G|H}.$$

Clearly, the level 3 version can simulate the level 2 version, meanwhile each concrete application of the level 3 contraction is recovered by successive applications of the level 2 contraction. On the other hand, the internal contraction can be alternatively defined by using variables of level 2:

$$\frac{G|X, X, Y \succ Z}{G|X, Y \succ Z}, \qquad \frac{G|X \succ Y, Y, Z}{G|X \succ Y, Z}.$$

In an analogous way to the remark above, the internal contractions of level 2 and level 1 are equivalent.

Now we are going to define the category of formal commutative hypersequent calculi. Recall that in [8] the following category of signatures was used:

Definition 2.8 The category **Sig** of signatures is the category whose objects are propositional signatures. A *morphism* $f : C \to C'$ in **Sig** is a function $f : |C| \to L(C')$ such that $f(c)$ is a formula which depends at most on schema variables $\xi_1, \ldots, \xi_n$ whenever $c \in C'_n$ (in particular, $f(c) \in L(C'_0)$ if $c \in C_0$). If $f_1 : C \to C'$ and $f_2 : C' \to C''$ are two morphisms in **Sig**, the composite morphism $f_2 \circ f_1 : C \to C''$ in **Sig** is the composite map $\hat{f}_2 \circ f_1 : |C| \to L(C'')$, where the function $\hat{f}_2 : L(C') \to L(C'')$ is defined as follows:

$$\hat{f}_2(\xi) = \xi, \quad \text{for } \xi \in \Xi; \qquad \hat{f}_2(c) = f_2(c), \quad \text{for } c \in C'_0;$$

$$\hat{f}_2\big(c(\varphi_1, \ldots, \varphi_n)\big) = f_2(c)\big(\hat{f}_2(\varphi_1), \ldots, \hat{f}_2(\varphi_n)\big) \quad \text{for } c \in C'_n, n \geq 1.$$

The identity morphism $id_C : C \to C$ for the signature C is the function $id_C : |C| \to L(C)$ such that $id_C(c) = c(\xi_1, \ldots, \xi_n)$ if $c \in C_n$. In particular, $id_C(c) = c$, if $c \in C_0$.

Recall from [8] that if $(A; \Gamma \succ \Delta; B)$ is a sequent over C and $f : C \to C'$ is a signature morphism, then $\hat{f}(A; \Gamma \succ \Delta; B)$ is, by definition, the sequent $(A; \hat{f}(\Gamma) \succ \hat{f}(\Delta); B)$ over C'. This can be naturally extended to hypersequents:

$$\hat{f}\big(\langle \mathcal{H}; \mathcal{S}\rangle\big) = \big\langle \mathcal{H}; \hat{f}(\mathcal{S})\big\rangle.$$

It is clear that $\hat{f}(h)$ is a hypersequent over C' provided that h is a hypersequent over C.

The category of commutative hypersequent calculi is defined as follows.

Definition 2.9 The category **CHC** of commutative hypersequent calculi is the category whose objects are commutative hypersequent calculi of the form $\mathcal{A} = \langle C, R\rangle$. A morphism $f : \langle C, R\rangle \to \langle C', R'\rangle$ in **CHC** is a morphism $f : C \to C'$ in **Sig** such that, for every $r = \langle\{h_1, \ldots, h_n\}, h\rangle$ in R, it is verified that $\hat{f}(h)$ is derivable in $\langle C', R'\rangle$ from $\{\hat{f}(h_1), \ldots, \hat{f}(h_n)\}$. The composition of morphisms and the identity morphism in **CHC** is defined as in **Sig**.

3 Unconstrained Fibring of Hypersequent Calculi

Taking advantage of the formal framework for defining hypersequent calculi described in the previous section, we are now ready to combine these proof systems. The combination method proposed here is known in the literature as (categorical) fibring (see [11, 16]).

Basically, categorical fibring can be performed in two (related) ways: the simpler one, called *unconstrained fibring*, consists in joining up the inference rules of the two systems being combined, were the rules must be rewritten in the language generated by the free combination of the symbols of both systems. In formal terms, it is the coproduct of both systems, in the category in which they are represented. In Sect. 6, we shall study the second (and more general) way of categorical fibring, called *unconstrained fibring*, in which some connectives of the systems to be combined are shared in the resulting system.

Prior to the definition of unconstrained fibring, it is necessary to introduce some results and concepts.

Definition 3.1 Given substitutions σ, σ' and instantiations ϱ, ϱ' over C, the product $(\sigma, \varrho) \cdot (\sigma', \varrho')$ is given by

$$(\sigma, \varrho) \cdot \big(\sigma', \varrho'\big) = \big(\sigma \cdot \sigma', \big(\varrho \cdot \varrho'\big)_{\sigma}\big)$$

where $\sigma \cdot \sigma'$ is the substitution over C given by $\sigma \cdot \sigma'(\xi) = \hat{\sigma}(\sigma'(\xi))$ and $(\varrho \cdot \varrho')_{\sigma}$ is the instantiation over C given by

$$\big(\varrho \cdot \varrho'\big)_{\sigma}(X) = \big(\{X\}^{\varrho'}_{\mathcal{X}}\big)^{\varrho}_{\mathcal{X}} \cup \big(\{X\}^{\varrho'}_{\mathcal{X}}\big)^{\varrho}_{L(C)} \cup \hat{\sigma}\big(\{X\}^{\varrho'}_{L(C)}\big).$$

The proof of the following useful result is straightforward:

Proposition 3.2 *Let σ, σ' be substitutions over C and let ϱ, ϱ' be instantiations over C. Then, for every $s \in Seq(C)$,*

$$\big[(\sigma, \varrho) \cdot \big(\sigma', \varrho'\big)\big](s) = (\sigma, \varrho)\big(\big(\sigma', \varrho'\big)(s)\big).$$

In order to consider variables of level 3, we introduce the following definition.

Definition 3.3 Given substitutions σ, σ', instantiations ϱ, ϱ' and sequent instantiations λ, λ' over C, the product $(\sigma, \varrho, \lambda) \cdot (\sigma', \varrho', \lambda')$ is given by

$$(\sigma, \varrho, \lambda) \cdot \big(\sigma', \varrho', \lambda'\big) = \big(\sigma \cdot \sigma', \big(\varrho \cdot \varrho'\big)_{\sigma}, \big(\lambda \cdot \lambda'\big)_{\sigma\varrho}\big)$$

where $\sigma \cdot \sigma'$ and $(\varrho \cdot \varrho')_{\sigma}$ are as in Definition 3.1; and $(\lambda \cdot \lambda')_{\sigma\varrho}$ is the sequent instantiation given by

$$\big(\lambda \cdot \lambda'\big)_{\sigma\varrho}(H) = \big(\{H\}^{\lambda'}_{\mathfrak{H}}\big)^{\lambda}_{\mathfrak{H}} \cup \big(\{H\}^{\lambda'}_{\mathfrak{H}}\big)^{\lambda}_{Seq(C)} \cup (\sigma, \varrho)\big(\{H\}^{\lambda'}_{Seq(C)}\big).$$

Using Proposition 3.2, it is easy to prove the following result.

Proposition 3.4 *Let σ, σ' be substitutions over C, let ϱ, ϱ' be instantiations over C and let λ, λ' be sequent instantiations over C. Then, for every $h \in HSeq(C)$,*

$$\big[(\sigma, \varrho, \lambda) \cdot \big(\sigma', \varrho', \lambda'\big)\big](h) = (\sigma, \varrho, \lambda)\big(\big(\sigma', \varrho', \lambda'\big)(h)\big).$$

The next proposition will be useful.

Proposition 3.5 *Let $h_1 \dots h_n$ be a derivation of h in $\mathcal{A}$ from Υ. Then, for every $(\sigma, \varrho, \lambda)$, the sequence $(\sigma, \varrho, \lambda)(h_1) \dots (\sigma, \varrho, \lambda)(h_n)$ is a derivation of $(\sigma, \varrho, \lambda)(h)$ in $\mathcal{A}$ from $(\sigma, \varrho, \lambda)(\Upsilon)$.*

Proof By induction on n, taking into account Definition 2.6 and Proposition 3.4. □

Given an hypersequent h, we denote by $Var(h)$ the set of all the scheme variables occurring in formulas in h. If Υ is a set of hypersequents, then $Var(\Upsilon)$ denotes the subset $\bigcup_{h \in \Upsilon} Var(h)$ of Ξ.

Corollary 3.6 *Let $\mathcal{A} = \langle C, R\rangle$ be a chc, and $\Upsilon \cup \{h\} \subseteq HSeq(C)$ such that $\Upsilon \vdash_{\mathcal{A}} h$. Then, there exists a derivation $h_1 \dots h_n$ of h in $\mathcal{A}$ from Υ such that $Var(h_i) \subseteq Var(\Upsilon) \cup Var(h)$, for every $1 \leq i \leq n$.*

Proof Consider a derivation $h_1 \dots h_n$ of h in $\mathcal{A}$ from Υ. Let σ be a substitution over C such that $\sigma(\xi) \in Var(\Upsilon) \cup Var(h)$ whenever $\xi \notin Var(\Upsilon) \cup Var(h)$, and $\sigma(\xi) = \xi$ otherwise. Let ϱ and λ be the identity instantiation and the identity sequent instantiation over C, respectively. By Proposition 3.5, the sequence $(\sigma, \varrho, \lambda)(h_1) \dots (\sigma, \varrho, \lambda)(h_n)$ is a derivation of h in $\mathcal{A}$ from Υ such that, for every $1 \leq i \leq n$, $Var((\sigma, \varrho, \lambda)(h_i)) \subseteq Var(\Upsilon) \cup Var(h)$. □

The proof of the next result is routine. In particular, item (i) is an immediate consequence of Proposition 3.5.

Theorem 3.7

(i) *If $\Upsilon \vdash_{\mathcal{A}} h$, then $(\sigma, \varrho, \lambda)(\Upsilon) \vdash_{\mathcal{A}} (\sigma, \varrho, \lambda)(h)$, for every triple $(\sigma, \varrho, \lambda)$.*
(ii) *If $\Upsilon \vdash_{\mathcal{A}_1} h$, then $\hat{f}(\Upsilon) \vdash_{\mathcal{A}_2} \hat{f}(h)$, for every morphism $f : \mathcal{A}_1 \to \mathcal{A}_2$ in* **CHC**.

Recall from [8] the following result:

Proposition 3.8 *The category* **Sig** *has finite coproducts.*

The coproduct of the signatures C and C' will be denoted by $C \oplus C'$, with the canonical injections $i : C \to C \oplus C'$ and $i' : C' \to C \oplus C'$.

The (unconstrained) fibring of hypersequent calculi is defined as expected:

Definition 3.9 Let $\mathcal{A} = \langle C, R\rangle$ and $\mathcal{A}' = \langle C', R'\rangle$ be two chcs. The (unconstrained) fibring of $\mathcal{A}$ and $\mathcal{A}'$ is the commutative hypersequent calculus $\mathcal{A} \oplus \mathcal{A}' = \langle C, R\rangle$ where:

- $C = C \oplus C'$,
- $R = \{\hat{i}(r) : r \in R\} \cup \{\hat{i}'(r) : r \in R\}$.

Here i and i' are the canonical injections from C and C' to $C \oplus C'$, respectively.

The characterization of unconstrained fibring as a coproduct can be proved by generalizing the corresponding proof for sequent calculi found in [8]:

Proposition 3.10 *Let $\mathcal{A} = \langle C, R\rangle$ and $\mathcal{A}' = \langle C', R'\rangle$ be two chcs. Then, $\mathcal{A} \oplus \mathcal{A}'$ is the coproduct in* **CHC** *of $\mathcal{A}$ and $\mathcal{A}'$ with canonical injections induced by the injections i and i' in* **Sig**.

4 Admissible and Derivable Rules and the Rule Elimination Property

Recall that a rule of inference is admissible in a formal system if the set of theorems of the system is closed under the rule. In our context, we arrive to the following definition:

Definition 4.1 Let $\mathcal{A} = \langle C, R\rangle$, and let $r = \langle\{h_1, \ldots, h_n\}, h\rangle$ be an inference rule over C (r can belong or not to R). We say that r is an *admissible inference rule of* $\mathcal{A}$ if, for every substitution σ, instantiation ϱ and sequent instantiation λ over C, it is verified that

$$\text{if } \vdash_{\mathcal{A}} (\sigma, \varrho, \lambda)(h_i) \quad \text{for all } i = 1, \ldots n, \quad \text{then } \vdash_{\mathcal{A}} (\sigma, \varrho, \lambda)(h).$$

That is to say, an admissible rule is one whose conclusion holds whenever the premises hold, and so that rule can be added to the system without changing theoremhood. It is easy to prove the following:

Proposition 4.2 *Let* $\mathcal{A} = \langle C, R\rangle$ *and let* r *be an inference rule over* C. *Let* $\mathcal{A}^r = \langle C, R \cup \{r\}\rangle$. *Then* r *is admissible in* $\mathcal{A}$ *iff, for every hypersequent* h *over* C, $\vdash_{\mathcal{A}^r} h$ *implies* $\vdash_{\mathcal{A}} h$.

Clearly, if $r \in R$ then r is admissible in $\langle C, R\rangle$. A related notion is that of derived rule. A rule r is said to be derivable in a chc if its conclusion can be derived from its premises using the other rules of the system. Formally:

Definition 4.3 Let $\mathcal{A} = \langle C, R\rangle$, and let $r = \langle \Upsilon, h\rangle$ be an inference rule over C such that $r \notin R$. We say that r is a *derived inference rule of* $\mathcal{A}$ if $\Upsilon \vdash_{\mathcal{A}} h$.

From now on, if $\mathcal{A} = \langle C, R\rangle$, we will denote by $\mathcal{A}_r$ the chc $\langle C, R \setminus \{r\}\rangle$.

Remark 4.4

(i) If $f : \mathcal{A} \to \mathcal{A}'$ is a morphism in **CHC** and $r \in R$, then we have that $\hat{f}(r)$ is a derived inference rule of $\mathcal{A}'$.
(ii) Every derived rule is admissible, but the converse does not hold.

Recall that the cut-elimination theorem (Hauptsatz) states that any sentence that possesses a proof in the sequent calculus (hypersequent calculus) that makes use of the cut rule also possesses a cut-free proof, that is, a proof that does not make use of the cut rule. It was originally proved by G. Gentzen (see [12]) for the systems LJ and LK formalizing intuitionistic and classical logic, respectively. It is well known that in a sequent calculus where cut elimination holds, the cut rule is admissible in the calculus obtained by removing the cut rule. Taking that into account we introduce the following definition:

Definition 4.5 (Rule elimination property) Let $\mathcal{A} = \langle C, R\rangle$ be a chc, and let r be a rule in R. We shall say that $\mathcal{A}$ *admits elimination of rule* r (or simply that $\mathcal{A}$ has the *r-elimination property*) if every time that $h \in HSeq(C)$ has a proof in $\mathcal{A}$, then there is a proof of h in $\mathcal{A}_r$.

Then, by Proposition 4.2, we have:

Corollary 4.6 *Let $\mathcal{A} = \langle C, R\rangle$ be a chc, and let r be a rule in R. Then, the following conditions are equivalent:*

(i) *r is admissible in $\mathcal{A}_r$;*
(ii) *$\mathcal{A}$ has the r-elimination property.*

Definition 4.5 can be generalized as follows:

Definition 4.7 Let $\mathcal{A} = \langle C, R\rangle$ be a chc, and let r be a rule in R. We shall say that $\mathcal{A}$ *admits full elimination of the rule r* (or simply that $\mathcal{A}$ has the *full r-elimination property*) if for every derivation of h from Υ in $\mathcal{A}$, there is a derivation of h from Υ in $\mathcal{A}_r$.

Clearly, if $\mathcal{A}$ admits full elimination of r then it admits the elimination of r: it is enough to take $\Upsilon = \emptyset$. The converse does not hold.

The next result is the counterpart of Corollary 4.6 in terms of the notion of derivability.

Proposition 4.8 *Let $\mathcal{A} = \langle C, R\rangle$ be a chc, and let r be a rule in R. Then, the following conditions are equivalent:*

(i) *r is derivable in $\mathcal{A}_r$;*
(ii) *$\mathcal{A}$ has the full r-elimination property.*

5 Preservation Features and Translating Derivations

In this section, some preservation features are explored. In particular, it is noted that r-elimination property is not preserved by fibring of commutative hypersequent calculi, provided that one (or both) of the calculi enjoys this property. Additionally, by using the notion of goedelization proposed in [5], we shall be able to translate derivations from a given calculus into another.

Observe that, in general, the rule-elimination property is not preserved by fibring of commutative hypersequent calculi provided that *just one* of the calculi has the rule-elimination property: for instance, cut-elimination property is not preserved by fibring of commutative hypersequent calculi such that just one of them enjoys cut-elimination.

Example 5.1 A. Avron (in [1]) constructed a commutative hypersequent calculus called *GLCW* enjoying cut-elimination property such that, by adding the connective $\wedge$ together with the usual rules, the resulting calculus GLC^* does not have the cut-elimination property. In our framework, this means that the fibring of *GLCW* (with cut-elimination) and a calculus of conjunction plus cut (without cut-elimination), results in a calculus without cut-elimination.

This proves the following assertion:

Fact 5.2 *In general, unrestricted fibring of commutative hypersequent calculi does not preserve the cut-elimination property, provided that just one of them has this metaproperty.*

The more interesting case is when both fibred systems admit r-elimination. However, it is not hard to see that the obtained system need not to admit r-elimination. Indeed, take the sequent calculus for the multiplicative fragment of abelian logic and the additive fragment of linear logic (see [15]). Both calculi have cut-elimination, but the result of fibring does not.

Fact 5.3 *In general, unrestricted fibring of commutative hypersequent calculi does not preserve the cut-elimination property, provided that both of them have this metaproperty.*

Remark 5.4 Observe that connectives are not shared in the unrestricted fibring. However, the rule r is present in both calculi, collapsing into just one rule in the fibring. Then, r must be a rule without occurrences of connectives, that is, a structural rule like, for instance, Cut or Contraction.

On the other hand, things change when we consider full rule elimination. From Proposition 4.8 we have:

Fact 5.5 *Let $\mathcal{A} = \langle C, R\rangle$ be a chc. If $\mathcal{A}$ has the full r-elimination property, then so does $\mathcal{A} \oplus \mathcal{A}'$, for any chc $\mathcal{A}' = \langle C', R'\rangle$.*

The next two results are tools for translating derivations from a given calculus into an extension of it, and vice versa. This technique is based on the notion of goedelization introduced in [5].

Definition 5.6 Let C and C' be two signatures. An *embedding* from C to C' is a signature morphism $l : C \to C'$ such that, for every $n \geq 0$ and $c \in C_n$, there exists a unique $c' \in C'_n$ such that $l(c) = c'(\xi_1, \dots, \xi_n)$, if $n > 0$, and $l(c) = c'$ if $n = 0$. We write $C \leq_l C'$ to denote that $l : C \to C'$ is an embedding.

Observe that if $C \leq_l C'$ then the underlying functions $l : |C| \to L(C')$ and $\hat{l} : L(C) \to L(C')$ are injective. Additionally, $C_j \leq_{i_j} C_1 \oplus C_2$ for $j = 1, 2$, for every signatures C_1 and C_2. That is, the canonical injections i_1 and i_2 of the coproduct $C_1 \oplus C_2$ are embeddings.

Definition 5.7 Let C and C' be two signatures such that $C \leq_l C'$, and consider a recursive bijection (from now on called *goedelization*) $g : L(C') \to \mathbb{N}$. The translation $\tau_g : L(C') \to L(C)$ is the function inductively defined as follows:

- $\tau_g(\xi_i) = \xi_{2i+1}$, for $\xi_i \in \Xi$;
- $\tau_g(l(c)) = c$, for $c \in C_0$;
- $\tau_g(c') = \xi_{2g(c')}$, for $c' \in C'_0 \setminus l(C_0)$;
- $\tau_g(l(c)(\gamma'_1, \dots, \gamma'_k)) = c(\tau_g(\gamma'_1), \dots, \tau_g(\gamma'_k))$, for $c \in C_k$, $k > 0$ and $\gamma'_i \in L(C')$;
- $\tau_g(c'(\gamma'_1, \dots, \gamma'_k)) = \xi_{2g(c'(\gamma'_1,\dots,\gamma'_k))}$, if $k > 0$, $c' \in C'_k$ is such that $c'(\xi_1, \dots, \xi_k) \neq l(c)$ for every $c \in C_k$, and $\gamma'_i \in L(C')$.

The substitution $\tau_g^{-1} : \Xi \to L(C')$ is the function defined by

- $\tau_g^{-1}(\xi_{2i+1}) = \xi_i$ and
- $\tau_g^{-1}(\xi_{2i}) = g^{-1}(i)$, for all $i \in \mathbb{N}$.

We shall denote by τ_g^{-1} the only extension of τ_g^{-1} to $L(C)$. It can be proved that τ_g and τ_g^{-1} are inverses of each other. The function defined by

$$\hat{\tau}_g(A;\Gamma \succ \Delta;B) = \big(A;\tau_g(\Gamma) \succ \tau_g(\Delta);B\big)$$

induces the function $\hat{\tau}_g : \mathcal{MP}_{\text{fin}}(\mathcal{X} \cup L(C')) \to \mathcal{MP}_{\text{fin}}(\mathcal{X} \cup L(C))$ in a natural way; and if we define

$$\hat{\hat{\tau}}_g\big(\langle \mathcal{H};\mathcal{S}\rangle\big) = \big\langle \mathcal{H};\hat{\tau}_g(\mathcal{S})\big\rangle,$$

it induces a function $\hat{\hat{\tau}}_g : \mathcal{MP}_{\text{fin}}(\mathfrak{H} \cup Seq(C')) \to \mathcal{MP}_{\text{fin}}(\mathfrak{H} \cup Seq(C))$.

Analogously, we define $\hat{\tau}_g{}^{-1} : \mathcal{MP}_{\text{fin}}(\mathcal{X} \cup L(C)) \to \mathcal{MP}_{\text{fin}}(\mathcal{X} \cup L(C'))$ and $\hat{\hat{\tau}}_g^{-1} : \mathcal{MP}_{\text{fin}}(\mathfrak{H} \cup Seq(C')) \to \mathcal{MP}_{\text{fin}}(\mathfrak{H} \cup Seq(C))$.

Remark 5.8 It is easy to see that $\hat{\tau}_g$ and $\hat{\tau}_g{}^{-1}$; and $\hat{\hat{\tau}}_g$ and $\hat{\hat{\tau}}_g^{-1}$ are inverse to each other, respectively.

Lemma 5.9 *Let C and C' be two signatures such that $C \leq_l C'$, and let $g : L(C') \to \mathbb{N}$ be a goedelization.*

(i) *If $\sigma' : \Xi \to L(C')$ is a substitution over C', then $\bar{\sigma} : \Xi \to L(C)$ given by $\bar{\sigma}(\xi) = \tau_g(\sigma'(\xi))$ is a substitution over C such that $\hat{\bar{\sigma}}(\varphi) = \tau_g(\hat{\sigma}'(\hat{l}(\varphi)))$ for all $\varphi \in L(C)$.*
(ii) *If $\varrho' : \mathcal{X} \to \mathcal{MP}_{\text{fin}}(\mathcal{X} \cup L(C'))$ is an instantiation over C' and $\hat{\tau}_g : \mathcal{MP}_{\text{fin}}(\mathcal{X} \cup L(C')) \to \mathcal{MP}_{\text{fin}}(\mathcal{X} \cup L(C))$ is defined as above, then the composite map $\bar{\varrho} = \hat{\tau}_g \circ \varrho'$ is an instantiation over C.*
(iii) *If $\lambda' : \mathfrak{H} \to \mathcal{MP}_{\text{fin}}(\mathfrak{H} \cup Seq(C'))$ is a sequent instantiation over C' and $\hat{\hat{\tau}}_g : \mathcal{MP}_{\text{fin}}(\mathfrak{H} \cup Seq(C')) \to \mathcal{MP}_{\text{fin}}(\mathfrak{H} \cup Seq(C))$ is defined as above, then the composite map $\bar{\lambda} = \hat{\hat{\tau}}_g \circ \lambda'$ is a sequent instantiation over C.*
(iv) *With the notation used in the above items, if $h \in HSeq(C)$ then*

$$\hat{\hat{\tau}}_g\big((\sigma',\varrho',\lambda')\big(\hat{l}(h)\big)\big) = (\bar{\sigma},\bar{\varrho},\bar{\lambda})(h).$$

Proof Straightforward. □

Using the above results we are now able to translate derivations.

Proposition 5.10 *Let $\mathcal{A} = \langle C, R\rangle$ and $\mathcal{A}' = \langle C', R'\rangle$ be two chcs such that $C \leq_l C'$ and $\hat{l}(R) \subseteq R'$, and let $\Upsilon \cup \{h\} \subseteq HSeq(C')$ such that $h_1 \dots h_n$ is a derivation of h in $\mathcal{A}'$ from Υ using exclusively rules from $\hat{l}(R)$. Then, $\hat{\hat{\tau}}_g(\Upsilon) \vdash_{\mathcal{A}} \hat{\hat{\tau}}_g(h)$, for any goedelization $g : L(C') \to \mathbb{N}$. Moreover, $\hat{\hat{\tau}}_g(h_1) \dots \hat{\hat{\tau}}_g(h_n)$ is a derivation of $\hat{\hat{\tau}}_g(h)$ in $\mathcal{A}$ from $\hat{\hat{\tau}}_g(\Upsilon)$.*

Proof We shall use induction over the length m of the derivation. If $m = 1$, then we have two possible cases:

Case 1. $h_1 \in \Upsilon$ and $h_1 = h$. Then, $\hat{\hat{\tau}}_g(h_1) \in \hat{\hat{\tau}}_g(\Upsilon) \subseteq HSeq(C)$ and $\hat{\hat{\tau}}_g(h) = \hat{\hat{\tau}}_g(h_1)$. Hence, $\hat{\hat{\tau}}_g(\Upsilon) \vdash_{\mathcal{A}} \hat{\hat{\tau}}_g(h)$.
Case 2. h_1 is the image of an axiom of R. That is to say, there are an $r \in R$, $r = \langle \emptyset, h'\rangle$, a substitution σ' over C', an instantiation ϱ' over C' and a sequent instantiation λ over C' such that $h_1 = (\sigma',\varrho',\lambda')(\hat{l}(h'))$. By Lemma 5.9(iv), $\hat{\hat{\tau}}_g(h_1) = \hat{\hat{\tau}}_g((\sigma',\varrho',\lambda')(\hat{l}(h'))) =$

$(\bar{\sigma}, \bar{\varrho}, \bar{\lambda})(h')$, where $\bar{\sigma}$, $\bar{\varrho}$ and $\bar{\lambda}$ are a substitution, an instantiation, and a sequent instantiation, respectively, over C. This means that $\vdash_{\mathcal{A}} \hat{\hat{\tau}}_g(h)$, and therefore $\hat{\hat{\tau}}_g(\Upsilon) \vdash_{\mathcal{A}} \hat{\hat{\tau}}_g(h)$.

Suppose now that the assertion stands for derivations of length $\leq m$ and let us see that it also stands for length $m+1$.

Let

$$h_1 \ldots h_m h_{m+1}$$

be a derivation of h in $\mathcal{A}'$ from Υ. Then $\hat{\hat{\tau}}_g(h_1) \ldots \hat{\hat{\tau}}_g(h_m)$ is a derivation of $\hat{\hat{\tau}}_g(h_m)$ in $\mathcal{A}$ from $\hat{\tau}_g(\Upsilon)$.

If either h_{m+1} is the image of an axiom of R or it belongs to Υ, the treatment is as above.

On the other hand, suppose that there are an $r \in R$, $r = \langle\{g_1, \ldots, g_k\}, g'\rangle$, a substitution σ', an instantiation ϱ', and a sequent instantiation λ' on C' such that $(\sigma', \varrho', \lambda')$ $(\hat{l}(g_j)) \in \{h_1, \ldots, h_m\}$, for $1 \leq j \leq k$, and $(\sigma,' \varrho', \lambda')(\hat{l}(g')) = h_{m+1} = h$. Then, $\hat{\hat{\tau}}_g((\sigma', \varrho', \lambda')(\hat{l}(g_j))) = (\bar{\sigma}, \bar{\varrho}, \bar{\lambda})(g_j) \in \{\hat{\hat{\tau}}_g(h_1), \ldots, \hat{\hat{\tau}}_g(h_m)\}$, for $1 \leq j \leq k$, and $\hat{\hat{\tau}}_g((\sigma', \varrho', \lambda')(\hat{l}(g'))) = (\bar{\sigma}, \bar{\varrho}, \bar{\lambda})(g') = \hat{\hat{\tau}}_g(h_{m+1}) = \hat{\hat{\tau}}_g(h)$, by Lemma 5.9(iv). By induction hypothesis, we may assert that

$$\hat{\hat{\tau}}_g(h_1) \ldots \hat{\hat{\tau}}_g(h_m) \hat{\hat{\tau}}_g(h_{m+1})$$

is a derivation of $\hat{\hat{\tau}}_g(h)$ in $\mathcal{A}$ from $\hat{\hat{\tau}}_g(\Upsilon)$. □

And conversely:

Proposition 5.11 *Let $\mathcal{A} = \langle C, R\rangle$ and $\mathcal{A}' = \langle C', R'\rangle$ be two chcs such that $C \leq_l C'$ and $\hat{l}(R) \subseteq R'$, and let $\Upsilon \cup \{h\} \subseteq HSeq(C)$ such that $h_1 \ldots h_n$ is a derivation of h in $\mathcal{A}$ from Υ. Then, $\hat{\hat{\tau}}_g^{-1}(h_1) \ldots \hat{\hat{\tau}}_g^{-1}(h_n)$ is a derivation of $\hat{\hat{\tau}}_g^{-1}(h)$ in $\mathcal{A}'$ from $\hat{\hat{\tau}}_g^{-1}(\Upsilon)$, for any goedelization $g : L(C') \to \mathbb{N}$.*

Proof It is similar to the previous one. □

6 Extension to Constrained Fibring

In Sect. 3, we introduced the notion of unconstrained fibring in **CHC**, that is, the combination of two chcs without sharing any connectives. However, it is frequently necessary to combine logics while sharing some connectives. Here is when the constrained fibring appears. In this section, we shall introduce the notion of constrained fibring in the category **CHC**, generalizing the results of the previous sections.

Let C be a signature and let $\equiv \subseteq |C| \times |C|$ be an equivalence relation on $|C|$. We shall say that $\equiv$ is a *signature congruence over* C if it verifies the following condition:

$$c_1 \equiv c_2 \quad \text{implies } c_1, c_2 \in C_n, \text{ for some } n \in \mathbb{N}.$$

It is clear that $C/{\equiv} = \{C_n/{\equiv}\}_{n \in \mathbb{N}}$ is a signature. The *canonical map* $q : |C| \to L(C/{\equiv})$ is the function defined as follows:

- $q(c) = [c]$, if $c \in C_0$,
- $q(c) = [c](\xi_1, \ldots, \xi_n)$, if $c \in C_n$, for $n \geq 1$

where $[c]$ denotes the equivalence class of c by $\equiv$. Clearly, q is a morphism $q : C \to C/{\equiv}$ in **Sig**.

Definition 6.1 Let $\mathcal{A} = \langle C, R\rangle$ be a chc and let $\equiv$ be a signature congruence over C. The *quotient commutative hypersequent calculus* (or simply the *quotient calculus*) determined by $\equiv$ is the chc $\mathcal{A}/{\equiv} = \langle C/{\equiv}, R'\rangle$ such that

$$R' = \left\{\hat{q}(r) : r \in R\right\}.$$

It is clear that $\mathcal{A}/{\equiv}$ is indeed a commutative hypersequential calculus and that q induces a morphism $q : \mathcal{A} \to \mathcal{A}/{\equiv}$ in **CHC**.

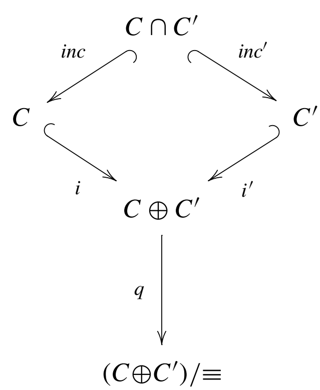
Proposition 6.2 *If $\mathcal{A}$ has the full r-elimination property, then so does $\mathcal{A}/{\equiv}$ for any congruence $\equiv$.*

Proof Straightforward. □

Let $\mathcal{A} = \langle C, R\rangle$ and $\mathcal{A}' = \langle C', R'\rangle$ be two chcs to be combined by sharing the connectives in the signature $C \cap C' = \{C_n \cap C'_n\}_{n\in\mathbb{N}}$. Let $inc : C \cap C' \to C$ and $inc' : C \cap C' \to C'$ be the inclusion morphisms. Consider now the coproduct $C \oplus C'$ and the canonical injections $i : C \to C \oplus C'$, $i' : C' \to C \oplus C'$. Then, the relation $\equiv$ given by

$$\equiv = \left\{\left(\lfloor i \circ inc(c)\rfloor, \lfloor i' \circ inc'(c)\rfloor\right) : c \in C \cap C'\right\} \cup \left\{\left(c', c'\right) : c' \in \left(C \cup C'\right) \setminus C \cap C'\right\}$$

where $C \cup C' = \{C_n \cup C'_n\}_{n\in\mathbb{N}}$ and $\lfloor c(\xi_1, \ldots, \xi_n)\rfloor = c$ for any connective c, is a congruence over $C \oplus C'$.



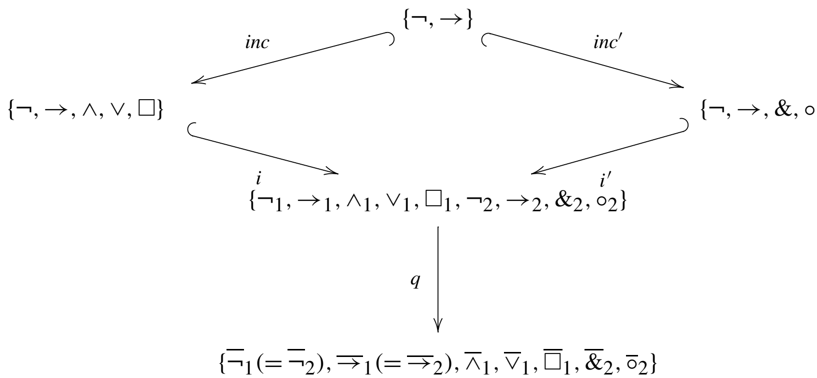
The *constrained fibring of $\mathcal{A}$ and $\mathcal{A}'$ by sharing the symbols in $C \cap C'$* is the chc

$$\mathcal{A} \overset{C\cap C'}{\oplus} \mathcal{A}' = \left(\mathcal{A} \oplus \mathcal{A}'\right)/{\equiv}.$$

Observe that if $c' \in (C \oplus C')_n$, we have the following cases:

(I) $[c'] = \{\lfloor i \circ inc(c) \rfloor, \lfloor i' \circ inc'(c) \rfloor\}$, for $c \in (C_n \cap C'_n)$;
(II) $[c'] = \{\lfloor i(c) \rfloor\}$ for a unique $c \in C_n \setminus C'_n$;
(III) $[c'] = \{\lfloor i'(c) \rfloor\}$ for a unique $c \in C'_n \setminus C_n$.

Example 6.3 Let $|C| = \{\neg, \to, \wedge, \vee, \Box\}$ and $|C'| = \{\neg, \to, \&, \circ\}$. Then,



Then, the formula $\overline{\Box}_1 \overline{\neg}_1 (\xi_1 \overline{\to}_1 (\xi_2 \,\overline{\&}_2\, \xi_3))$ of $L((C^1 \oplus C^2)/\equiv)$ stands for identifying the following formulas of $L(C^1 \oplus C^2)$:

- $\Box_1 \neg_1 (\xi_1 \to_1 (\xi_2 \,\&_2\, \xi_3))$,
- $\Box_1 \neg_1 (\xi_1 \to_2 (\xi_2 \,\&_2\, \xi_3))$,
- $\Box_1 \neg_2 (\xi_1 \to_1 (\xi_2 \,\&_2\, \xi_3))$, and
- $\Box_1 \neg_2 (\xi_1 \to_2 (\xi_2 \,\&_2\, \xi_3))$.

From Fact 5.5 and Proposition 6.2, we can state the following:

Fact 6.4 *Let $\mathcal{A}$ a chc. If $\mathcal{A}$ has the full r-elimination property then so does $\mathcal{A} \overset{C \cap C'}{\oplus} \mathcal{A}'$, for every chc $\mathcal{A}'$.*

7 The Non-commutative Case

In the previous section, ‘concrete’ sequents were considered as formed by pairs of (finite) multisets of formulas, while ‘concrete’ hypersequents were defined as (finite) multisets of sequents. It is a natural question how to generalize the previous approach to general (non-commutative) sequents and hypersequents, where finite sequences are taken instead of multisets. It is worth noting that the case of general sequents was already addressed in [8]. The aim of this section is to generalize the previous definitions and results to general hypersequents, composed of general sequents.

Given a set X, we denote by X^* the set of all finite sequences formed by elements of X, and by X^2 the Cartesian product $X \times X$. The empty sequence is denoted by ϵ. The concatenation of two finite sequences $s, s' \in X^*$ is the finite sequence denoted by $s \cdot s'$. Note that $s \cdot \epsilon = \epsilon \cdot s = s$, for every $s \in X^*$.

Recall from Sect. 2 that Ξ, $\mathcal{X}$ and $\mathfrak{H}$ are the set of scheme variables, context variables and sequent variables, respectively.

Definition 7.1 Let C be a signature. A *general sequent* (over C) is a pair of finite sequences whose elements are either formulas over C or elements of $\mathcal{X}$. The set of all general sequents over C will be denoted by $GSeq(C)$. That is,

$$GSeq(C) = \big((L(C) \cup \mathcal{X})^*\big)^2.$$

A general sequent $\langle s_1 \dots s_n, s'_1 \dots s'_m \rangle$ will be usually written as $s_1 \dots s_n \succ s'_1 \dots s'_m$. The sequent $\langle \epsilon, \epsilon \rangle$ is called the *bottom sequent*, and it is denoted by $\bot_{\mathbf{s}}$.

Definition 7.2 Let C be a signature. A *general hypersequent* (over C) is a finite sequence whose elements are either general sequents over C or elements of $\mathfrak{H}$. The set of all general hypersequents over C will be denoted by $GHSeq(C)$. That is,

$$GHSeq(C) = \big(GSeq(C) \cup \mathfrak{H}\big)^*.$$

A general hypersequent $s_1 \dots s_n$ will be usually written as $s_1 | \cdots | s_n$. The hypersequent ϵ is called the *bottom hypersequent*, denoted by $\bot_{\mathbf{h}}$.

Definition 7.3 Let C be a signature. A (n-ary) inference rule of general hypersequents over C is a pair $r = \langle \{h_1, \dots, h_n\}, h \rangle$ such that $h_i, h \in GHSeq(C)$. If $n = 0$ then r is called *an axiom*. A *general hypersequent calculus* (ghc) is a pair $\mathcal{A} = \langle C, R \rangle$ where C is a signature and R is a finite set of inference rules of general hypersequents over C.

As it was done for chcs, rules of the form $\langle \{h_1, \dots, h_n\}, h \rangle$ and $\langle \emptyset, h \rangle$ will be simply denoted by

$$\frac{h_1 \quad \dots \quad h_n}{h} \quad \text{and} \quad \frac{}{h}.$$

Recall that a substitution over a signature C is a map $\sigma : \Xi \to L(C)$, and that its unique homomorphic extension to $L(C)$ is denote by $\hat{\sigma} : L(C) \to L(C)$. Additionally, if σ and σ' are substitutions over C then $\sigma \cdot \sigma'$ is the substitution over C given by $\sigma \cdot \sigma'(\xi) = \hat{\sigma}(\sigma'(\xi))$ which satisfies the following:

$$\widehat{\sigma \cdot \sigma'} = \hat{\sigma} \circ \hat{\sigma}'.$$

Adapting [8] and the previous definitions of instantiation ϱ, sequent instantiation λ as well as the mappings of the form $(\sigma, \varrho, \lambda) : HSeq(C) \to HSeq(C)$, we introduce the following notions to deal with inference rules of general hypersequents.

Definition 7.4 (Context substitutions) Let σ be a substitution over C and let $\varrho : \mathcal{X} \to (L(C) \cup \mathcal{X})^*$ be a mapping (called *context instantiation* over C). A pair $\langle \sigma, \varrho \rangle$ is called a *context substitution* over C.

A context substitution $\mu = \langle \sigma, \varrho \rangle$ generates naturally a function $\bar{\mu} : L(C) \cup \mathcal{X} \to (L(C) \cup \mathcal{X})^*$ as follows: $\bar{\mu}(s) = \hat{\sigma}(s)$, if $s \in L(C)$, and $\bar{\mu}(s) = \varrho(s)$, if $s \in \mathcal{X}$ (note that this is well-defined, since it is assumed that $L(C) \cap \mathcal{X} = \emptyset$). It induces a unique function $\tilde{\mu} : (L(C) \cup \mathcal{X})^* \to (L(C) \cup \mathcal{X})^*$ as follows: $\tilde{\mu}(s_1 \dots s_n) = \bar{\mu}(s_1) \dots \bar{\mu}(s_n)$, for $n \geq 0$. Note that $\tilde{\mu}(\epsilon) = \epsilon$. Finally, the last mapping induces a unique function $\hat{\mu} : GSeq(C) \to GSeq(C)$ as expected: $\hat{\mu}(\langle s_1 \dots s_n, s'_1 \dots s'_m \rangle) = \langle \tilde{\mu}(s_1 \dots s_n), \tilde{\mu}(s'_1 \dots s'_m) \rangle$. Observe that $\hat{\mu}(\bot_{\mathbf{s}}) = \bot_{\mathbf{s}}$.

The following notion comes from [8]. Let $\mu = \langle \sigma, \varrho \rangle$ and $\mu' = \langle \sigma', \varrho' \rangle$ be context substitutions over C, and consider the context instantiation $(\varrho \cdot \varrho')_\sigma$ over C defined as follows: if $s \in \mathcal{X}$ and $\varrho'(s) = s_1 \dots s_n$ then $(\varrho \cdot \varrho')_\sigma(s) = f_\varrho^\sigma(s_1) \cdots f_\varrho^\sigma(s_n)$, where, for every $s' \in L(C) \cup \mathcal{X}$,

$$f_\varrho^\sigma(s') = \begin{cases} \varrho(s') & \text{if } s' \in \mathcal{X}, \\ \hat{\sigma}(s') & \text{if } s' \in L(C). \end{cases}$$

It is easy to prove that the context substitution $\mu'' = \langle \sigma \cdot \sigma', (\varrho \cdot \varrho')_\sigma \rangle$ over C is such that $\hat{\mu}'' = \hat{\mu} \circ \hat{\mu}'$.

Definition 7.5 (Sequent substitutions) Let σ be a substitution over C, ϱ a context instantiation over C and let $\lambda : \mathfrak{H} \to GHSeq(C)$ be a mapping (called *general sequent instantiation* over C). A triple $\langle \sigma, \varrho, \lambda \rangle$ is called a *sequent substitution* over C.

A sequent substitution $\kappa = \langle \sigma, \varrho, \lambda \rangle$ generates a function $\bar{\kappa} : GSeq(C) \cup \mathfrak{H} \to GHSeq(C)$ as follows: $\bar{\kappa}(s) = \widehat{\langle \sigma, \varrho \rangle}(s)$, if $s \in GSeq(C)$, and $\bar{\kappa}(s) = \lambda(s)$, if $s \in \mathfrak{H}$. From this, a unique function $\hat{\kappa} : GHSeq(C) \to GHSeq(C)$ is defined as follows: $\hat{\kappa}(s_1 \dots s_n) = \bar{\kappa}(s_1) \cdots \bar{\kappa}(s_n)$, for $n \geq 0$. Note that $\hat{\kappa}(\bot_{\mathbf{h}}) = \bot_{\mathbf{h}}$.

Now, let $\kappa = \langle \sigma, \varrho, \lambda \rangle$ and $\kappa' = \langle \sigma', \varrho', \lambda' \rangle$ be sequent substitutions over C, and consider the general sequent instantiation $(\lambda \cdot \lambda')_{\sigma\varrho}$ over C defined as follows: if $\lambda'(s) = s_1 \dots s_n$ then $(\lambda \cdot \lambda')_{\sigma\varrho}(s) = \theta_\lambda^{\sigma\varrho}(s_1) \cdots \theta_\lambda^{\sigma\varrho}(s_n)$, where, for every $s' \in GSeq(C) \cup \mathfrak{H}$,

$$\theta_\lambda^{\sigma\varrho}(s') = \begin{cases} \lambda(s') & \text{if } s' \in \mathfrak{H}, \\ \widehat{\langle \sigma, \varrho \rangle}(s') & \text{if } s' \in GSeq(C). \end{cases}$$

It is immediate that the sequent substitution $\kappa'' = \langle \sigma \cdot \sigma', (\varrho \cdot \varrho')_\sigma, (\lambda \cdot \lambda')_{\sigma\varrho} \rangle$ over C is such that $\hat{\kappa}'' = \hat{\kappa} \circ \hat{\kappa}'$.

Now, we can define the notion of derivation in ghcs:

Definition 7.6 Let $\mathcal{A} = \langle C, R \rangle$ be a general hypersequent calculus over C and let $\Upsilon \cup \{s\} \subseteq GHSeq(C)$. We say that h is derivable in $\mathcal{A}$ from Υ, and write $\Upsilon \vdash_{\mathcal{A}} h$, if there is a finite sequence $h_1 \dots h_n$ of elements of $GHSeq(C)$ such that $h_n = h$ and for all $1 \leq i \leq n$, either $h_i \in \Upsilon$, or there exist an inference rule $r = \langle \{h'_1, \dots, h'_k\}, h' \rangle$ in R, a substitution σ, a context instantiation ϱ and a general sequent instantiation λ over C such that $(\sigma, \varrho, \lambda)(h'_j) \in \{h_1, \dots, h_{i-1}\}$ (for $1 \leq j \leq k$) and $(\sigma, \varrho, \lambda)(h') = h_i$. If $\Upsilon = \emptyset$ we shall just say that h is provable in $\mathcal{A}$.

The category of general hypersequent calculi can now be defined.

If $s_1 \dots s_n \succ s'_1 \dots s'_m$ is a general sequent over C and $f : C \to C'$ is a signature morphism, then $\hat{f}(s_1 \dots s_n \succ s'_1 \dots s'_m)$ is, by definition, the general sequent $\bar{s}_1 \dots \bar{s}_n \succ \bar{s}'_1 \dots \bar{s}'_m$ over C' where, for every $s \in L(C) \cup \mathcal{X}$, $\bar{s}$ is $\hat{f}(s)$, if $s \in L(C)$, and s otherwise. This can be naturally extended to hypersequents:

$$\hat{f}(s_1 \dots s_n) = \tilde{s}_1 \dots \tilde{s}_n$$

where, for every $s \in GSeq(C) \cup \mathfrak{H}$, $\tilde{s}$ is $\hat{f}(s)$, if $s \in GSeq(C)$, and s otherwise. It is clear that $\hat{f}(h)$ is a general hypersequent over C' provided that h is a general hypersequent over C.

Definition 7.7 The category **GHC** of general hypersequent calculi is the category whose objects are general hypersequent calculi. A morphism $f : \langle C, R\rangle \to \langle C', R'\rangle$ in **GHC** is a morphism $f : C \to C'$ in **Sig** such that, for every $r = \langle\{h_1, \ldots, h_n\}, h\rangle$ in R, it is verified that $\hat{f}(h)$ is derivable in $\langle C', R'\rangle$ from $\{\hat{f}(h_1), \ldots, \hat{f}(h_n)\}$. The composition of morphisms and the identity morphism in **GHC** is defined as in **Sig**.

It should be clear that all the previous definitions and results on (constrained and unconstrained) fibring of chcs, translation of derivations and preservation theorems can be reproduced in the framework of general hypersequent calculi. We left the details to the interested reader.

8 Hypersequent Calculi and Hypertranslations

In [8], the notion of meta-fibring based on meta-translations was proposed. Within this framework, designed to deal with sequent calculi, every morphism f (called *meta-translation*) in the category of sequent calculi has the following property: every (formal) sequent rule of the form

$$(r) \quad \frac{s_1 \quad \ldots \quad s_n}{s}$$

is preserved by f. By interpreting a sequent rule as above as a meta-property of the logic associated to the given calculus, a morphism f can be, therefore, seen as a translation between logics which preserves all the meta-properties as above. This was the startpoint of [8], additionally developed in [6], where meta-translations were called *contextual translations*. In [3], the notion of meta-translation was also used to analyze the combination of the logics of (classical) conjunction and disjunction.

The main difference between meta-translations and usual translations is that the latter just preserve simple metaproperties of the logic of the form $\Gamma \vdash \varphi$, while the former preserves logical combinations of them described by sequent rules as above. As argued in [6], contextual translations refine the usual concept of translation between logics, helping to analyze the complex question of how a logic should be translated into another one as well as the question of how a logic can be extended faithfully. As it was proved recently, the simpler notion of conservative translation was shown not to be informative enough, since any two reasonable deductive systems can be conservatively translated into each other (cf. [14]). This is not obviously the case for meta-translations: in order to be contextually translatable, the target logic must satisfy at least all the structural rules satisfied by the source logic (see [6]). This is why the inclusion morphism between the sequent calculi *INT* for intuitionistic propositional logic and *CPL*, the sequent calculus for classical propositional logic, is a meta-translation and so *INT* can be considered a "good" sublogic of *CPL*, since every meta-property of the former is enjoyed by the latter (cf. [6]).

But things are not so simple. As it is well-known, Gödel was the first to observe (cf. [13]) that, unlike to what happens in classical logic, intuitionistic propositional logic has the disjunction property, namely:

(DP) If $(\alpha \vee \beta)$ is a theorem, then α is a theorem or β is a theorem.

It is easy to see that (DP) cannot be expressed as a metaproperty in the language of (formal) sequents introduced in [8]. In fact, the metaproperty (DP) has the form

$$\frac{\vdash \alpha \vee \beta}{\vdash \alpha \quad \text{or} \quad \vdash \beta}$$

which lies outside the scope of the language of sequent rules of the form (r). In this perspective, *INT* could not be considered such a good sublogic of *CPL* since the former satisfies the metaproperty (DP) which is not satisfied by the latter. This distinction can be made precise within the framework of hypersequents.

Recall the notion of morphism in the category **CHC** of chcs given in Definition 2.9, as well as the definition of morphism in the category **GHC** of ghcs given in Definition 7.7. In both cases, every (formal) hypersequent rule of the form

$$(r') \quad \frac{h_1 \quad \dots \quad h_n}{h}$$

is preserved by such a morphisms. As it was done with sequents, an hypersequent rule as (r') could be seen as a meta-property of the logic associated to the given calculus, but written in a richer (meta)language which allows expressing metaproperties such as (DP). In fact, (DP) can be represented by the following hypersequent rule:

$$(\text{DP}') \quad \frac{\succ \xi_1 \vee \xi_2}{\succ \xi_1 \quad | \quad \succ \xi_2}.$$

By Definition 2.9 or 7.7, a morphism f will force the target logic to satisfy the following metaproperty:

$$(\text{DP}') \quad \frac{\succ \varphi(\xi_1, \xi_2)}{\succ \xi_1 \quad | \quad \succ \xi_2}$$

where $\varphi(\xi_1, \xi_2)$ is the formula associated by f to the disjunction operator $\vee$, and so $\hat{f}(\xi_1 \vee \xi_2) = \varphi(\xi_1, \xi_2)$. In particular, if f is the inclusion morphism, the rule (DP) will be satisfied by the target logic, since in this particular case we have that $\hat{f}(\xi_1 \vee \xi_2) = \xi_1 \vee \xi_2$. In other words, if intuitionistic propositional logic (presented as an hypersequent calculus) is extended through an inclusion morphisms of hypersequent calculi, the target calculus must also satisfy the disjunction property. This justify to call the morphisms of hypersequent calculus as *hypertranslations*.

From the above discussion, we believe that the present framework of formal hypersequent calculi can throw some light on the subject of translations between logics and its significance.

9 Concluding Remarks

The present paper generalizes in a natural way the formal treatment of sequent calculi and their fibring introduced in [8]. Additionally, some preservation features were analyzed. Finally, the relevance of this approach concerning the theory of translations between logics was stressed.

As observed in Remark 2.5, we propose here an intensional approach to inference rules, in contrast with the traditional, extensional approach to inference rules. Being so, a single linguistic object represents infinite concrete rules which are obtained by instantiation of their metavariables, that is, variables in the metalanguage. The advantage of using formal variables instead of informal metavariables is crucial in the context of combining logic systems: in our framework, the rules are prepared to be combined, being ready to accept new connectives by means of substitutions over the language resulting from the combination procedure.

Several other questions remain open, and deserve future research. The use of hypersequents instead of sequents opens interesting possibilities to the study of how a logic can be constructed (or deconstructed) from (into) its fragments, along the lines of the studies initiated in [8]. The preservation by fibring of some meta-properties of hypersequent calculi (interpolation, for instance) should also be addressed.

Acknowledgements The first author was financed by FAPESP (Brazil), Thematic Project LogCons (2010/51038-0), as well as by an individual research grant from The National Council for Scientific and Technological Development (CNPq), Brazil. The second author wants to acknowledge the hospitality of CLE (UNICAMP, Brazil), where the writing of this paper was begun. Also, he wants to thank the National University of the South—UNS (Bahía Blanca, Argentina) for supporting his stay at Campinas, Brazil.

References

1. Avron, A.: Hypersequents, logical consequence and intermediate logics for concurrency. Ann. Math. Artif. Intell. **4**, 225–248 (1991)
2. Avron, A.: The method of hypersequents in the proof theory of propositional nonclassical logics. In: Logic: From Foundations to Applications, European Logic Colloquium. Oxford Science Publications, pp. 1–32. Clarendon Press, Oxford (1996)
3. Béziau, J.-Y., Coniglio, M.E.: To distribute or not to distribute? Log. J. IGPL **19**(4), 566–583 (2011)
4. Carnielli, W.A., Coniglio, M.E.: Combining logics. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (2007). http://www.plato.stanford.edu/entries/logic-combining/
5. Carnielli, W.A., Coniglio, M.E., Gabbay, D., Gouveia, P., Sernadas, C.: Analysis and Synthesis of Logics—How to Cut and Paste Reasoning Systems. Applied Logic Series, vol. 35. Springer, Amsterdam (2008)
6. Carnielli, W.A., Coniglio, M.E., D'Ottaviano, I.M.L.: New dimensions on translations between logics. Log. Univers. **3**(1), 1–18 (2009)
7. Ciabattoni, A., Galatos, N., Terui, K.: From axioms to analytic rules in nonclassical logics. In: IEEE Symposium on Logic in Computer Science (LICS'08), pp. 229–240. IEEE, New York (2008)
8. Coniglio, M.E.: Recovering a logic from its fragments by meta-fibring. Log. Univers. **1**(2), 377–416 (2007). Preprint available as: "The meta-fibring environment: preservation of meta-properties by fibring", *CLE e-Prints*, vol. 5., no. 4, 2005. http://www.cle.unicamp.br/e-prints/vol_5,n_4,2005.html
9. Coniglio, M.E., Figallo, M.: Fibring of hypersequent calculi and preservation of rule-elimination. CLE e-Prints. vol. 10, no. 7 (2010). http://www.cle.unicamp.br/e-prints/vol_10,n_7,2010.html
10. Figallo, M.: Hipersecuentes y la Lógica Tetravalente Modal $\mathcal{TML}$ (Hypersequents and the Tetravalent Modal Logic $\mathcal{TML}$, in Spanish). PhD thesis, Universidad Nacional del Sur, Argentina (2013)
11. Gabbay, D.M.: Fibring Logics. Oxford Logic Guides, vol. 38. Oxford (1998)
12. Gentzen, G.: Recherches sur la Déduction Logique Traduit de L'Allemand par R. Feys et J. Ladrière. Presses Universitaires de France, Paris (1955)
13. Gödel, K.: On the intuitionistic propositional calculus (1932). In: Feferman, S., et al. (eds.) K. Gödel's Collected Works, vol. 1, pp. 223–225. Oxford University Press, Oxford (1986)
14. Jeřábek, E.: The ubiquity of conservative translations. Rev. Symb. Log. **5**(4), 666–678 (2012)
15. Metcalfe, G., Olivetti, N., Gabbay, D.: Proof Theory for Fuzzy Logics. Applied Logic Series, vol. 36. Springer, Berlin (2009)
16. Sernadas, A., Sernadas, C., Caleiro, C.: Fibring of logics as a categorical construction. J. Log. Comput. **9**(2), 149–179 (1999)

M.E. Coniglio (✉)
Department of Philosophy and Centre for Logic, Epistemology and the History of Science (CLE), State University of Campinas (UNICAMP), Campinas, Brazil
e-mail: meconiglio@gmail.com

M. Figallo
Department of Mathematics and Institute of Mathematics (INMABB), National University of the South (UNS), Bahía Blanca, Argentina
e-mail: figallomartin@gmail.com

Investigating Knowledge and Opinion

John Corcoran and Idris Samawi Hamid

Abstract This work treats correlative concepts of knowledge and opinion, in various senses. In all senses of 'knowledge' and 'opinion', a belief known to be true is *knowledge*; a belief not known to be true is *opinion*. In this sense of 'belief', a *belief* is a proposition thought to be true—perhaps, but not necessarily, known to be true. All knowledge is truth. Some but not all opinion is truth. Every proposition known to be true is believed to be true. Some but not every proposition believed to be true is known to be true. Our focus is thus on *propositional* belief ("belief-that"): the combination of propositional knowledge ("knowledge-that") and propositional opinion ("opinion-that"). Each of a person's beliefs, whether knowledge or opinion, is the end result of a particular thought process that continued during a particular time interval and ended at a particular time with a conclusive act—a *judgment* that something is the case. This work is mainly about beliefs in substantive informative propositions—not empty tautologies.

We also treat *objectual* knowledge (knowledge of objects in the broadest sense, or "knowledge-of"), *operational* knowledge (abilities and skills, "knowledge-how-to", or "know-how"), and *expert* knowledge (expertise). Most points made in this work have been made by previous writers, but, to the best of our knowledge, they have never before been collected into a coherent work accessible to a wide audience.

Mathematics Subject Classification Primary 03B42 · Secondary 03A05

Keywords Belief · Knowledge/opinion · Propositional · Operational · Objectual · Cognition

> There was a time when I *believed* that you belonged to me. But, now I *know* your heart is shackled to a memory.—Hank Williams

1 Preliminaries

It is only with respect to propositions[1] that the distinction between [sc. propositional] knowledge and opinion applies—for example, a person's abilities cannot be said to be

[1] We use the word 'proposition' in the abstract sense in which one proposition might be expressed by many different sentences. Our usage—which follows Church [8, 9] where 'proposition' is a near synonym for 'thought' in the sense of Frege's [27] "The Thought"—is explained more fully in Corcoran [19] and [20]. Also, see Frege [28, 325].

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_5

opinion or knowledge in the correlative sense. In fact, it is only with respect to propositions that belief applies: every belief is a proposition that is or was believed to be true by some person or persons, and conversely, of course.

The last sentence is lexically ambiguous[2] in virtue of its one occurrence of the ambiguous common noun 'belief': in the more basic sense, a belief is someone's attitude of believing of a proposition; in the more derivative sense used above, a belief is a proposition someone believes. In the basic *attitudinal* sense, a person's belief that Plato admired Socrates is inseparable from the person. In the derivative *propositional* sense, a person's belief that Plato admired Socrates is independent of the person.

In the basic attitudinal sense, no two persons can have the same belief and every belief comes into being when its believer starts believing its propositional content and it perishes no later than its believer's demise. Moreover, in the attitudinal sense, no belief per se has a truth-value: the belief's propositional content might be either true or false.

In the derivative propositional sense, two persons can have the same belief. Moreover, it would be wrong to say that every belief comes into being when its believer starts believing its propositional content and it perishes no later than its believer's demise. Rather, being a belief is an extrinsic[3] property of a proposition: it is not that the belief comes into being and then perishes but that the proposition becomes and then ceases to be a certain person's belief. More explicitly, the proposition comes to be believed and then ceases to be believed. Moreover, in the propositional sense, every belief has a truth-value. Having a truth-value is an intrinsic property of propositions—as explained further in Corcoran [20].

In the basic attitudinal sense, having a certain believer is an intrinsic property of a belief. In the derivative propositional sense, having a certain believer is an extrinsic property of a belief.

Besides *propositional* knowledge, each of us has objectual knowledge—of the objects that propositions are about—and operational knowledge, or "know-how". Propositional knowledge that five plus seven is twelve might be held to presuppose objectual knowledge—of five, seven, and twelve—as well as operational knowledge—knowing how to add, or at least knowing how to add five to seven [23]. As said above, being a belief applies to propositions alone: all belief is propositional.[4]

The distinction between propositional and objectual knowledge is routinely illustrated by reference to two interpretations of the ancient Greek injunction "Know thyself" [43, p. 240]. Taking 'know' in the propositional sense, "Know thyself" can be understood as "Know *for* thyself", i.e. take the responsibility to verify your own beliefs; do not rely on the word of others. This is related to the contemporary saying, "Trust but verify". Taking 'know' in the objectual sense, "Know thyself" can be understood as "Know *of* or *about*

---

[2]Here, ambiguity is having multiple normal meanings and a sentence is lexically ambiguous if it contains an ambiguous word (lexical item). See Corcoran [19].

[3]Being an even square is an intrinsic property of the number four, being the number of Evangelists is an extrinsic property of four. Having four letters is an intrinsic property of the word 'four', being a name of the number of Evangelists is an extrinsic property of 'four'. Changes in intrinsic properties are known as ordinary changes whereas changes in extrinsic properties are variously called "Cambridge changes", "Pickwickian changes", "relative changes", and others. For further discussion and examples, see Corcoran [12].

[4]In the sense of 'belief' used in this work, propositions are exclusively the objects of believing: what a person believes is a proposition. Thus, a person's acceptance as true of an incoherency such as 'Socrates is equal' is not belief in the sense of this paper.

thyself", i.e. learn your own nature; determine your abilities and limitations; do not allow others to define your identity for you.

There is no corresponding "objectual belief" and no corresponding "operational belief". Objectual knowledge does not contrast with "objectual opinion", nor does operational knowledge contrast with "operational opinion". The expression 'my knowledge of [a certain entity]' can be used to refer to propositional knowledge or to objectual knowledge or even in some cases to operational knowledge. However, the expressions 'my belief about [a certain entity]' and 'my opinion about [a certain entity]' can be used coherently to refer only to propositional belief and propositional opinion, respectively.[5] In a given context, the appropriate senses of the word 'knowledge' are correlated with corresponding appropriate senses of the word 'opinion'. In appropriately correlated senses, a belief is knowledge if and only if it is not opinion.

Nevertheless, strictly speaking, it would be wrong to say that no knowledge is opinion—if 'knowledge' refers to everyone's propositional knowledge and 'opinion' refers to everyone's opinions. As a rule, we have the following situation for a given person, say A: Some of A's opinions are known to be true by others, and some of other people's opinions are known to be true by A. What is true is that no proposition is knowledge and opinion for the same person at the same time.

Another basic point is that every proposition that can be knowledge can be opinion: every subject matter that can be the subject matter of a proposition known to be true can be the subject matter of a proposition believed but not known to be true. Moreover, a fundamental assumption of this investigation is that no subject matter is beyond the scope of belief.[6]

The words 'knowledge' and 'opinion' have complementary senses determined by the propositional sense of 'know' being used: we will consider the strict traditional sense as well as other widely used senses. In effect, 'knowledge' means "belief that is not opinion" and 'opinion' means "belief that is not knowledge".[7] In any given context, the senses of the words 'know', 'knowledge', and 'opinion' are interdependent.

In every sense considered, propositional knowledge implies truth and belief. If a person *knows* that something is the case, it *is* the case and the knower *believes* that it is. However, neither truth nor belief, nor even the combination of truth and belief, implies knowledge. In contrast, propositional opinion implies belief but does not imply truth. As suggested above, no proposition known to be true is false; but many opinions, which are propositions believed to be true, are false. Each of a person's beliefs, whether knowledge or opinion, is the end result—the culmination, so to speak—of a particular thought process that ended at a particular time with a judgment: the process that produces knowledge has distinctive characteristics.

---

[5] An expression 'my belief of [a certain entity]' is questionable English if English at all unless the entity is something special such as a person's statement: 'my belief of the number one' is ungrammatical. An expression 'my opinion of [a certain entity]' would be taken to refer to a propositional opinion. For example, our opinion of knowledge is that people's lives are improved by it.

[6] For Corcoran, this is a working hypothesis; for Hamid, it is firmly held belief.

[7] Writing in 1846, Augustus De Morgan [24, 1–3] thought that negative substantive expressions such as 'non-human' were logically defective and would not occur in a logically perfect language [39, p. 183]. Instead, each substantive would have its own equally "positive" complementary substantive—as 'knowledge' has its complementary 'opinion'.

As said, this work treats four forms of knowledge: propositional, objectual, operational, and expert. Each instance of the first three (propositional, objectual, and operational knowledge) can involve instances of one or both of the other two. The paradigm case of the fourth—expert knowledge—is the knowledge possessed by an active qualified practitioner—the medical knowledge possessed by a physician or the mathematical knowledge possessed by a mathematician. A person's expertise includes their experience—practical and theoretical. Moreover, it involves all of the other three kinds of knowledge. Perhaps most importantly, experts know the limitations of their own expertise. Typically, the expert has a stock of unsolved problems and unsettled hypotheses to be investigated [23]. In fact, the hallmark of the expert is the ability to call to mind open questions, or hypotheses, propositions not known to be true and not known to be false.[8] The expert has no hesitation and no embarrassment saying, "I do not know"—when appropriate. Experts do not just share their knowledge: Experts unashamedly share their "ignorance". Being an expert excludes being a know-it-all. In the traditional university, the central concern is propositional knowledge; in the traditional professional school, the central concern is expert knowledge.

> One who does not embrace the fundamental principles of Logic or any other Science, whatever he may have taken on authority and learned by rote, knows, properly speaking, nothing of [Logic or] that Science.—Whately [48, xvii].

2 Introduction

The first sense of the word 'knowledge' considered in this paper is the traditional strict sense. In this sense, a given person knows that a given proposition is true only if they have judged that the proposition is true by means of a cognitive judgment. A cognitive judgment is one that was the culmination of a process including the following: understanding the proposition, gathering sufficient evidence based on personal experience of the facts the proposition is about, and bringing that evidence to bear on the issue of whether the proposition is true.[9] In this work, belief that is knowledge in the strict sense is called *certain knowledge* or *cognition*.

The strict sense of 'knowledge' is what is referred to in the ancient Greek injunction "Know thyself" in at least one of its interpretations that take 'know' in a propositional sense. Moreover, in the strict sense of 'knowledge', knowledge is personal, e.g. reliable testimony does not produce knowledge: one person cannot gain knowledge of a proposition simply by being told that it is true by another person, even if the other person knows in the strict sense. Cognitions are non-transferable. Although a proposition, whether true or false, is communicable, knowledge of the truth of a proposition is not communicable [13, pp. 113f].

[8]The terminology for propositions not known to be true and not known to be false is awkward and unsettled. What is called an open question is often not a question in any of the more usual senses. Moreover, what is called a hypothesis was never hypothesized by anyone. See the article "Conjecture" in Audi [5].

[9]The late Dr. Ray Lucas asked whether having made a cognitive judgment is sufficient for having a cognition. Unfortunately, the answer is no, people can lose cognitions. People can lose belief in a proposition they once knew because of a later mistake or because of memory decay. We suspect that there are propositions we once knew but no longer believe and thus no longer know.

Other belief merits being called knowledge only to the extent that its acquisition approximates that of cognition, knowledge in the strict sense. Cognition fulfills an ideal only approximated by other knowledge—just as mathematical circles fulfill an ideal only approximated by certain visible shapes, we nevertheless also call such visible shapes circles. Once we leave the realm of certain knowledge, the border between what is called knowledge and what in contrast is called opinion shifts according to the degree of certainty required. Moreover, it loses its sharpness. More and more beliefs that had been "opinions" become "knowledge". This is explained more fully below.[10]

The process that results in cognition deserves the name *the cognitive process*. It is important to emphasize that the cognitive process begins with the process of grasping, understanding, a proposition—a process that has its own component processes and its own antecedents. Besides the grasping of a proposition, which is also called apprehension,[11] the cognitive process also involves evidence gathering and evidence marshaling—the bringing of the evidence to bear on the proposition to be decided. To be explicit, the cognitive process has four components: proposition grasping, evidence gathering, evidence marshaling, and proposition judging. The four are more succinctly but less explicitly called respectively: apprehension, observation, marshaling, and judging. The positions of the first and last are fixed. The first is required to set the goal of the whole process and the last achieves that goal. Cognition is a goal-directed activity.

In a way, it was somewhat arbitrary to begin the cognitive process with apprehension, grasping a proposition. For some purposes, it is important to include whatever prompts or inspires the grasping. Philosophers, including Aristotle, Peirce, Dewey, and Kuhn, have started the cognitive process before the grasping stage with experience of an obstacle, an "aporia", a blockage, a difficulty, frustration, confusion, failure, disappointment, or some other sort of discontentment or unpleasantness.[12] For reasons that will become clear, we emphatically exclude doubt from the list.

Above we used the word 'cognition' as an epistemic mass noun, which—like 'knowledge'—does not have a plural form ('knowledges' is not even an English word) and which does not follow grammatical indefinite articles ('a knowledge is true' is not well formed). In fact, 'cognition' was an exact synonym for 'knowledge' in the strict sense. Below, it is convenient to follow J.S. Mill [34, vol. I, 317], and others in using the word 'cognition' as a common or count noun that does pluralize and that does take articles. A *cognition* is a proposition known to be true: cognitions are propositions known to be true. Although knowledge is cognition and cognition is knowledge, because the word

[10] Nothing said above should be interpreted as suggesting that any given cognition is more meritorious or more worthy than any non-cognitive item of knowledge. The relative worth of two items of knowledge is beyond the scope of this essay. However, there are clearly cases in which knowledge of a certain mathematical theorem is less valuable than knowledge that a certain pill stops a certain pain.

[11] Applied to propositions, apprehension is an action while comprehension is an attitude or state [41, p. 140]. Normally, after someone apprehends a given proposition, they comprehend it for a time—often a long time during which they are only occasionally aware of the fact that they comprehend it. Apprehending a proposition takes place in a time interval—often relatively short—whose end coincides with the onset of comprehension. In a way, apprehension is to comprehension as judging is to believing: judging is an act not an attitude, believing is an attitude not an act [20].

[12] Of course, all or at least many such experiences require the subject to be pursuing a goal prior to the event. Thus, some thinkers might prefer to start the process with a goal, or even with the pursuit of the goal, or even with the desire that prompted the pursuit.

'knowledge' does not pluralize or take articles, it thus cannot be substituted in the previous sentence. The following is ungrammatical: "A knowledge is a proposition known to be true: knowledges are propositions known to be true".

The word 'truth' is used as a count noun as in 'truths are propositions that are true'; it is also used as a mass noun as in 'some truth is knowledge'.

In every sense considered, every proposition known to be true is true. Moreover, knowledge in every such sense is personal; it represents a cognitive accomplishment by the knower. There is no way to buy knowledge or even to impart it. A teacher can assist in various ways, e.g. by directing students' attention and by encouraging students to become more autonomous by thinking and doing for and by themselves. Not all 'learning' is gaining knowledge. Very little of what is learned in school is knowledge. The acquisition of knowledge is autonomous, self-affirming, disciplined, courageous, and dignified. It presupposes intellectual autonomy, intellectual freedom, and responsibility.

Propositional knowledge in the broadest sense spans a spectrum between two limiting cases.[13] One limiting case is included, the other excluded. At the included end, we have cognition—which is thought to be more common in logic and mathematics than in other fields. On the excluded end, we have groundless true belief—which is not knowledge at all. Such belief is called *credence* here. *Credence* is groundless belief in true propositions. In this broad sense, knowledge includes all cognition, at one extreme; but it excludes all credence, at the other extreme. Between the two extremes, we have non-certain knowledge—also called *probable knowledge*.[14] This might include most true beliefs. Our essential daily decisions are based more often on probable knowledge than on knowledge in the strict sense.

Absolute, *apodictic*, or *mathematical certainty* is the state of having knowledge in the strict sense. As we move away from mathematical certainty along the spectrum of probable knowledge, we come to *scientific certainty*, which has a slightly lower level of warranted assertibility. In some cases, a belief held with scientific certainty is close to cognition. We can have scientific certainty that smoking is deleterious to health. A person's decision to give up smoking is often based on probable knowledge. As we move further along, we come to *moral certainty*, the state of holding true belief that is sufficiently grounded to serve as basis for responsible action and warranted assertion. In many cases, a member of a jury should vote to acquit unless moral certainty has been achieved.[15] After moral certainty, the level of warranted assertibility continues to decline.

Eventually we find true beliefs that were acquired through processes that hardly qualify them to be called probable knowledge. It is worth emphasizing that in the sense of 'probable' used here, probable knowledge is actually knowledge, non-certain knowledge,

---

[13] See "Limiting case" and "Borderline case" in Audi [5].

[14] The word 'probable' is used in the original sense going back to around 1600 before the invention of "probability theory" gave it another meaning. In this sense, it applies to beliefs and contrasts with 'certain'. See any dictionary that dates senses, for example, Definition 1 in the *Merriam-Webster Collegiate Dictionary* [33]. It is still widely used in the original sense where there is no question of assigning numbers to "events". Frege uses it in the original sense in his classic 1918 paper "The Thought" [27] (p. 306).

[15] We do not know the history of the expression 'moral certainty'. Antoine Arnauld (1612–1694) and Pierre Nicole (1625–1695) use the expression without comment as if it were a common locution in their 1662 masterpiece *The Art of Thinking* known as *The Port-Royal Logic* [4, 264, 270]. Dessi [25, xvii] traces it to John Locke (1632–1704). Dessi (loc. cit.) and Whately [48, 243] both use it in the sense just explained.

and thus is actually true. We need to be clear that this sense of 'probable' is not one of the senses used in connection with probabilities of events. A belief that is probable knowledge might be almost a cognition, known with certainty, or it might be little more than mere credence, a true belief not known at all. Being probable knowledge is not the same as probably being knowledge. With every level of warranted assertibility, there is a corresponding level of opinion. Some members of a jury who do not have moral certainty of the guilt of the accused might nevertheless have a somewhat grounded, true belief, which would then be opinion in comparison with moral certainty.

It seems pointless to try to coin terminology for the levels of opinion that complement scientific certainty and moral certainty. It is clear that in these contexts the word 'scientific' carries epistemically honorific connotations not warranted by many opinions lacking scientific certainty. Thus, the expression 'scientific opinion' is often inappropriate. Likewise, the word 'moral' carries epistemically honorific connotations not warranted by many opinions lacking moral certainty. Thus, the expression 'moral opinion' is often inappropriate.

The expressions 'absolute certainty', 'mathematical certainty', 'scientific certainty', and 'moral certainty' are peculiarly ambiguous constructions. In the sense used here, 'absolute certainty' does not mean "certainty about absolutes", and 'mathematical certainty' does not mean "certainty about mathematics". Likewise, for 'scientific certainty' and 'moral certainty'. Moreover, 'certain knowledge' does not mean "knowledge about certainties" and 'probable knowledge' does not mean "knowledge about probabilities".

> We read fine things but never feel them to the full until we have gone the same steps as the author.—John Keats

3 The Spectrum of True Beliefs

As said above, the distinction between knowledge and opinion applies only to propositional knowledge (or knowledge-that) and not to either objectual knowledge (knowledge-of) or operational knowledge (knowledge-how-to).[16] It was also said that in many cases, each of the three requires the other two. To see how propositional knowledge involves the other two, consider propositional knowledge that no square number is twice a square number. This involves among other things objectual knowledge of the system of numbers (positive integers) and operational knowledge how to square numbers. It also involves further operational knowledge: ability to count and to perform other arithmetic operations, ability to understand propositions, ability to make judgments, and ability to deduce conclusions from premises—to mention a few of the skills used acquiring arithmetic knowledge [23].

As already emphasized, in every sense of 'know' used in this essay, every proposition known to be true actually is true. Truth is a precondition of knowledge. Moreover, except in rare cases, every such proposition was true before it was known. Fact is prior to knowledge. As Frege explicitly noted, the process of achieving knowledge in no way

---

[16]The complementarity and interconnectedness of objectual, operational, and propositional knowledge has been a cornerstone of our thinking for many years. See Corcoran [10].

intrinsically alters the proposition known.[17] Consider the proposition that no square number is twice a square number. This very proposition, which Corcoran came to know only a few years ago, was unknown until first known to be true by one of our predecessors thousands of years ago, perhaps in ancient Greece or ancient China. Later, it came to be known by countless others, one after another, down through the years in Asia Minor, North Africa, and in many other places. One and the same proposition that was at one time not known at all became known by one person and then by more and more people, speaking in different languages, at different times, and in different places—each time by one person acting autonomously. Although knowledge cannot be transmitted, a person who knows a proposition to be true can sometimes help others to come to know the same proposition.[18]

However, by becoming a cognition of a new person the proposition is changed *extrinsically*: more people know it than knew it earlier. A proposition not widely known becomes widely known. However, everything people relate to changes extrinsically every time someone relates to it: thinking of the number one changes it extrinsically.

Knowledge is *objective* in that it is of objective reality. Moreover, there is no such thing as disembodied knowledge. Every proposition known to be true is known to be true by a person. Knowledge in all senses is personal: it is *subjective* in the sense of being achieved by a knowing *subject*. The objectivity of knowledge is prior but no less essential than its subjectivity. The knowing subject willingly defers to the object known, so to speak.

Thus, knowledge *per se* is both objective and subjective, in appropriate senses of these troublingly ambiguous words. In contrast, truth *per se* is objective but not subjective. If everyone who knows a given true proposition were to forget it, there would no longer be knowledge of it, but its truth would not be altered. Being known to be true is an extrinsic property of propositions; being true is an intrinsic property.

Truth is necessary for knowledge, but knowledge is not necessary for truth. As said above, strictly speaking, knowledge is sufficient for truth. However, this should not be taken to mean that knowledge produces truth, which is a common fallacy. Rather, knowledge is sufficient for truth because knowledge has truth as a precondition.

To indicate that 'know' is being used in the strict sense emphasized in this work, words such as 'with certainty', 'categorically', or 'conclusively' can be added—as long as this is meant to refer to what the knower did objectively as opposed to what the knower felt subjectively. Below we will distinguish knowing with certainty from believing with certitude. Strict knowing requires that the knower accurately judge based on conclusive evidence. Philosophers can agree to use the words 'know' and 'knowledge' in the strict sense and yet disagree on whether a given proposition is known with certainty to be true by a given person at a given time or whether there is any such knowledge outside mathematics or even whether anyone has ever known any proposition with certainty. To some philosophers it seems amazing, implausible, or even incredible to think that people might have certain knowledge.

---

[17]Frege makes this point several places. However, without qualification, it is misleading to say that he said propositions do not change *intrinsically*—he did not emphasize the intrinsic/extrinsic distinction explained below.

[18]In fact, even belief cannot be transmitted in this sense. One person's blood can be transfused into another's body but one person's belief cannot not be moved to another person or copied on another person's mind. Beliefs are formed by judging and each judging must be autonymous. Of course, magicians, for example, trick people into making judgments they would not otherwise make.

As already noted, true opinion complementing strict knowledge spans the spectrum *starting* after cognition or absolute certainty—the limiting case of entirely adequate, evidence-based deliberation included in knowledge in the broad sense. The spectrum of true belief continues through scientific certainty and through moral certainty. It is important to recall that scientific certainty and moral certainty are only opinion in relation to knowledge in the strict sense: scientific certainty is knowledge in a less strict sense—a sense in which moral certainty is in the realm of opinion. The spectrum of true opinion finally ends with true credence or groundless opinion that happens to be true—the limiting case of total lack of evidence-based deliberation.

As an analogy, compare the spectrum of true belief to the fractions between one and zero inclusive of end-points. Imagine that the sizes somehow represent amounts of "evidentiality". The spectrum of probable knowledge excludes both one (the cognitions) and zero (the credences), but the spectrum of knowledge in the broad sense excludes only zero.

All humans [*anthropoi*] by nature desire knowledge.—Aristotle, *Metaphysics* A.1 [2]

4 The Spectra of Certainty

In a given person's lifetime or even in a relatively short interval of years or days, their degree of certainty in a given known belief may vary from near credence to absolute or near absolute certainty through increasing degrees as evidence accumulates. In fact, sometimes credence taken on trust is transformed by the cognitive process into certain knowledge. Of course, in another case, the process might reverse and a true proposition, one time known with certainty, through deteriorating memory or other factors, may over time end up being taken purely on faith.

Thus, for every person, and for each of their true beliefs, there is a spectrum of degrees of certainty within which their actual degree of certainty at a given time is located. Of course, each of these many spectra, or spectrums, is similar to what we called the spectrum of true belief above.

In appropriate orthogonal senses of 'objective' and 'subjective', there is no contradiction in saying that one and the same thing is both objective and subjective, e.g., perception, inference, and cognition.—Albert Hammond (paraphrase).

5 Senses of 'Objective' and 'Subjective'

Above we used the words 'objective' and 'subjective' as non-opposing adjectives in the domain of cognitions.

As opposing adjectives, the words 'objective' and 'subjective' are applied in different senses in different domains or ranges of applicability [11, pp. 38f]. A judgment is objective to the extent that it is based on logic and evidence or other factors pertaining to the objects that the judgment concerns; a judgment is subjective to the extent that it is based on loyalties and feelings or other factors pertaining to the subject making the judgment. People are objective to the extent they make objective judgments; people are subjective to the extent they make subjective judgments. Advice, testimony, journalism, and the like are objective to the extent that they reflect objective judgments, and subjective to the extent that they reflect subjective judgments. A person is objective to the extent that their judgments are objective, and they are subjective to the extent that their judgments are subjective. These are all interrelated but distinguishable uses of 'objective' and 'subjective'.

A distinction drawn between two objects in a given domain is objective insofar as it is based on intrinsic features of those objects, and it is subjective insofar as it based on the tastes, loyalties, feelings, etc. of the person who drew it. For example, in the domain of integers, the distinction between even and odd integers is objective but any distinction between large and small is subjective. Again, in the domain of propositions about integers, the distinction between true and false is objective but any distinction between interesting and uninteresting or between simple and complicated is subjective. Of course, we are not denying any author's right to stipulate objective usage to words whose normal use is to mark subjective distinctions. For example, certain numbers are called perfect without any suggestion that they are better than the non-perfect numbers: the perfect versus non-perfect distinction in mathematics is objective even though the same words are applied elsewhere for subjective value judgments.

Church, Tarski, and other mathematically oriented logicians use cognates of the words *object* and *subject*—such as *objective*, *subjective*, *objectively*, *subjectively*, *objectivity*, and *subjectivity*—in senses related to those in traditional subject–object epistemology where thinking *subjects* make judgments about factual *objects*. In Corcoran–Hamid [22], we survey use and conspicuous non-use of such words in logic.

For example, modern formalizations of geometry—and other historically established disciplines—require distinguishing between the *underlying logic* and the *overlying science*. Church said such distinctions are "subjective and essentially arbitrary" [9, pp. 58ff]. Tarski implied there is no "objective" basis for them [46, esp. pp. 412 and 418f]. Neither gives grounds for his claim—and neither explains his crucial term.

Distinguishing between the underlying logic and the overlying science requires distinguishing between the *logical* and the *scientific* concepts or—what amounts to the same thing with respect to interpreted formalized languages—between the *logical* and the *scientific* constants. Tarski explicitly stated that he knew of no "objective" basis for the latter distinction [46, esp. pp. 412, 418f].[19]

Remarkably, Church [9] uses *subjective* frequently while completely avoiding *objective* whereas in Tarski [46] the exact opposite holds: there *objective* occurs frequently while *subjective* is completely absent. The above epistemic uses of cognates of *subject* and *object* contrast with other essentially unrelated uses. For example, in logical syntax the word *zero* is called the subject of the sentence *zero precedes two* and the word *two* its object. Moreover, the goal or aim of a work is often called its *objective*. There seems to be no parallel contrasting usage of 'subjective'. Curiously, although the objectives of Quine and Ullian's 1970 book *The Web of Belief* [44] overlap with those of this work, it uses 'objective' only twice, both in the sense of "aim", and it uses 'subjective' only twice, both in the sense of "not objectively grounded".

---

[19]However, in the posthumously published "What are logical notions? Tarski [47], he proposed a condition for distinguishing logical from non-logical objects: individuals, sets, relations, functions, etc. This proposal does not imply a condition for distinguishing logical from non-logical concepts (senses) and thus does not yield a condition for distinguishing logical from non-logical terms (expressions).

In primary senses, *objective* and *subjective* are correlative adjectives like *old* and *young*. It is difficult to determine what is being conveyed by calling something *objective* or *subjective* unless writers give (i), for each, examples where they would apply one as opposed to the other and (ii) some sense of their criteria for applying each word. Moreover, when a writer asserts something without giving any objective grounds, the reader is justified in suspecting that the assertion was empty rhetoric or that it was based on a subjective judgment. In the case of the distinction between logical and scientific concepts, there is a history of such distinctions being made or implicitly used going back through modern, renaissance, medieval, and ancient logic. If all those logicians were mistaken, we need some kind of explanation of how they went wrong. The apparent fact that Tarski could not think of an objective basis for distinguishing between the *logical* and the *scientific* concepts does not warrant his conclusion that no such basis exists. Likewise, the apparent fact that Church could not think of an objective basis for distinguishing between the *underlying logic* and the *overlying science* does not warrant *his* conclusion that there is none.

> In the broad sense of ‘animal’, every human is an animal; in the narrow sense, no human is an animal.—Frango Nabrasa, 2001 (personal communication).

6 Broad and Narrow Senses of ‘Believe’

The verb ‘believe’ is used in several senses even when its direct object is the proposition. In all such senses, belief can manifest itself in action. If the right circumstances occur, a person who believes something will react in one way and a person not believing it will react in another way—even if in some cases the action is entirely private and not observable by others. One broad inclusive sense and two narrow exclusive senses of ‘believe’ are pertinent. In the broad and inclusive sense used above and throughout this work, believing is accepting as true regardless of whether done with absolute certainty, with a lower level of certainty, or with no certainty. A person believes every one of his or her own cognitions and every one of his or her own opinions.

In both of the narrow and exclusive senses, discussed but not used here—sometimes signaled by a word such as ‘merely’—no proposition is known to be true *and* believed to be true by the same person: “beliefs exclude knowledge”. A person aware of merely believing will sometimes answer, “I believe so” where a person who knows would answer simply “Yes”.

Of course, it makes a difference whether only certain knowledge is excluded or whether all knowledge including probable knowledge is excluded. The less exclusive narrow sense in which only certain knowledge is excluded is used mainly in mathematics and philosophy. In this sense, most scientific and medical knowledge is mere belief. In the more exclusive narrow sense of ‘belief’, all knowledge, certain and probable, is being excluded. In this sense, the mere beliefs are only the credences. In this sense, all mere belief is entirely subjective in the sense of coming from within the believing subject: the process leading to it includes no component, however small, based on objective evidence and deliberation. For example, in some cases, all of what was taken as evidence was in some sense constructed by the believer. In other conceivable cases, the believers made no effort to square their judgments with objective reality: loyalties, preconceptions, fears, hopes, or other subjective factors were dominant.

Many mere beliefs can be expected to be false, regardless of whether 'mere belief' is taken in the less or the more exclusive sense. However, nothing precludes subjective beliefs from being true—in some cases by accident, so to speak. True mere beliefs in the more exclusive sense are what we called true credences above. By contrast, in the broad and inclusive sense of 'believe' used in this work—sometimes signaled by suffixing words such as 'in the broad sense'—every proposition known to be true is believed to be true: "beliefs include knowledge". More explicitly, for a given person, the set of their beliefs includes their cognitions. In order to speak of belief that a proposition is true, we sometimes speak of *belief in* the proposition. Sometimes, "I think that" is interchangeable with "I believe that".

Certitude is not the test of certainty.—O.W. Holmes, Jr.

7 Certitude and Certainty

Certitude is the subjective feeling of assurance of the truth of a proposition.[20] Certitude can be the result of thorough objective investigation which started from a suspension of belief or even from doubt or disbelief, and it can also arise without investigation or be the result of deception, rationalization, indoctrination, error, or hallucination, to mention a few. O.W. Holmes, Jr., reminded us that "Certitude is not the test of certainty".

Certainty, as used by Holmes and in this work, is not a feeling at all; it is the state of having knowledge in the broad sense. The expression 'false certitude' is sometimes used for misdirected certitude: a feeling of assurance toward a false proposition. It is also used for improperly derived certitude: certitude arrived at by persuasion, enthusiasm, illusion, fallacious reasoning, and the like.

False certitude is analogous to false trust, false distrust, false security, false danger, false guilt, false righteousness, false pride, false shame, and the rest. Holmes could have added, for example, that guilt is not the test of immorality: the world has known people who had done nothing wrong but who were consumed with guilt. Freud and Tarski lamented these cases.

Aristotle said that every person by nature desires knowledge. Peirce disagreed; he said that every person by nature desires belief. According to Aristotle, the goal of inquiry is the possession of truth. According to Peirce, the goal of inquiry is the cessation of doubt. Aristotle sought certainty, the possession of truth.[21] Peirce sought certitude, the subjective feeling of certainty.[22] Aristotle's view relates more to the spectra of certainty;

[20]The expression 'the truth of a proposition' should not be detached from its context. By 'assurance of the truth of a proposition', we mean assurance that the proposition is true. It is a mistake to think that the truth of a proposition is an entity separate from the proposition itself.

[21]There are two difficult issues here for the Aristotle scholar. The easier is whether our indirect quotation of Aristotle—that every person by nature desires knowledge—is a fair interpretation of the famous first line of *Metaphysics* [2]. The other issue is whether Aristotle held the general view that we attribute to him. Scholars we consulted do not all agree. After a nuanced discussion, David Hitchcock (per. comm.) concluded: "So I think that it is fair to say that Aristotle took one cognitive goal of all human beings to be the possession of truth." Very recent scholarship concurs. See Anagnostopoulous [1, 102f].

[22]Perhaps the classic expression of Peirce's views is in Sects. III, IV, and V of his famous 1877 article "The Fixation of Belief" [37] reprinted in the 1992 Houser and Kloesel volume [40, 109–123]. The

Peirce's relates more to spectra of certitude. Of course, it is a matter for scholars to decide whether, for the respective thinker, either view is a "mature" one.

Absolute, *apodictic*, or *mathematical certainty* is the state of having knowledge in the strict sense. *Scientific certainty* and *moral certainty* are two states of having probable knowledge. It is unfortunate that the words "I am certain" often indicate certitude, not certainty. Certitude is a feeling of confidence in a belief. Some authors such as Tarski write of "intuitive certainty" and of being "intuitively certain" to refer to certitude. It is also unfortunate that certitude, the feeling, is often not distinguished from the scientific or moral certainty it sometimes reflects. It is even more unfortunate that the two words 'certitude' and 'certainty' are sometimes used interchangeably with one meaning. It is confusing that sometimes when the two meanings are distinguished, the words are used with the meanings reversed.

As Holmes said, certitude is not always based on certainty. Moreover, certainty does not always give rise to certitude; and when it does, time might pass between the achieving of certainty and the feeling of certitude. Even absolute certainty is not always or not immediately accompanied by certitude—especially in cases where the knower is at first surprised, delighted, or dismayed to find out that the proposition is true. In fact, for modest and objective investigators into important issues, sometimes the more certainty they achieve, the more they grasp the complexities and the less certitude they might feel. As implied above, philosophers disagree on whether absolute certainty is achievable. John Stuart Mill had certitude that "There is no such thing as absolute certainty".

Many of a person's beliefs, including all those based on testimony, are not their certain knowledge, and some, even some that are true, are probably not certain knowledge for most. An example is the famous Fermat Theorem: given any three numbers that are all the same power exceeding two, no one is the sum of the other two. This implies that no cube is a sum of two cubes, that no fourth power is a sum of two fourth powers, and so on. Corcoran, for example, thinks that he is fully justified in believing this. Part of his justification is based on his knowledge that mathematicians he has reason to respect have testified in print that it has been proved to be true by a proof that has been carefully studied and found to be cogent by qualified experts. Here is a justified and true belief of Corcoran's that is not his certain knowledge, i.e. not his cognition; it is, of course, one of his scientific certainties. In contrast, for the other author, Hamid—whose expertise does not extend as far into mathematics as Corcoran's—the Fermat Theorem is not a scientific certainty: it is merely a moral certainty. Probably most people who believe Fermat's Last Theorem have less evidence than Corcoran or even Hamid.

But the proposition in question is a justified true belief of dozens of mathematicians for whom it *is* knowledge. With regard to the Fermat Theorem, while Corcoran has certitude but not absolute certainty, some mathematicians have certitude and absolute certainty. Corcoran has probable knowledge that the Fermat Theorem is true. In fact, he thinks he has moral certainty, even scientific certainty.

---

interpretation putting Peirce in diametrical opposition to Aristotle is almost universally shared by Peirce scholars as being a view that Peirce actually held at the time of the article. However, Peirce's writings are replete with subtlety, irony, and scathing sarcasm. So much so, that it is hard to be certain that he was not actually expressing the opposite of what he wrote. Moreover, it might well be that he later came to embrace in a nuanced form a view he had formerly ridiculed in a naïve and exaggerated form.

There is no such thing as absolute certainty.— John Stuart Mill.

8 The Spectra of Certitude

In a given person's lifetime, or even a relatively short interval of years or days, their degree of certitude in a given belief may vary from absolute or near absolute certitude though lesser degrees until there is no certitude whatever, when the person's belief might be said to lack all conviction. At such a point, doubt may start to accumulate until a maximum is approached.

Of course, a similar observation applies to each disbelief. Moreover, as we all know, one day's belief might be replaced over time by the diametrically opposed disbelief. However, let us focus on a given belief belonging to a given person. Analogous to the spectra of certainty there are spectra of certitude, which range from a maximum of certitude through ever more diminishing level until a state of subjective neutrality is passed and doubt begins to increase to its maximum.

To picture a spectrum of [objective] certainty with the corresponding spectrum of [subjective] certitude, we may represent the former spectrum as a horizontal axis that extends from credence (starting point) on the left to cognition on the right. We may then represent the certitude spectrum as a vertical line on the left whose neutral midpoint lies at the point of credence. The high point of the vertical axis above the horizontal axis represents maximal certitude; the low point of the vertical axis below the horizontal represents maximal doubt.

As a given cognitive process moves the knower's level of certainty from credence to cognition, certitude sometimes may increase, and the curve rises above the axis of the certainty spectrum. However, in some cases, as evidence increases doubt increases and the curve goes below the axis of the certainty spectrum.

An extreme skeptic may experience doubt when they have certainty, even when there is cognition. An extreme *paraskeptic*—to coin a word—may experience certitude even when there is only credence. Such extreme skeptic's subjective feeling of doubt is disconnected from their objective state of certainty in one direction; such an extreme paraskeptic's subjective feeling of certitude is disconnected from their objective state of certainty in the other direction.

Philosophers counsel thinkers to adjust their level of certitude to be proportional to their level of certainty without suggesting how this can be accomplished.

The certitude-doubt axis is independent of the certainty axis. Besides, a person can have unbounded certitude in a false proposition and unbounded doubt in a true proposition.

Doubt is more often the mark of knowledge than certitude is. Certitude is more often the mark of error than doubt is.—Frango Nabrasa

9 Certain Knowledge

How does a person go about arriving at certain knowledge of the truth of a proposition even if, in at least some cases, knowledge in the strict sense is the ideal limit of a process that can never be completed—except perhaps in mathematics? Let us use the word 'hypothesis' for a proposition not known to be true and not known to be false by a given knower. In the first place, it is necessary to understand the hypothesis to be investigated. Next, it is necessary to connect with the reality that the proposition is about in order to acquire from it evidence sufficient to ground a judgment that the proposition is true. Third, it is necessary to marshal the evidence, to bring the evidence to bear on the hypothesis. Finally, it is necessary to see that the evidence is conclusive and to accurately judge on the basis of the understanding and marshaling that the proposition is true. A belief that resulted from successful completion of this process is said to be *cognitively* grounded or justified. Husserl was referring to knowledge in this strict sense when he said that in having knowledge, "we possess truth as the object of a correct judgment" [29, 130].

In the case of the proposition that no square number is twice a square number, which was probably known to be true by Socrates, Plato, Aristotle, Leibniz, Pascal and many others, the evidence phase included reviewing previously known arithmetic propositions and the marshaling phase included inferring the hypothesis from them by logical deduction.

Certain knowledge is cognitively justified true belief. In this context, the word 'true' is redundant in the sense that every cognitively justified belief is true. A belief that a given proposition is true is cognitively justified only if there was a successfully completed four-step method or its equivalent. Moreover, a proposition that is a certain person's true belief but not now their knowledge can become cognitively justified and thus become knowledge if the believer successfully completes the four-step method.

Notice that in regard to cognitive justification, praise and blame are often beside the point. A person might flawlessly make every effort to apply the four-step method to a proposition that they believe and yet fail because of circumstances beyond their control such as the necessary evidence having been destroyed.

The verb 'justify' is ambiguous. In several other senses of 'justified', the word 'true' is not redundant in the sentence 'knowledge is justified true belief'. Unless 'true' is redundant, the sentence expresses a false and misleading proposition, as we show below—in agreement with Plato, who criticized such formulations toward the end of the *Theatetus*. Preus [43, 93] wrote: "In the *Theaetetus*, the hypotheses that knowledge might be "true belief" (*alēthē doxa*) or "true belief plus an account" (*logos*) are discussed and refuted". Other qualified scholars concur; see, for example, [42, 59].

In some of the other senses, justifying a belief involves explaining something to others: perhaps why I should not be blamed for having the belief or why other people in my circumstances would have come to the same conclusion. No matter how these explanatory senses of justification are spelled out, it is clear that knowledge is not justified true belief. Gaining knowledge that a proposition is true does not require explaining anything to anyone.

As explained above, acquisition of knowledge comes about through a personal process that begins with apprehending a proposition and ends with a judgment. There is no need for the knower to explain anything to anyone. There is no room in the process for accounting to others. In fact, the suggestion that knowledge requires explanation is contrary to the principle of the autonymy of knowing—a principle emphasized in this paper.

More generally, in any sense of 'justify' in which a false belief is justified, it is not the case that justified true belief is knowledge in the strict sense—there are indefinitely many propositions that could become justified true beliefs without thereby becoming my certain knowledge. The reason is based on the fact that from any false proposition indefinitely many true propositions are deducible. If we justifiably believe that you own a new pen when in fact your pens are all old, then we could justifiably believe without knowing the true proposition that you own a pen—if we were to deduce the latter truth from the former falsehood. Our belief that you own a pen would have been formed by a flawed process—in this case inferring from a false premise, one form of begging-the-question [11, 22f].

In the cognitive sense of 'justify', every justified belief is true and, in the strict sense of 'knowledge', justified belief is knowledge; knowledge is justified belief. To be even more explicit, cognition is cognitively justified belief. The formula 'knowledge is justified true belief' is dangerously flawed and should never be used without adequate qualification. In statements using the formula without qualification, 'justify' can be interpreted either cognitively or non-cognitively: if the former 'true' is misleadingly redundant; if the latter, the formula expresses a falsehood.

10 Speaking of Propositional Knowledge

Previous paragraphs present one of many ways of organizing propositional knowledge for purposes of discussion. Every item of propositional knowledge is someone's true belief that something is the case. These beliefs have all been established with some degree of certainty. Each was derived by a process involving some objective considerations.

We started with knowledge in the strict sense, the most firmly and objectively established beliefs, which we called cognitions.[23] These beliefs are known with certainty and are thus also known as certain knowledge. If the word 'knowledge' is used in the strict sense, 'certain knowledge' is a redundancy, a mere rhetorical flourish, and 'probable knowledge' is an oxymoron. When 'knowledge' is used in a less strict sense, we can speak of certain knowledge without redundancy and of probable knowledge without contradiction. Even then, putting the adverb 'absolutely' in front of 'certain knowledge' is a mere rhetorical flourish exactly analogous to other uses of the adverb in expressions such as 'absolutely true' and 'absolutely perpendicular'.

The established true beliefs that are not cognitions are all called probable knowledge. The degree of firmness starts very high but then shades off imperceptibly through the less and less probable, tending toward but excluding true credences, which are not established at all and not knowledge. The most firmly established of true beliefs that are not cognitions have scientific certainty. Moving on from there, we reach those with moral certainty, less firmly established but still sufficiently so to serve as the basis for responsible action and warranted assertion. After that we eventually come to beliefs that hardly merit being called probable knowledge but are still awkwardly (but strictly) so-called. In this usage, putting 'scientifically' and 'morally' in front of 'certain' weakens the adjective in the same way that 'partly' and 'lightly' weakens 'cooked', 'scrubbed', 'seeded', and other such expressions.

[23] A person's most firmly established beliefs are rarely those they believe most firmly, i.e. those of which they are most strongly convinced. This point relates to the contrast, dealt with elsewhere in this work, between the objective state of certainty and the subjective feeling of certitude.

As observed above, in the sense of ‘probable’ used here, all probable knowledge is knowledge. There are other senses of ‘probable’ recognized by dictionaries. In another sense, saying that a given proposition is probable knowledge might mean that it is likely to have been investigated and found to be true. In the latter sense, probable knowledge is *probably* knowledge, but it might not actually be knowledge; in this sense, then, not all probable knowledge is knowledge. In fact, in this sense, not all probable knowledge is true. In the sense of ‘probable’ used in this work, all probable knowledge is true.

Cognitions are said to be known with mathematical certainty, not because such cognitions are necessarily about mathematics, but only because mathematical cognitions are often taken as paradigm examples. The mathematical community is known to aspire to the highest standards of clarity, transparency, rigor, and cogency. Likewise, beliefs said to be known with scientific certainty are not necessarily about science; they are so-called only because this level of certainty is often aspired to or achieved by the scientific community. Finally, beliefs said to be known with moral certainty are hardly ever about morality; they are so-called only because this level of certainty is often required for beliefs on which moral action is to be based.

People might have cognitions that they just locked up their house. Shortly after locking up, they might have scientific certainty that the house is locked. But after enough time has passed, perhaps they can have only moral certainty that the house is locked. And after even more time has passed they might not be able to have even moral certainty.

Notice that this organization applies only to propositional knowledge, not to objectual knowledge or operational knowledge. Moreover, it also does not apply to knowledge in the sense of scientific competence or expertise. As explained above, by ‘expertise’ we refer to the result of years of observation, experimentation, investigation, and deliberation that fine-tunes instincts of the dedicated scientist. For example, Newton probably had more scientific competence and expertise in mechanics than any person who came before him, and yet, after Einstein, many of Newton’s beliefs about mechanics have been shown to be false, even though they are close enough for many applications. In fact, many scientists who have extensive knowledge in their respective fields expect that future researchers will find to be false many if not all of the propositions they currently believe to be true. If scientific knowledge were to be measured by propositional knowledge, the propositions currently regarded as established truths, then there would be little or no progress in science.

Focus on propositional knowledge to the exclusion of expertise is unfortunate, even in mathematics or philosophy. Studies of propositional knowledge, including this work, will probably be valued less for specific achievements than for preparing the way for a broader and more inclusive study of knowledge and for underlining the need for more attention to expertise, the fourth kind of knowledge mentioned. The fact that contemporary epistemology focuses almost exclusively on propositional knowledge is especially regrettable.

It is true that traditional philosophy often extolled the seeking of wisdom, which is inseparable from expertise. But the fact that expertise and the accompanying wisdom require practice, experience, learning from mistakes, and other mundane activity is rarely mentioned. Even more unfortunate from our perspective is that in many discussions of wisdom-seeking no mention is made of propositional, operational, and objectual knowledge.

11 Cognitivism, Probabilism, and Skepticism

Among the tensions that pervade discussions of knowledge and belief is the perennial issue of whether knowledge in the strict sense is ever achievable or whether it is merely an ideal to which objective people can only strive but never fully reach. Our personal opinion is that it is achievable—but often only with great difficulty and often not at all. We believe that the *main* role of the concept of cognition in our lives is serving as an ideal standard by which to measure our performances: an ideal goal to strive for and a constant reminder of the shakiness of many of our beliefs. It would serve these important purposes even were there no cognitions, no knowledge in the strict sense. When Aristotle wrote that every human by nature desires to know, he was thinking of knowing in the strict sense.

With respect to any given proposition, we can distinguish three philosophic viewpoints a given person might have: cognitivism, probabilism, and skepticism. *Cognitivism* holds that the proposition is or can be known in the strict sense, i.e. that it or its negation is or can become certain knowledge or cognition. *Probabilism* holds that the proposition cannot be known in the strict sense, but that it or its negation is or can become morally certain or even scientifically certain. *Skepticism* holds that the proposition cannot be known in the strict sense, and that neither it nor its negation is or can become even morally certain, much less scientifically certain.[24]

Both authors agree with cognitivism with respect to many but not all mathematical propositions. They also agree with probabilism with respect to many but not all propositions about the material world. For example, they think that they have scientific certainty that smoking is deleterious to health. However, they differ on philosophical propositions: Corcoran agrees with skepticism but Hamid agrees with probabilism with respect to most but not all philosophical propositions. For example, Corcoran thinks that it is impossible to know even with moral certainty whether there is cosmic justice, i.e. whether each good act will be rewarded in proportion to its degree of goodness and each bad act will be punished in accord with its degree of badness [21]. However, Hamid thinks that this "cosmic justice hypothesis" can be settled with moral certainty, perhaps even with scientific certainty.

12 Understanding Propositions

Understanding a proposition is grasping its truth-condition, knowing what its being true would be and what its being false would be. A person who understands a given proposition will often be able to imagine what it would be like to affirm it, or to deny it. A person who understands a given proposition can wonder whether it is true or false. Such a person can recognize in simple cases its implications—what it implies, what follows from it. They

[24] The word 'skepticism', or 'scepticism', derives from the Greek verb meaning "to consider carefully", which was taken by some to mean "to consider so carefully that no conclusion is reached". The Greek skeptic (*skeptikos*) did not subscribe to the view called 'scepticism' above; the skeptic meticulously avoided subscribing to any view at all. Today, the word 'skepticism', or 'scepticism', is used in various senses, often as above, rarely if ever in the etymological sense. See Preus [43, 237–8].

can also recognize in simple cases its implicants—what implies it, what it is implied by. In addition, again in simple cases, they can recognize what it contradicts, what is inconsistent with it. In some cases, a person who understands a proposition can look for methods to settle it affirmatively or negatively.

Before a person can begin to marshal or even acquire evidence by which to judge that a proposition is true, it is necessary to understand the proposition. In fact, as Frege said in several places, it is necessary to understand a proposition before one can make a judgment concerning it.[25] He suggested without explicitly saying so that understanding is necessary before we can wonder whether it is true or false, before one can become aware that it is a hypothesis. There are many propositions that are widely understood but are not known to be true and not known to be false by anyone—or so it is said. Clearly, in most if not all cases, it would be impossible to know with absolute certainty, to have a cognition, that a given proposition is neither known to be true nor known to be false by anyone. This would require exhaustive knowledge of the mental states of every person now alive.

The Goldbach Hypothesis [23, p. 39] is that every even number exceeding two is the sum of two prime numbers. This proposition is easy to understand but difficult to settle. Despite the fact that many able mathematicians have spent long years trying to determine whether it is true, none have succeeded.

It is possible to understand all of the concepts in a proposition without understanding the proposition—just as it is possible to understand each step in a proof without understanding the proof and to grasp each note in a melody without grasping the melody. In all cases, even adding to apprehension of the constituents a grasping of their order does not account for an understanding of the whole. One of the most perplexing problems in the theory of propositional knowledge is that of how a proposition is understood—assuming that its constituent concepts are understood. However, the problem is even more challenging to philosophers such as Frege who believed, or at least said, that we apprehend a proposition's constituent concepts only after apprehending the proposition, that the concept is grasped by analyzing a proposition containing it. Cf. [41, p. 140].

In order to understand a proposition it is normally not necessary to know what understanding a proposition is. In order for a person, a young student for instance, to understand the proposition that there are four single-digit square numbers it is not necessary for the student to know what it is to have understanding of a proposition. However, in order for a person, Hamid for instance, to understand the proposition that a proposition must be understood before it can be known to be true it is necessary for him to know what it is to have understanding of a proposition. However, this is not a typical case. A child can come to know that the family dogs each have four legs without it understanding the proposition that it understood the proposition that the family dogs each have four legs. Our assumption in this article is that understanding propositions precedes having knowledge of what understanding propositions is.

---

[25] Frege [27, 62] speaks of understanding as "grasping" the proposition and he speaks of judging as "acknowledging" its truth. See Frege [28, 329]. He seems oblivious of the fact that understanding is an act-process that takes time to complete. He seems likewise oblivious of the process intervening between understanding and judging. Moreover, he is vague about the nature of judging.

At the 1990 Buffalo Church Symposium, Hartley Rogers said that John Myhill once called him up to ask whether a certain hypothesis that had just occurred to him had been settled. Rogers replied that it was known to the field as Myhill's Theorem.—Sriram Nambiar

13 Knowing That One Knows

Let us start with an example of a cognition and then consider what else would be required to know that it is a cognition. Corcoran *believes* that he knows that every [*sc*. geometric] square is equal in area to the sum of two smaller squares of different sizes. He also believes he knows that each square is equal in area to the sum of two smaller squares that are equal to each other. *Every* square is equal to twice a square. If you draw the two diagonals, you make four isosceles right triangles, any two of which make a square.

He remembers discovering this and proving it in connection with one of his many readings of Plato's *Meno*. This geometric proposition is related to the Pythagorean Theorem, which might have been discovered and proved by Pythagoras.

In order for Corcoran to know that this proposition is true, it was not necessary for him to recall the details of his own thought process. But for him to *know that he knows* this proposition to be true, it is necessary for him to be able to recall the processes by which he gained that knowledge and to verify that the steps were properly carried out and completed. For example, he must verify that, *at the exact time of the process*, he succeeded in understanding the proposition, he succeeded in gathering sufficient evidence, and he succeeded in properly bringing the evidence to bear on the issue of the truth of the proposition. We doubt whether this is even possible.

The proposition itself is about geometric squares; it is not about Corcoran, and it is not about a proposition. The proposition that he knows that it is true is about him (his past cognitive history) *and* about the proposition; it is not about geometric squares. Knowing that one knows is different than knowing, and far more complicated and problematic. Corcoran believes that he knows the geometrical proposition with mathematical certainty; but he thinks he has at best moral certainty that he knows that he knows.

Similar points apply in the case of belief. We imagine that there are propositions that we believe but concerning which the issue of *whether* we actually do believe them has never come up. In such cases, we believe that it is true, but we do not even believe, much less know, *that* we believe it. There are difficult issues about knowing that one knows even though it is often very easy to know that one does *not* know. In order to know that one knows, it is necessary to know that one believes. But, by what criterion do we determine that we believe a given proposition? We have made mistakes about what we thought we believed. For example, Corcoran now knows that he was wrong in believing *that* he believed certain propositions. He now knows that there are propositions that he did not even understand but that he believed that he believed. People he trusted led him to believe that he believed propositions that, as he now knows, he did not believe. As he reflects on his youthful views, he is frequently surprised at what in his youth he thought were his beliefs.

One of the most confusing mistakes that we can make regarding cognition or even knowledge in the broad sense is thinking that in order to know it is necessary to know that we know. There are many people who have cognitions who do not even understand, much less know to be true, the proposition that they know the cognition to be true. A person who knows that five plus seven is twelve need not understand the proposition that the proposition that five plus seven is twelve is known to be true. The idea that knowledge that a proposition is true requires knowledge that the proposition is known to be true leads to an infinite regress and thus to skepticism, the view that knowledge is impossible. Persons who know that we are writing about knowledge and opinion need not know that they know that they know that they know that we are writing about knowledge and opinion. Why anyone should have thought otherwise has baffled us for years.[26] We should never forget that the skeptics do not feel bound by logic or by the requirement of testifying in accord with their own knowledge—which they deny they have. But, it is not just skeptics who held that knowledge requires knowledge of itself. Aristotle held that having knowledge of an axiom (*arche*) required knowledge of the knowledge [1, 68].

The above treats knowing with absolute certainty that one knows with absolute certainty. Knowing with scientific certainty that one knows with scientific certainty is equally problematic. Knowing with scientific certainty that one knows a given proposition with scientific certainty requires knowing that the proposition is true, but it does not require knowing with absolute certainty that the proposition is true. Thus, it might be possible to know with scientific certainty, given a proposition that cannot be known with absolute certainty but is known with scientific certainty, that it is known with scientific certainty. However, it is impossible to know with absolute certainty, given a proposition that cannot be known with absolute certainty but is known with scientific certainty, that it is known with scientific certainty.

The problem of knowing *that* one knows is closely related to the problem of knowing *how* one knew. For example, can we determine of a given cognition whether it was inferred from previous cognitions or whether it was achieved by some other means?

Beliefs formed by logical deduction from previous beliefs were called *inferences* and those formed without deduction were called *intuitions*. In this sense, not every intuition is known to be true and, in fact, not every intuition is true. Intuitions that are known with absolute certainty, i.e. that are cognitions, are called *cognitive intuitions* or *intuitive cognitions*. Cognitions formed by logical deduction from previous cognitions can be called *cognitive inferences* or *inferential cognitions*.

The word 'intuition' has other meanings, of course. One relatively common usage is in a way broader and in a way narrower. It is broader in that it applies to objectual knowledge as well. It is narrower in that it does not apply to beliefs that are not cognitions. The common usage just mentioned is in Bruce Russell's entry "Intuition" [45] in the *Cambridge Dictionary of Philosophy*. It also agrees with the 1868 Peirce article "Questions Concerning Certain Faculties Claimed for Man", which has a long and interesting footnote on the history of the word. Peirce wrote [40, p. 11]: "*Intuition* here will be nearly the same as 'premise not itself a conclusion'; the only difference being that the premises and conclusions are judgments, whereas an intuition may be any kind of cognition whatever".

Some philosophers from previous centuries believed that it is possible to determine with certainty whether a given cognition was an inference or an intuition. In other words, it was held that in every case a person who has a given cognition can have a

[26]We are aware of the literature on "knowing that one knows" centered on or stemming from "epistemic logic", for example, Hilpinen [31]. The definitions of knowing used in that literature are so alien to those used here that little written there is relevant.

cognition about that given cognition, viz. either a cognition that the given cognition is an inference or a cognition that it is an intuition. Moreover, Aristotle, Frege [26, §3, §4], and others have been interpreted as holding that, in case of a cognitive inference, knowers have the capacity to trace their chains of reasoning back and back until they come upon cognitive intuitions, propositions that they knew to be true without inference. The cognitive intuitions were called axioms, principles, first principles, primitive truths, or something similar. Even though the overwhelming majority of mathematical cognitions were held to be the results of inference from cognitive intuitions, not one example of such backwards tracing has been presented and no one has ever proposed a criterion for determining of a given belief whether it is an axiom or an inference. Needless to say, we are skeptical concerning the hypothesis that knowers have the capacity to trace each of their cognition-producing chains of reasoning back to cognitive intuitions.

Each of the four steps in the method of cognition admits of slippage. Coming to know that a proposition is true is like securing a house having four complicated locks. If A person's attention wavers at any lock in the process, a person might have locked the house without knowing so. And a person might be mistakenly certain of having locked the house regardless of whether they actually had locked it.

14 Results, Intuitions, Inductions, and Inferences

We need a word for a belief that is the result of a cognitive process however complete or incomplete, successful or unsuccessful, it may have been. In other words, we need a word for beliefs that are not entirely subjective. We propose *result*. Every cognition is a result, and so is every moral certainty and every scientific certainty. Every result that is not the conclusion of a chain of deduction is believed on the basis of experience, whether sense-based, intuition-based, or mixed, whether reliable or unreliable. Using this terminology, we can say that every belief is either a result or a credence.

We propose *induction* for a belief arrived at through experience and not through deduction. We are not the first to adopt this usage. Aristotle's word for this kind of belief, *epagoge*, is routinely translated 'induction'. Further, we can say that every result is either a deduction or an induction, using the word 'deduction' in the broad sense in which 'fallacious deduction' is not oxymoronic.

One famous induction is Aristotle's belief that all swans are white (*Prior Analytics* A4.26b7-14 [3]). Another example of an induction is Archimedes' belief in his Law of Buoyancy that an immersed body is buoyed by a force equal to the weight of the displaced fluid. Corcoran's initial belief in Archimedes' Law was also an induction in this sense, but unlike Archimedes' belief, his involved very little experience: it was largely based on his teacher's testimony and the experiment the teacher showed the class how to do.

There is an important difference between inductions involving sense experience such as the Archimedes example and those based on abstract experience such as those traditionally attributed to Thales, Pythagoras, Euclid and other mathematicians. Those involving sense experience are not normally given a special term, but they may be called *empirical inductions*. Those involving abstract experience are often called *mathematical intuitions*. Mathematical intuitions are sometimes misleadingly said to be self-evident even though nothing is evident to anyone unless they have gone to the trouble to understand it, to experience the relevant reality, and to complete the cognitive process. The ambiguous words 'apriori' and 'aposteriori', which are not *used* in this investigation, have been used to distinguish mathematical from empirical inductions—and in several other ways as well [16, 1–3].

Every mathematical cognition is either an inference or a mathematical intuition. Of course, no cognition is both an inference and a mathematical intuition. But there is no reason we can see for not thinking that one and the same proposition that is the content of one person's cognitive inference might also be the content of another person's cognitive intuition. In fact, we would say that one and the same proposition which is the content of one person's cognitive intuition might also be deducible from one or more of the same person's cognitive intuitions.

> When is a statement true? There is a temptation to answer, 'When it corresponds to the facts'. And, as a piece of standard English, this can hardly be wrong. Indeed, I must confess I do not really think it is wrong at all.—Austin [6] and [7, 89].

15 Truth and Knowledge

In this work, the word 'true' is used in the traditional classical sense traced by Tarski back to Aristotle. If something is the case, then it is *true* that it is the case. And conversely, if it is true that something is the case, then it *is* the case. To separate the classical sense of 'true' from its near neighbors, we can observe that in the classical sense the following are obvious to everyone who understands them.

The proposition that Aristotle read Plato is true if and only if Aristotle read Plato.
The proposition that Aristotle read Euclid is true if and only if Aristotle read Euclid.
The proposition that Euclid read Aristotle is true if and only if Euclid read Aristotle.
The proposition that Euclid didn't read Aristotle is true if and only if Euclid didn't read Aristotle.

These are obvious even though the four proposition they are about are not obvious to many. The proposition that Aristotle read Plato is known with moral certainty to be true by persons who have read Plato and Aristotle. The proposition that Aristotle read Euclid is known with moral certainty to be false: Aristotle died before Euclid wrote anything. The proposition that Euclid read Aristotle is a hypothesis—highly improbable to many given what the two wrote, but far from settled.

Every proposition is either true or false. But not every proposition is either known to be true or known to be false, by either of us or by any other person. In fact, not every true proposition is known to be true. Corcoran knows that there is a true proposition not known by him to be true. He can say this even though he knows that he cannot give an example. It is logically impossible for Corcoran to give a proposition that he knows to be a true proposition not known by him to be true. There are many counterexamples for the proposition that every true proposition is known to be true. And, for all we know, "every perfect number is even" might be one of them. But we can never give a counterexample and know that it is a counterexample [14].

Truth includes knowledge and goes well beyond knowledge: the set of true propositions includes as a relatively small subset of the set of propositions known to be true. And it includes much more—namely the vast expanse of true propositions not known to be true.

No proposition is both true and false. And no proposition is both known to be true and known to be false, by either of us or by any given person. In fact, no proposition is both known to be true by either of us and known to be false by some other person. Nevertheless, there are many propositions believed by one of us to be true and believed by others to be false.

There are many modern philosophers who believe that Corcoran was wrong to say that he has certain knowledge of the fact that no square number is twice a square number. It is not that they have any doubt of the fact, rather they disbelieve that Corcoran has absolute certainty of it. We are glad to discuss the issue with them. Without them this work would be less interesting.

Above we explained what we mean by saying that a certain proposition is true. For example, the proposition "No square is twice a square" is true. Better, the proposition "No number which is the sum of two numbers of the same power exceeding two is also of that same power" is true. Even better, the proposition "No perfect number is odd" is true. The reason we moved through these examples has to do with the following facts: we think we know the first to be true, we think that the second is true but that it is not known to be true by us, and we neither believe nor disbelieve the third. Moreover, we think that the third is not known to be true or known to be false by anyone. Nevertheless, as Tarski taught, we are fully warranted in stating the following:

> The proposition "No perfect number is odd" is true if and only if no perfect number is odd.
>
> The proposition "Every perfect number is even" is true if and only if every perfect number is even.

These contrast with the following:

> The proposition "No perfect number is odd" is known to be true if and only if some person knows that no perfect number is odd.
>
> The proposition "Every perfect number is even" is known to be true if and only if some person knows that every perfect number is even.

In order for either of the two arithmetic propositions to be true, it is not necessary for anyone to do anything. In order for these propositions to be *known* to be true, it *is* necessary for someone to do something—something very difficult that no one has yet managed to accomplish. In addition, for these propositions to be *known* to be true it *is* necessary for them to be true. It is not that the knowing would be what would make them true, as some philosophers and mathematicians have held.

In many cases, in order for a person to responsibly *state* that a proposition is true, it is necessary for that person to know that it is true. But this should not be taken as evidence that knowledge is necessary for truth. What was just said of the two arithmetic propositions could be said with equal warrant of the four propositions about the two arithmetic propositions.

> A cerebral habit of the highest kind, which will determine what we do in fancy as well as what we do in action, is called a *belief*. The representation to ourselves that we have a specified habit of this kind is called a *judgment*. Peirce [38, §1].

16 Beliefs and Disbeliefs

For a given person saying, "I *believe* a certain proposition" amounts to saying "I believe it to be true", or "it is one of my beliefs". Saying "I *disbelieve* a certain proposition" amounts to saying "I believe it to be false", or "it is one of my disbeliefs". It is important to be explicit about some elementary points. Although every proposition that is not true is false, it is not the case that every proposition not one of a person's beliefs is one of their disbeliefs. There are many propositions a certain person has never thought of, and among those they have thought of, there are many they have no belief concerning. One way this point is missed is that the words "I do not believe it" are used to say the same as "I disbelieve it", not simply "it is not one of my beliefs". It would be better either to say "I do not believe it; I need to see how the evidence available is sufficient for concluding it" or else to say "I disbelieve it; I have sufficient evidence to the contrary". The atheist disbelieves what the theist believes; the agnostic does not believe or disbelieve what the theist believes—and therefore does not believe or disbelieve what the atheist believes.

In judging, a fresh belief is formed, often a belief in the truth of a proposition not previously believed by the person judging, but people often form a *new* belief in (the truth of) a proposition they previously believed. Judgment creates belief and thus contrary to Peirce's usage in [38] and [39], belief begins as judgment ends. Of course, once beliefs are formed, it can happen that the believers perceive their having those beliefs and then form judgments to that effect. This creates beliefs about beliefs, which we have called *belief-beliefs*. However, as a general point, this article's usage conflicts with Peirce's 1880 usage quoted above.

As Frege, Husserl, and others taught, propositions are timeless, "beyond time"—to use Husserl's phrase. In contrast, statements, judgments, and beliefs are dated [20]. The belief Corcoran formed years ago when he first learned that 10° Celsius is 50° Fahrenheit started years before the new belief he formed of the same proposition today. We imagine that many readers had to do a calculation to achieve belief that 10° Celsius is 50° Fahrenheit and could substitute their own name for Corcoran's above.

In contrast, the belief Corcoran formed years ago when he first learned that seven plus five is twelve might have persisted uninterrupted to this day even though weeks may pass without him reflecting on the fact or on the circumstances of his learning it. Often, our beliefs persist for years even when not put to use, so to speak.

Each belief comes into existence sometime during the life of a person who believes it and perishes no later than the death of the believer—earlier if memory fades. Each belief depends for its existence on its believer. No two persons have the same belief although in many cases two persons have different beliefs with the same propositional content—as emphasized earlier in this article.

The word 'belief' is frequently used elliptically for 'the content of a belief', the proposition believed. In this sense, there are beliefs that no one still believes: there are beliefs in the *non-temporal* sense that are not beliefs in the *temporal* sense. In the temporal sense, none of a person's mere beliefs can ever become anything else; a mere belief can never become a cognition—even if the person subjects the propositional content to a cognitive process that produces knowledge. In such a case, the belief in the non-temporal sense became a cognition, but the cognition is a belief with a later starting date. Of course, the words ‘statement’ and ‘judgment’ both have temporal and non-temporal senses.

In contrast, the act of stating per se does not form new beliefs, although there are simple cases in which judging and stating are simultaneous, or very nearly so. Usually the date of a statement is not the same as that of the belief stated. Sometimes the statement is made before the belief is formed. Of course, propositions are not dated at all even though, as just mentioned, the word ‘belief’ is sometimes used non-temporally to refer not to the dated belief but to the undated proposition believed. This is the case when we say that Euclid and Pythagoras had the same belief about right triangles even though one was born centuries after the other died.

As said above, propositions do not change in the process of becoming belief or knowledge. The proposition about right triangles was not changed as Pythagoras came to know it. The proposition about right triangles known by Pythagoras was not changed as Euclid came to know it. John Dewey is sometimes interpreted as saying the opposite, that every proposition is transformed by becoming knowledge.

Moreover, just as propositions are undated, propositions that are true did not *become* true. A proposition known to be a truth became a *belief*, but it did not become a *truth*. A proposition known to be a falsehood became a *disbelief*, but it did not become a *falsehood*. Every proposition either *is* a truth or *is* a falsehood, but not every proposition either is a belief or is a disbelief. William James is sometimes interpreted as disputing these points.

Consensus among objective investigators researching the same hypothesis from different points of view is taken to be a mark of truth or probable truth. We might wish that in the fullness of time the community of investigators will come to share common beliefs in the central true hypotheses under current investigation; we might wish that consensus will be achieved—and also that it will be correct. But the idea that in the long run every true proposition will become a belief, much less a consensus belief, is absurd—most true propositions will never even be understood much less believed. The view, sometimes attributed to Peirce, that being true is coextensive with being believed to be true in the long run by the community of investigators must be mistaken. Perhaps it should be attributed to wishful thinking. Besides, in the fullness of time there might be no investigators.

Others have attributed to Peirce the even more irresponsible view that ‘being true’ means “being believed in the long run by the community of investigators”. This view implies that every attribution of truth to a proposition is a statement about the future attitudes or mental states of people who have not yet been born [18, 36].

17 Lying and Telling the Truth

This section is about deliberate statements. It excludes inadvertent remarks, misstatements, statements made under distracting conditions, and the like. The two expressions ‘telling a falsehood’ and ‘telling the truth’ can be misleading. The first does not mean “saying something false”, and the second does not mean “saying something true”. Telling a falsehood is lying, and that is not necessarily saying something false. And a person can say something false without lying. Likewise, telling the truth is stating what one believes to be true and beliefs are not necessarily true. And a person can say something true without telling the truth. Lying and telling the truth are forms of statement-making: they are human actions called *speech-acts*. Some philosophers have misunderstood the nature of the lie. A lie is a speech-act, not merely a sentence, or a proposition. A lie is a statement of a proposition that is not a belief of the speaker. Speakers who state their false beliefs are not lying. Likewise, speakers who state true propositions that they do not believe are lying—regardless of whether the non-belief is disbelief. Persons who state propositions on which they have no opinion are lying as much as those who state propositions they believe to be false.

Lies of ignorance are statements that are neither believed nor disbelieved by the speaker. *Lies of knowledge* are statements contrary to the speaker's beliefs or disbeliefs. Lies of ignorance can be just as harmful as lies of knowledge and just as effective in promoting the aims of the liar. Moreover, because of confusion about the nature of lying, it is often easier to get away with lies of ignorance. As a matter of terminology, it might be better to call lies of ignorance *lies of unbelief* and to call lies of knowledge *lies of belief or disbelief*.

Perhaps paradoxically, there is often no way to tell a lie of ignorance without indirectly telling a lie of knowledge. The reason is that any given statement that something is the case carries with it the indirect statement that the speaker believes the given statement. When speakers *know* that they do not believe their direct statements, their indirect statements are lies of knowledge.

If a person states that the house is locked up, there are at least two propositions that they have stated: a *primary* or direct statement about the house—that it is locked up—and a *secondary* or indirect statement about themselves—that they believe that the house is locked up. If they have no such belief, the primary statement is a lie of ignorance. If they know that they have no such belief, the secondary statement is a lie of knowledge.

18 Objectual Knowledge and Operational Knowledge

In order to have propositional knowledge, e.g. knowledge that two is the only even prime number, it is necessary to have objectual knowledge of several objects including but by no means limited to the following: the number two, the properties of being even and of being prime, the system of numbers, and the *exemplification* relation, i.e. the logical relation of a number to a numerical property that belongs to it. The number two exemplifies the property of being even. The word 'object' is being used in a very broad sense. Objectual knowledge is knowing of objects, including properties, relations, concepts, and anything else. We have objectual knowledge of everything we are acquainted with directly or indirectly as well as everything we know *of* by inference or reflection, such as the concept of truth. Some philosophers have implied that it is possible to know of an object by describing it, but this has always seemed to us to have the facts reversed. We do not see how we could describe something and thereby acquire knowledge of it unless we did not previously know of it. At this point it is not necessary to be more precise about the limits of objectual knowledge.

In the process of acquiring propositional knowledge that two is the only even prime number, Corcoran used various know-how, skills, or abilities that he had acquired previously. The skills and abilities that he has are what we have been calling his operational knowledge or knowledge-how-to. In the case being discussed, several items of operational knowledge might have come into play: the ability to factor a number, the ability to survey the progression of numbers starting with one, the ability to deduce consequences of propositions, the ability to understand propositions, the ability to make judgments.

It might well be that some or all of his objectual knowledge derives from exercising operational knowledge, for example, that he learns of numbers by counting, or conversely that some or all of his operational knowledge somehow derives from or depends on reflecting on his objectual knowledge, for example, that he learns how to count through reflecting on numbers. It is beyond the scope of this work to reflect on such issues. However, to forestall possible confusion, it is important to notice that he has objectual knowledge of propositions and of skills, which of course are objects in the broad sense. As is evident from the above, we *know of* propositions concerning which we do not have propositional knowledge, for example the Goldbach Hypothesis. In addition, we know of skills that we have not acquired. For example, Corcoran does not have operational knowledge of playing the violin or touch-typing but he has objectual knowledge of those skills.

My first act of free will shall be to believe in free will.—William James, 1870.

19 Choosing Beliefs and Disbeliefs

A person presented with several beliefs in the propositional sense might want to choose one for any number of purposes, e.g. to discuss first or to think about first. There is no problem here: it is always possible to choose randomly. However, when beliefs in the attitudinal sense are considered, the situation has changed.

Forming or shedding a belief and holding or lacking a belief are not acts like turning a switch on or turning it off. Forming and shedding beliefs are more like waking up and falling asleep. Holding and lacking a belief are more like staying awake and staying asleep. Turning a switch is voluntary and arbitrary in a way that belief formation is not. I can never form a belief in a given proposition by deciding to believe it and then throwing some sort of switch. This point has been denied by William James, who believed in "the will's primacy, even in choosing what to believe"—to use Rebecca Goldstein's formulation. According to Goldstein [30, 25], James wrote in his journal in 1870, "My first act of free will shall be to believe in free will".

Once a proposition has been understood and the evidence gathered and marshaled, the judgment is almost automatic—if it happens. Sometimes no result is reached. When a result is reached, it could be contrary to the desires of the believer. However, it is important to realize that although judgment is voluntary, it is not arbitrary at least when knowledge is achieved. An attitude or state of mind that was caused and not autonomously achieved through judgment is not even belief in the sense of this investigation.

Likewise, once there is awareness of a serious deficiency in the process that led to a belief, say that the source of testimony has been discredited or the instruments used found flawed, often the belief is lost or at least its certitude is diminished regardless of how attached to it the believer is. Charles Sanders Peirce agreed here, but he might have gone a little too far with it when he wrote [37, 119–120]: "Now, there are some people [...] who, when they see that any belief of theirs is determined by any circumstance extraneous to the facts, will from that moment [...], experience real doubt of it, so that it ceases to be a belief".

If a coin is flipped and the outcome covered, as long as the evidence is unavailable there is no way for a person to form the belief that it is heads, say, no matter how much they might want it to be heads. More generally, there is no way to form a belief in a proposition that is now a hypothesis simply by deciding to believe it. The statement that some people believe what they want to believe is a misleading half-truth. To the extent that it would be fair to say it, it would be just as fair to say that some people believe what they want to disbelieve. People sometimes confuse hopes and fears with evidence. The expressions “too awful to be true” and “too good to be true” are common enough.

Forming a belief is not like deciding to purchase a given item and then putting it in the shopping cart. Some philosophers have disagreed, holding that a credence such as a religious belief is an exception, that a belief can be freely chosen when there is no hope of finding evidence. However, some religious thinkers have disputed that conclusion, saying that humans are powerless to construct such beliefs, that humans must await divine intervention, and that “the gift of faith” cannot be chosen but is freely bestowed.

Persons who find that two or more of their beliefs are inconsistent have all of those beliefs undermined. They are not free to decide which they prefer to keep and which to drop. They cannot arbitrarily decide, contrary to what some logicians seem to say. There is no switch to pull that reinstates some as beliefs while rendering others as disbeliefs. The issue here is not whether it is morally acceptable to adopt arbitrarily a belief and it is not whether a person who does this risks losing intellectual integrity. The issue concerns how beliefs are formed.

Some logicians recommend that we adjust our degrees of certitude so that they are proportional to our degrees of certainty. They say that the more certainty we have of a belief the more certitude we should feel and, accordingly, the less certainty we have the less certitude we should feel. However, we have no more control over the intensity of our feelings of certitude than we have over our judgments. Moreover, no way of measuring relative certitude or relative certainty has been devised. Those logicians might as well have recommended that we adjust our level of fear in proportion to the level of danger or that we adjust our level of happiness in proportion to our level of well-being.

The above should not be taken to deny that we often choose *to try* to understand one proposition while choosing *to try* not to understand another. Selective attention and willful ignorance are common enough. Nor should it be taken to deny that we can choose to seek evidence or arguments for one proposition while choosing to ignore evidence and arguments for another. The reality of partisanship, rationalization, and self-deception must be acknowledged, and it is just to hold people responsible in such cases. There is something unsavory about trying to choose a belief; it seems to violate intellectual integrity. Choosing to adopt or shed a belief seems to be a kind of self-deception, a kind of lie of ignorance.

The selfless autonomy of the objective judgment is counterpoised against the selfish arbitrariness of the subjective decision. Selfishness, laziness, impatience, lack of discipline, and other character defects that interfere with the successful completion of the cognitive process are open to blame.

However, with exceptions such as those just noted, it is absurd to blame people for holding, lacking, adopting, or shedding belief in a given proposition regardless of how deleterious or beneficial we might think it would be. It is even more absurd to try to require someone to adopt or shed a belief in a given proposition. This would be like trying to require people to enjoy something they find repulsive or requiring them to be repulsed by something they enjoy. Attempting to coerce belief or disbelief compounds absurdity with injustice. Compare Peirce [39, p. 188].

20 Background

It might be supposed that discussion of knowledge versus opinion took center stage in American philosophy with the 1877 publication by Charles Sanders Peirce of his seminal paper "The Fixation of Belief", which has justly become somewhat a *locus classicus* for the issue. It is evident to scholars that the above discussion is heavily indebted to the Peirce paper both in spirit and in particular views. The above resonates with several of Peirce's points including his "method of science", his supposition that "there is some one thing to which a proposition should conform", and his view that what is believed is in no way determined or changed by our thinking—to mention only three. In terms of the present discussion, roughly speaking, "The Fixation of Belief" presents four methods that can be used to increase *certitude* in a proposition already believed. One of those methods, the method of science, increases *certainty*.

Nevertheless, Peirce does not explicitly raise the issue of distinguishing "knowledge" from "opinion". Perhaps surprisingly, he does not even use the word 'knowledge' or a synonym, and he does not make a distinction analogous to certitude/certainty.

However, the above discussion does not relate to certain "pragmatic themes" in later Peirce writings, in which focus on the *nature* of truth as "conforming to facts" gives way to focus on the *criterion* of truth as leading us to fulfilling our aims. *A fortiori*, the above does not relate to other classic American philosophers such as William James and John Dewey who worked in paradigms that might even be incommensurable with those currently flourishing in the United States. James and Dewey would dispute that "truth is a precondition to knowledge" and that "the proposition known is not changed by becoming known".

The present discussion of the knowledge/opinion distinction addresses the more analytic side of post-World War II American philosophy. One of the most accessible of relevant texts is the 1978 monograph *The Web of Belief* by Quine and Ullian [44]. An excellent contemporary American treatment of propositional knowledge, which explicitly treats the knowledge/opinion distinction, is the 1991 treatise *Knowledge and Evidence* by Paul Moser [35].

Acknowledgements We would like to thank those who have made important contributions to the dialogue leading up to this work. Robert Barnes, Leigh Duffy, John Glanville (1921–2010), Carol Gould, Forest Hansen, Peter Hare (1935–2008), Amanda Hicks, Daniel Merrill, Joaquin Miller, Hassan Masoud, Paul Moser, Mary Mulhern, Carlo Penco, Anthony Preus, José Miguel Sagüillo, Michael Scanlan, John Shook, Andrew Spear, Yevgeny Yarmolinets, Joseph Zeccardi, and John Zeis all deserve credit. Many thanks also to all who participated in the crafting of this article, especially David Brewer, Joseph Corcoran, David Hitchcock, Leonard Jacuzzo, Timothy Madigan, James McNabb, Kristo Miettinen, Frango Nabrasa, Sriram Nambiar, Daniel Novotny, Charles Pailthorp, Paul Penner, David Plache, and Roberto Torretti. David Hitchcock and Daniel Novotny deserve special thanks for alerting us to errors and obscurities in the penultimate draft; they are almost co-authors even though they disagree with some of our conclusions.

Many of the scholars mentioned disagree with some, many, or most of the conclusions of this work. We should also acknowledge many previous writers to whom we are indebted for most of the points made in this work. Our main contribution was to have collected their thoughts into a coherent work accessible

to a wide audience. This paper is a substantial reworking of "An Essay on Knowledge and Belief" [15]. The word 'essay' in its title was carefully chosen: the essay has no footnotes or references and it is much shorter than this investigation. The essay was inspired by responses to early drafts of the encyclopedia article "Knowledge and Belief" [17].[27]

References

1. Anagnostopoulous, G. (ed.): Companion to Aristotle. Wiley-Blackwell, New York (2013)
2. Aristotle: Metaphysics
3. Aristotle: Prior Analytics
4. Arnauld, A., Nicole, P.: Logic or the Art of Thinking: Port Royal Logic. Cambridge UP, Cambridge (1662/1996). Trans. J. Buroker.
5. Audi, R. (ed.): Cambridge Dictionary of Philosophy, 3rd edn. Cambridge UP, Cambridge (2015)
6. Austin, J.L.: Philosophical Papers. Oxford UP, Oxford (1961). Ed. J. Urmson and G. Warnock
7. Austin, J.L.: Truth. Proc. Aristotelian Soc. Suppl. **24** (1950/1961). Reprinted Austin 1961 [6]
8. Church, A.: Introduction to Mathematical Logic. Princeton UP, Princeton (1956)
9. Church, A.: Propositions and sentences. In: Bochenski, I., Church, A., Goodman, N. (eds.) The Problem of Universals. Notre Dame UP, Notre Dame (1956)
10. Corcoran, J.: Discourse grammars and the structure of mathematical reasoning I: mathematical reasoning and the stratification of language. J. Struct. Learn. **3**(1), 55–74 (1971)
11. Corcoran, J.: Argumentations and logic. Argumentation **3**, 17–43 (1989)
12. Corcoran, J.: Semantic arithmetic: a preface. Agora **14**, 149–156 (1995)
13. Corcoran, J.: Critical thinking and pedagogical license. Manuscrito **XXII**, 109–116 (1999)
14. Corcoran, J.: Counterexamples and proexamples. Bull. Symb. Log. **11**, 460 (2005)
15. Corcoran, J.: An essay on knowledge and belief. Int. J. Decis. Ethics **II.2**, 125–144 (2006)
16. Corcoran, J.: A priori/a posteriori. In: Lachs, J., Talisse, R. (eds.) Encyclopedia of American Philosophy. Routledge, New York (2007)
17. Corcoran, J.: Knowledge and belief. In: Lachs, J., Talisse, R. (eds.) Encyclopedia of American Philosophy. Routledge, New York (2007)
18. Corcoran, J.: Review of Nolt 2008. Math. Rev. (2008)
19. Corcoran, J.: Ambiguity: lexical and structural. Bull. Symb. Log. **15**, 235–236 (2009)
20. Corcoran, J.: Sentence, proposition, judgment, statement, and fact. In: Carnielli, W., et al. (eds.) The Many Sides of Logic, pp. 71–103. College Publications, London (2009)
21. Corcoran, J., Frank, W.: Cosmic justice hypotheses. Bull. Symb. Log. **20**, 247 (2014)
22. Corcoran, J., Hamid, I.S.: Objectivity-subjectivity distinctions. Bull. Symb. Log. **20**, 248 (2014)
23. Davenport, H.: Higher Arithmetic. New York, Harper (1952/1960)
24. De Morgan, A.: On the Syllogism and Other Logical Writings. Routledge and Kegan Paul, London (1966). Ed. P. Heath
25. Dessi, P.: Introduction. In: Whately, R.: Elements of Logic. Editrice CLLEB, Bologna (1826). Ed. P. Dessi (1988)
26. Frege, G.: Foundations of Arithmetic. Basil Blackwell, Oxford (1884). Trs. J.L. Austin
27. Frege, G.: The thought: a logical inquiry. Mind **65**, 289–311 (1956). Trans. by A. and M. Quinton
28. Frege, G.: The Frege Reader. Blackwell, Oxford (1997). Ed. M. Beaney
29. Giorgi, A.: The question of validity in qualitative research. J. Phenomenol. Psychol. **33**, 124–134 (2002)
30. Goldstein, R.N.: The pragmatist. New York Times Book Review, 17 December 2006
31. Hilpinen, R.: Knowing that one knows and the classical definition of knowledge. Synthese **21**, 109–132 (1970)
32. Lachs, J., Talisse, R. (eds.): Encyclopedia of American Philosophy. Routledge, New York (2007)
33. Merriam-Webster Collegiate Dictionary, 11th edn. Merriam-Webster, Springfield (2008)
34. Mill, J.S.: A System of Logic, 10th edn. 2 vols. Longmans Green, London (1843/1879)

[27] Corcoran's encyclopedia article 2007 "Knowledge and Belief" [17] was written earlier than the 2006 "An Essay on Knowledge and Belief" [15], but assigned a later publication date [32].

35. Moser, P.: Knowledge and Evidence. Cambridge UP, Cambridge (1991)
36. Nolt, J.: Truth as an epistemic ideal. J. Philos. Log. **37**, 203–237 (2008)
37. Peirce, C.S.: The fixation of belief. Pagination refers to the reprinting in Peirce, C.S.: The Essential Peirce: Selected Philosophical Writings (1867–1893), vol. I. Indiana UP, Bloomington (1992). Eds. N. Houser and C. Kloesel
38. Peirce, C.S.: On the algebra of logic. Am. J. Math. **3**, 15–57 (1880)
39. Peirce, C.S.: On the algebra of logic: a contribution to the philosophy of notation. Am. J. Math. **7**, 180–202 (1885)
40. Peirce, C.S.: The Essential Peirce: Selected Philosophical Writings (1867–1893), vol. I. Indiana UP, Bloomington (1992). Eds. N. Houser and C. Kloesel
41. Peirce, C.S.: The Essential Peirce: Selected Philosophical Writings (1893–1913), vol. II. Indiana UP, Bloomington (1998). Eds. N. Houser and C. Kloesel
42. Peters, F.E.: Greek Philosophical Terms. New York UP, New York (1967)
43. Preus, A.: Historical Dictionary of Ancient Greek Philosophy. Rowman & Littlefield, Lanham (2007)
44. Quine, W.V., Ullian, J.S.: The Web of Belief. Harvard UP, Cambridge (1978)
45. Russell, B.: Intuition. Cambridge Dictionary of Philosophy (2015)
46. Tarski, A.: Logic, Semantics, Metamathematics. Hackett, Indianapolis (1983)
47. Tarski, A.: What are logical notions? Hist. Philos. Logic **7**, 143–154 (1986)
48. Whately, R.: Elements of Logic. Editrice CLLEB, Bologna (1826). Ed. P. Dessi, 1988

J. Corcoran (✉)
Philosophy, University at Buffalo, Buffalo, NY 14260-4150, USA
e-mail: corcoran@buffalo.edu

I.S. Hamid
Philosophy, Colorado State University, Fort Collins, CO 80523-1781, USA
e-mail: ishamid@colostate.edu

The Algebra of Opposition (and Universal Logic Interpretations)

Răzvan Diaconescu

Abstract We give clear algebraic sense to logical opposition by providing negation-free lattice-theoretic definitions for the concepts around the notorious square of opposition. These include contradiction, contrariety, subcontrariety, the square and the hexagon of opposition, etc. This constitutes a platform for an analysis of the mathematical properties of logical opposition. We also discuss several examples of squares of opposition arising from universal logics studies, including Boolean as well as less conventional non-Boolean squares. The latter kind arise from a very general study of negation and consistency within the context of many-valued consequence relations. This work is a tribute to Jean-Yves Béziau, friend and colleague, on the occasion of his 50th birthday.

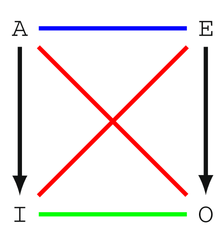
Keywords Square of opposition · Universal logic · Many-valued logic · Lattice theory · Consequence · Consistency

Mathematics Subject Classification (2000) 03B22 · 03B50 · 06B05

1 The Square of Opposition

The meta-logical square of opposition has its origins in the work of the ancient Greek philosopher Aristotle [1] and it has been used since antiquity until now in philosophical logic as tool and doctrine. In recent years, there has been a new interest in the square of opposition reflected by various new interpretations and a diversity of extensions. These activities have recently found a home in a series of dedicated World Congresses very much due to the effort of Jean-Yves Béziau, a glimpse of them can be seen in [5].

The modern look of the square of opposition is like below, with the red lines representing logical *contradiction*, the black ones *subalternation* (or implication), the blue one *contrariety*, and the green one *subcontrariety*.



A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_6

1.1 The Consequence Square of Opposition

The original intention behind the square of opposition is to describe falsifiability. This is a logical issue which is not really dependent upon a particular concrete logic, so it can be approached at a very general level by using Tarski's abstract axiomatization of consequence relations [19]. Here we use its entailment style formulation [18] rather than its original formulation in terms of closure operators, these being equivalent formulations of the same concept. Note that the theory of abstract consequence, either in its closure operator or entailment relation form, constitutes one of the important origins of the *universal* trend in logic [3], that in the recent decades has gained a prominent status, again very much due to the dedication of Jean-Yves Béziau. Universal logic represents a general abstract study of logic phenomena that is completely independent of any particular logical systems; this is a top-down non-substantialist conceptual process very different for the bottom-up substantialist thinking underlying the conventional approach to logic.

Definition 1.1 (Entailment) A pair $(S, \vdash)$ is an *entailment relation* when S is a set and $\vdash \subseteq 2^S \times S$ such that for all $E, \Gamma \subseteq S$ and $\rho \in S$:

- (*Reflexivity*) $E \vdash \rho$ if $\rho \in E$; and
- (*Cut*) $E \vdash \rho$ if $\Gamma \vdash \rho$ and $E \vdash \gamma$ for each $\gamma \in \Gamma$.

The set S in Definition 1.1 abstracts the sentences, or closed formulæ, of actual logical systems. Actual entailment relations can be obtained either model or proof theoretically. The model-theoretic entailments can be determined in a very general way by employing the following abstract concept of satisfaction.

Definition 1.2 ([14]) A *room* is a relation $\models \subseteq \mathcal{M} \times S$ where $\mathcal{M}$ is a class (of 'models') and S is a set (of 'sentences').

The following is folklore in some very abstract approaches to model theory, such as institution theory [14]; its proof is completely straightforward and can be checked easily by the reader.

Proposition 1.3 *Every room* $\models \subseteq \mathcal{M} \times S$ *determines a canonical entailment relation* $(S, \models)$, *called* semantic entailment, *by*

$$T \models \rho \quad \textit{if and only if or each } M \in \mathcal{M}, M \models \rho \quad \textit{whenever} \quad M \models \tau \quad \textit{for each } \tau \in T.$$

Often in logic there is no distinction between triviality and inconsistency; this is due to the fact that in classical logics, and not only, these two concepts coincide. However, the distinction between them, which is quite apparent when employing an abstract universal logic perspective, constitutes in fact the main idea underlying paraconsistent logics [17].

Definition 1.4 (Triviality) In any entailment relation $(S, \vdash)$ any $T \subseteq S$ is called a *theory*. A theory T is *non-trivial* when there exists a ρ such that $T \not\vdash \rho$, otherwise is *trivial*.

Definition 1.5 (Consistency) Given an entailment relation $(S, \vdash)$ and a function $\neg : S \to S$, a theory $T \subseteq S$ is $\neg$*-consistent* when there does *not* exist ρ such that $T \vdash \rho$ and $T \vdash \neg\rho$; otherwise T is called $\neg$*-inconsistent*.

The function $\neg$ is *consistent* (with respect to $\vdash$) when for each $\rho \in S$, the set $\{\rho, \neg\rho\}$ is a trivial theory in $(S, \vdash)$; otherwise $\neg$ is called *paraconsistent*.

The following, which constitutes a general clarification of the relationship between non-triviality and consistency, gets a straightforward proof from the definitions above.

Proposition 1.6 *In any entailment* $(S, \vdash)$, *the following are equivalent properties for a function* $\neg : S \to S$:

1. $\neg$ *is consistent*; *and*
2. *For each theory* T, T *is non-trivial if and only if* T *is* $\neg$*-consistent.*

The negations of classical logics, as well as of most non-classical logics are consistent. Non-classical logics with consistent negation in the sense of Definition 1.5 include both local and modal semantic consequences, intuitionistic logics, etc., but exclude paraconsistent logics. The following represents a general way to obtain consistent semantic entailments that is applicable to all classical logics, but not only.

Definition 1.7 (Semantic negation) A room $\models \subseteq \mathcal{M} \times S$ *admits negation* when there exists a function $\neg : S \to S$ such that for each $M \in \mathcal{M}$ and each $\rho \in S$

$$M \models \neg\rho \quad \text{if and only if } M \not\models \rho.$$

Fact 1.8 *For any room with a negation* $\neg$, *this is consistent with respect to the induced semantic entailment.*

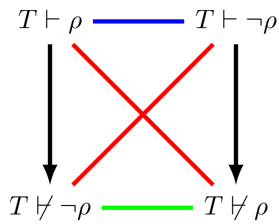
Corollary 1.9 *For any room with negation* $\neg$, *in the corresponding semantic entailment a theory* T *is non-trivial if and only if it is* $\neg$*-consistent.*

The consistency of negation for the global consequence relations in modal logics and for the consequence relations of intuitionistic logics holds but falls outside the scope of Fact 1.8. The same can be said about the semantic consequence in the weakened version of classical propositional logic proposed in [2] where S may be taken as the set of ordinary propositional logic sentences for a given set P of propositional variables and the models are the valuations $M : S \to \{0, 1\}$ that respect that usual truth table semantics of all the Boolean connectives except negation, for which they are required to respect one half of the usual condition, namely $M(\rho) = 1$ implies $M(\neg\rho) = 0$.

The following straightforward characterization of $\neg$-consistency is particularly relevant for the consequence-theoretic interpretations of the square of opposition. Note that this holds independently of the consistency property for $\neg$.

Corollary 1.10 *For any entailment relation* $(S, \vdash)$ *and for any function* $\neg : S \to S$, *a theory* $T \subseteq S$ *is* $\neg$*-consistent if and only if for each* $\rho \in S$, $T \vdash \neg\rho$ *implies* $T \not\vdash \rho$.

In an entailment relation $(S, \vdash)$, for any $\neg : S \to S$, any $\neg$-consistent theory $T \subseteq S$ and any $\rho \in S$ determine the following binary square of opposition in the sense that each node represents one of the values 0 (for false) or 1 (for true).



(1.1)

The essential relations defining this as an opposition square are those of

- (*Contradiction*)
  - $(T \vdash \rho) \wedge (T \nvdash \rho) = 0$; $(T \vdash \rho) \vee (T \nvdash \rho) = 1$ and
  - $(T \vdash \neg\rho) \wedge (T \nvdash \rho) = 0$; $(T \vdash \neg\rho) \vee (T \nvdash \rho) = 1$;

  and
- (*Subalternation*)
  - $(T \vdash \rho) \leq (T \nvdash \neg\rho)$ and
  - $(T \vdash \neg\rho) \leq (T \nvdash \rho)$.

The relations of

- (*Contrariety*) $(T \vdash \rho) \wedge (T \vdash \neg\rho) = 0$ and
- (*Subcontrariety*) $(T \nvdash \neg\rho) \vee (T \nvdash \rho) = 1$

are determined from the previous two relations just by calculations in the binary Boolean algebra.

In order to lift this binary square to a non-binary (but yet Boolean) square, we may eliminate the parameter ρ:

Corollary 1.11 *Given an entailment system* $(S, \vdash)$ *and* $\neg : S \to S$, *for each theory* $T \subseteq S$ *let us denote as follows*:

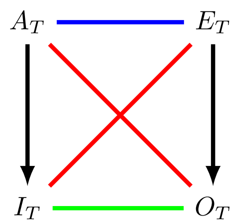
- $A_T = \{\rho \in S \mid T \vdash \rho\}$;
- $E_T = \{\rho \in S \mid T \vdash \neg\rho\}$;
- $I_T = \{\rho \in S \mid T \nvdash \neg\rho\}$; *and*
- $O_T = \{\rho \in S \mid T \nvdash \rho\}$.

We then have the following relations:

1. $A_T \cap O_T = \emptyset$, $A_T \cup O_T = S$, $I_T \cap E_T = \emptyset$, $I_T \cup E_T = S$;
2. *If* T *is* $\neg$*-consistent then* $A_T \subseteq I_T$, $E_T \subseteq O_T$, $A_T \cap E_T = \emptyset$ *and* $I_T \cup O_T = S$.

Proof Relation 1 is trivial. The first part of relation 2 follows from Corollary 1.10, the other ones are Boolean consequences of all previous ones. □

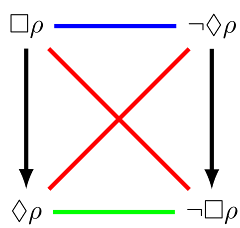
The conclusion of Corollary 1.11 may be graphically represented by the following square of opposition whose nodes are elements of the Boolean algebra S^S.



(1.2)

1.2 The Modal Square of Opposition

Another notorious interpretation of the square of opposition, a modern one, comes from modal logic (as usually $\Box$ denotes necessity and $\Diamond$ possibility):



(1.3)

In what follows, we present the most general structure that provides a precise mathematical sense to the square (1.3). For this we employ the general method of 'modalization' of logics common to works such as [10, 13, 15], etc.; this is also a universal logic device.

Definition 1.12 Given a room $\models \subseteq \mathcal{M} \times S$, we define:

- $M(S)$ the set of terms formed from S as constants, $\neg$, $\Box$, $\Diamond$ as unary functions, and $\wedge$ as binary function;
- $K(\mathcal{M})$ is the class of pairs (W, M) where
  - $W = (|W|, W_\lambda)$ with $|W|$ being a set and $W_\lambda \subseteq |W| \times |W|$ a binary relation on $|W|$, and
  - $M : |W| \to \mathcal{M}$ being a function,

  and
- $(W, M) \models^i \rho$ for each $(W, M) \in K(\mathcal{M})$, $\rho \in M(S)$ and $i \in |W|$ defined by induction on the structure of ρ as follows:
  - $((W, M) \models^i \rho) = (M_i \models \rho)$ when $\rho \in S$;
  - $((W, M) \models^i \rho_1 \wedge \rho_2) = ((W, M) \models^i \rho_1) \wedge ((W, M) \models^i \rho_2)$;
  - $((W, M) \models^i \neg\rho) = \neg((W, M) \models^i \rho)$;
  - $((W, M) \models^i \Box\rho) = \bigwedge_{(i,j)\in W_\lambda}((W, M) \models^j \rho)$; and
  - $((W, M) \models^i \Diamond\rho) = \bigvee_{(i,j)\in W_\lambda}((W, M) \models^j \rho)$.

The elements of $M(S)$ may be referred to as 'the (propositional) modal sentences over S' and the elements of $K(\mathcal{M})$ as the 'Kripke structures over $\mathcal{M}$'.

Proposition 1.13 *For any Kripke structure* (W, M) *and any sentence* ρ, *let us denote as follows*:

- $A = \{i \in |W| \mid (W, M) \models^i \Box\rho\}$;
- $E = \{i \in |W| \mid (W, M) \models^i \neg\Diamond\rho\}$;
- $I = \{i \in |W| \mid (W, M) \models^i \Diamond\rho\}$; *and*
- $O = \{i \in |W| \mid (W, M) \models^i \neg\Box\rho\}$.

We then have the following relations:

1. $A \cap O = \emptyset$, $A \cup O = |W|$, $I \cap E = \emptyset$, $I \cup E = |W|$;
2. *If the accessibility relation* W_λ *is reflexive then* $A \subseteq I$, $E \subseteq O$, $A \cap E = \emptyset$, *and* $I \cup O = |W|$.

Proof Relation 1 is trivial. The first two facts of relation 2 follow from the fact that $\Box\rho \Rightarrow \Diamond\rho$ holds in the model, this being a consequence of $\Box\rho \Rightarrow \rho$ and of $\rho \Rightarrow \Diamond\rho$, that hold because of the reflexivity of the accessibility relation W_λ. The other relations follow by straightforward Boolean calculations. □

The relations of Proposition 1.13 can be also represented by a square of opposition, but with the nodes in the Boolean algebra $2^{|W|}$.

2 The Square of Opposition as a Lattice Theoretic Structure

Both examples of consequence and modal theoretic squares of opposition discussed above, in fact, represent sets of equations in Boolean algebras that involve four constants corresponding to the nodes of the squares of opposition. These equations can be given in any bounded lattice, which leads us to an abstract definition of the concept of square of opposition as a lattice-theoretic structure. Note that *this is done without assuming a negation*, in our view the concept of negation not being an inherent part of the concept of logical opposition, at least not from a general perspective.

Definition 2.1 In any bounded lattice $(L, \leq, \wedge, \vee, 0, 1)$, we define the following relations:

- (*Contrariety*) $B_L = \{(x, y) \mid x \wedge y = 0\}$;
- (*Subcontrariety*) $G_L = \{(x, y) \mid x \vee y = 1\}$; and
- (*Contradiction*) $R_L = B_L \cap G_L$.

Note that in the literature on the square of opposition (e.g. [4]) the relation $x \leq y$ is called 'y is *subalterned* by x'.

According to Definition 2.1, we can see that

$$\boxed{\text{Contradiction} = \text{Contrariety} + \text{Subcontrariety.}}$$

Definition 2.2 (Square of opposition) In any bounded lattice L, a *square of opposition* is a 4-tuple (a, e, i, o) such that

- $(a, o), (i, e) \in R_L$; and
- $a \leq i, e \leq o$.

The following simple result shows that in a square of opposition the contrariety and the subcontrariety relations are determined by the contradiction and subalternation relations, hence an economical definition of the square of opposition does not need to include the former relations. However, the common presentations of the square of opposition do include all these relations, which I think may be due to not viewing the square of opposition as an algebraic structure, the algebraic perspective making redundancies rather transparent.

Proposition 2.3 *In any square of opposition* (a, e, i, o), *we have the following*:

1. $(a, e) \in B_L$; *and*
2. $(i, o) \in G_L$.

Proof The conclusion follows by simple calculations:

1. $a \wedge e \leq a \wedge o = 0$, hence $a \wedge e = 0$.
2. $i \vee o \geq i \vee e = 1$, hence $i \vee o = 1$. □

The following result shows that, assuming the distributivity of the lattice (which in lattice theory is a heavy condition with many implications), given two Contradictions (red) a corresponding square of opposition may be determined only by a Contrariety relation (blue) or by a Subcontrariety (green) one. In fact, the original Aristotle's formulation of the square of opposition was given in the form of two Contradictions and a Contrariety.

Proposition 2.4 *In any distributive bounded lattice* L, *given a* 4-*tuple* (a, e, i, o) *such that* $(a, o), (i, e) \in R_L$, *the following are equivalent*:

1. (a, e, i, o) *is a square of opposition*;
2. $(a, e) \in B_L$; *and*
3. $(i, o) \in G_L$.

Proof $\boxed{1. \Rightarrow 2.}$ From Proposition 2.3.

$\boxed{2. \Rightarrow 3.}$ By the following calculation:

$i \vee o = i \vee o \vee 0 = (i \vee o) \vee (a \wedge e) = (i \vee o \vee a) \wedge (i \vee o \vee e) = (i \vee 1) \wedge (o \vee 1) = 1 \wedge 1 = 1$.

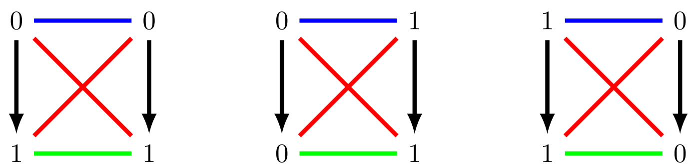
$\boxed{3. \Rightarrow 1.}$: By the following calculations:

$a = a \wedge 1 = a \wedge (i \vee o) = (a \wedge i) \vee (a \wedge o) = (a \wedge i) \vee 0 = a \wedge i$;
$e = e \wedge 1 = e \wedge (i \vee o) = (e \wedge i) \vee (e \wedge o) = 0 \vee (e \wedge o) = e \wedge o$. □

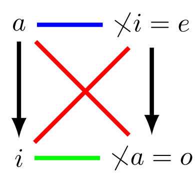
The squares of opposition may be mirror-reflected:

Fact 2.5 *If* (a, e, i, o) *is a square of opposition then* (e, a, o, i) *is also a square of opposition.*

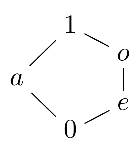
In each bounded lattice, there are three trivial squares of opposition:



Recall that a *complementation* in a bounded lattice L is a function $\nprec : L \to L$ such that for each element x, we have $x \wedge \nprec x = 0$, $x \vee \nprec x = 1$. In each lattice with an complementation $\nprec$ that is order-reversing (i.e. $\nprec x \leq \nprec y$ if $y \leq x$), each relation $a \leq i$ (or, in the mirror, $e \leq o$) determines a square of opposition as follows:



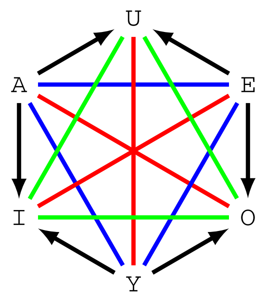
Note that, in general, an order reversing complementation $\nprec$ does not necessarily have to fulfill $\nprec\nprec x = x$, as shown by the simple example of the pentagon lattice N5



with the complementation $\nprec$ defined by $\nprec a = o$. Then $\nprec\nprec e \neq e$. However, if the lattice is a Boolean algebra, then there is a unique complementation which moreover enjoys $\nprec\nprec x = x$, so in this case there is a unique square of opposition determined by the relation $a \leq i$. For example, the squares of opposition (1.1) associated to an entailment relation are of the Boolean kind, being completely determined by the fact that $T \vdash \rho$ implies $T \nvdash \neg\rho$, which is equivalent to the fact that T is $\neg$-consistent. Also the modal squares of opposition (1.3) are Boolean, being completely determined by the implication $\Box\rho \Rightarrow \Diamond\rho$.

2.1 *The Lattice Theoretic Hexagon of Opposition*

There are several extensions of the square of opposition, among them the hexagon discovered independently by Sesmat and Blanché [6] represents a response to several logical issues that do not get a proper answer within Aristotle's square. The hexagon of opposition consists of three squares of opposition as in the figure below.



(2.1)

Definition 2.6 (Hexagon of opposition) In any bounded lattice, a *hexagon of opposition* is a 6-tuple (a, e, i, o, u, y) such that (a, e, i, o), (a, y, u, o) and (e, y, u, i) are squares of opposition.

A common way to extend a square to a hexagon of opposition (see [4]) is captured in this lattice theoretic approach by Proposition 2.7 below. Note, however, that this requires (again) the distributivity property for the lattice.

Proposition 2.7 *In any distributive bounded lattice, if* (a, e, i, o) *is a square of opposition then* $(a, e, i, o, a \vee e, i \wedge o)$ *is a hexagon of opposition.*

Proof We have only to check that $(a, i \wedge o, a \vee e, o)$ and $(e, i \wedge o, a \vee e, i)$ are squares of opposition.

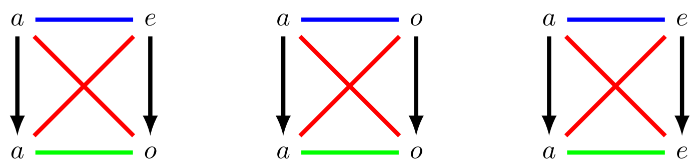
$\boxed{(a, i \wedge o, a \vee e, o)}$ Obviously, $a \leq a \vee e$ and $i \wedge o \leq o$. That $(a, o) \in R_L$ holds by hypothesis. That $(i \wedge o, a \vee e) \in R_L$ holds by the following calculations:

$$(i \wedge o) \wedge (a \vee e) = (i \wedge o \wedge a) \vee (i \wedge o \wedge e) = (i \wedge 0) \vee (0 \wedge o) = 0 \vee 0 = 0;$$
$$(i \wedge o) \vee (a \vee e) = (i \vee a \vee e) \wedge (o \vee a \vee e) = (1 \vee a) \wedge (1 \vee e) = 1 \wedge 1 = 1.$$

$\boxed{(e, i \wedge o, a \vee e, i)}$ Obviously, $e \leq a \vee e$ and $i \wedge o \leq i$. That $(e, i) \in R_L$ holds by hypothesis. That $(i \wedge o, a \vee e) \in R_L$ has been already proved. □

3 Non-Boolean Squares via Many-Valued Consequence

We may say that squares of opposition are truly interesting in non-Boolean lattices because of the greater independence between the components of the square (recall that in Boolean algebras squares of opposition may be determined just by one subalternation). For example, in the pentagon lattice N5 we have six non-trivial squares of opposition, the following three plus their mirror reflections:



In what follows, we will show how such non-Boolean squares of opposition may actually arise from universal logic, not necessarily as mere dry lattice theoretic structures. However, in order to achieve such examples, we have to go beyond the realm of binary truth.

3.1 Many-Valued Consequence

We may refine the Boolean square (1.1) to non-Boolean ones by considering a many-valued consequence rather than a binary one. This means that $T \vdash \rho$ is allowed to take other values than just true or false. Such many-valued approach to consequence, albeit marginal for the tradition of many-valued or fuzzy logic (see [11] for a discussion about this), has a history that can be traced back to the superb work of Pavelka [16]. Moreover, many-valued consequence arises naturally and has important applications in approximate reasoning in artificial intelligence, medicine, etc., formal methods based upon temporal logic. In the recent work [9], the author discusses several such examples from a logic perspective.

The original approach to many-valued consequence of [16] is on the side of closure operators in the style of Tarski [19] rather than on the entailment theoretic side. Unlike in the binary case, as several works in the literature show (e.g. [9]), in a many-valued context the equivalence between these two is highly problematic.

Definition 3.1 [16] Given a partial order $(P, \leq)$ a *closure operator* is a function $C : P \to P$ such that

- (*Reflexivity*) $x \leq Cx$ for each $x \in P$;
- (*Monotonicity*) $Cx \leq Cy$ for each $x \leq y \in P$; and
- (*Idempotency*) $CCx = Cx$ for each $x \in P$.

Given a lattice L and a set S, an L*-consequence operator* on S is just a closure operator on the power lattice L^S.

The following is a many-valued generalization of Definition 1.1. It should be noted that Definition 3.2 below introduces a mathematically weaker concept than the many-valued approach to entailment of [7–9, 12] often referred to as 'graded consequence'.

Definition 3.2 (Many-valued entailment) Given a bounded lattice L, a pair $(S, \vdash)$ is a *weak* L*-entailment relation* when S is a set and $\vdash : 2^S \times S \to L$ such that for all $E, \Gamma \subseteq S$ and $\rho \in S$:

- (*Reflexivity*) $(E \vdash \rho) = 1$ if $\rho \in E$; and
- (*Cut*) $(\Gamma \vdash \rho) \leq (E \vdash \rho)$ if $(E \vdash \gamma) = 1$ for each $\gamma \in \Gamma$.

The following gives a natural interpretation of L-consequences as weak L-entailments, which generalizes the corresponding interpretation from the binary case.

Notation 3.3 Given any set S and any bounded lattice L, let use the following notation: for any $T \subseteq S$ let $\tilde{T} : S \to L$ be defined by $\tilde{T}(\rho) = 1$ if $\rho \in T$ and $\tilde{T}(\rho) = 0$ otherwise.

Proposition 3.4 *For any bounded lattice L, any L-consequence C on a set S determines a weak L-entailment $(S, \vdash)$ defined for each $T \subseteq S$ and each $\rho \in S$ by*

$$(T \vdash \rho) = (C\tilde{T})\rho. \tag{3.1}$$

Proof *Reflexivity* of $\vdash$ is easy. For each $\rho \in E \subseteq S$, we have that $\tilde{E}(\rho) = 1$, hence by the *Reflexivity* of C we get that $(C\tilde{E})\rho = 1$ hence $E \vdash \rho$.

For establishing *Cut* for $\vdash$, let us assume $E, \Gamma \subseteq S$ such that $(E \vdash \gamma) = 1$ for each $\gamma \in \Gamma$. This means that $\tilde{\Gamma} \leq C\tilde{E}$. By the *Monotonicity* of C, it follows that $C\tilde{\Gamma} \leq CC\tilde{E}$. By the *Idempotency* of C, it further follows that $C\tilde{\Gamma} \leq C\tilde{E}$, which is the conclusion of *Cut*. □

The following gives a natural interpretation of weak L-entailments as L-consequences, which generalizes the corresponding interpretation from the binary case.

Proposition 3.5 *For any complete lattice L, any weak L-entailment $(S, \vdash)$ determines an L-consequence on S defined for each $X : S \to L$ by*

$$(CX)\rho = \bigwedge \{T \vdash \rho \mid X(\gamma) \leq T \vdash \gamma \textit{ for each } \gamma \in S\}. \tag{3.2}$$

Proof The *Reflexivity* of C is immediate from (3.2). The *Monotonicity* of C follows by (3.2) by noting that if $X \leq Y$ then

$$\{T \mid Y(\gamma) \leq T \vdash \gamma \text{ for each } \gamma \in S\} \subseteq \{T \mid X(\gamma) \leq T \vdash \gamma \text{ for each } \gamma \in S\}.$$

For the *Idempotency*, let us note that by (3.2) it follows that $X(\gamma) \leq T \vdash \gamma$ for each $\gamma \in S$ implies $(CX)\gamma \leq T \vdash \gamma$ for each $\gamma \in S$. Since $X \leq CX$ (by *Reflexivity*) it follows that $X(\gamma) \leq T \vdash \gamma$ for each $\gamma \in S$ and $(CX)\gamma \leq T \vdash \gamma$ for each $\gamma \in S$, are equivalent properties. Hence

$$\begin{aligned}(CX)\rho &= \bigwedge \{T \vdash \rho \mid X(\gamma) \leq T \vdash \gamma \text{ for each } \gamma \in S\} \\ &= \bigwedge \{T \vdash \rho \mid (CX)\gamma \leq T \vdash \gamma \text{ for each } \gamma \in S\} \\ &= (CCX)\rho.\end{aligned}$$

□

Proposition 3.6 *The interpretation of L-consequences as weak L-entailments (of Proposition 3.4) is a retract to the interpretation of weak L-entailments as L-consequences (of Proposition 3.5).*

Proof We have to show that for each weak L-entailment $(S, \vdash)$, when we interpret it as an L consequence C by (3.1) and then when we interpret C as a weak L-entailment $(S, \vdash')$ by (3.2), we have that $\vdash = \vdash'$. According to (3.1) and (3.2), we have that

$$(T \vdash' \rho) = \bigwedge \{E \vdash \rho \mid \tilde{T}(\gamma) \leq E \vdash \gamma \text{ for each } \gamma \in S\}.$$

Since $\tilde{T}(\gamma) \leq T \vdash \gamma$ for each $\gamma \in S$ it follows that $(T \vdash' \rho) \leq (T \vdash \rho)$. On the other hand, for each E such that $\tilde{T}(\gamma) \leq E \vdash \gamma$ for each $\gamma \in S$ it follows that $E \vdash \gamma = 1$ for each $\gamma \in T$. By *Cut*, it follows that $(T \vdash \rho) \leq (E \vdash \rho)$ hence $(T \vdash \rho) \leq (T \vdash' \rho)$. □

3.2 Many-Valued Consistency and Squares of Opposition

We have seen that consistency plays a crucial role in the consequence-theoretic square of opposition. The following concept of many-valued non-triviality was introduced in [16]:

Definition 3.7 [16] Any mapping $X: S \to L$ is *non-trivial* with respect to an L-consequence operator C on S if and only if $CX \neq 1$; otherwise X is *trivial*.

The following is both an entailment styled replica of Definition 3.7 and a generalization of Definition 1.4 to many-valued truth.

Definition 3.8 In any weak L-entailment $(S, \vdash)$, $T \subseteq S$ is *non-trivial* when there exists $\rho \in S$ such that $(T \vdash \rho) \neq 1$; otherwise T is *trivial*.

The following gives the link between the two concepts of many-valued triviality above.

Fact 3.9 *If $(S, \vdash)$ is the weak L-entailment system determined by an L-consequence operator C on S then any $T \subseteq S$ is non-trivial in $(S, \vdash)$ if and only if $\tilde{T}$ is non-trivial with respect to C.*

A careful look at the binary consequence square of opposition (1.1) reveals that it involves two negations, an object level one $\neg$, and a meta level one $\nvdash$. The former one is syntactic while the latter functions implicitly as the ordinary binary negation. In a many-valued context of an arbitrary lattice L of truth values, the meta level negation may be abstracted to a function $\not\!\!\times: L \to L$. We have already seen that the consequence square of opposition (1.1) relies crucially upon a consistency property for T relative to the syntactic negation $\neg$ (Definition 1.5). While this can be replicated without much thinking to a many-valued context, it would be highly inadequate both from a technical and an interpretational point of view. Much more adequate would be to lift the equivalent characterization of consistency given by Corollary 1.10 in the form of a definition of many-valued consistency. This requires also an abstraction of the binary negation as a unary function on the lattice of truth values (the function $\not\!\!\times$ below).

Notation 3.10 For any weak L-entailment $(S, \vdash)$ and any $\not\!\!\times: L \to L$, we let $T \nvdash \rho$ denote $\not\!\!\times(T \vdash \rho)$.

Definition 3.11 In any weak L-entailment $(S, \vdash)$, for any $\neg : S \to S$ and any $\not\times : L \to L$, a theory $T \subseteq S$ is $(\neg, \not\times)$-*consistent* when for each $\rho \in S$ we have that

$$(T \vdash \neg\rho) \leq (T \not\vdash \rho).$$

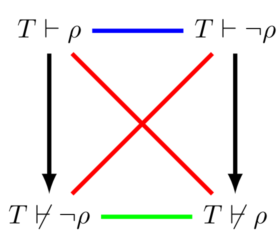
In a general many-valued setting, the equivalence between non-triviality and $\neg$-consistency given by Proposition 1.6 is not possible; however, one implication holds quite easily and is given below. In the next section, we will be able present a set of sufficient conditions for such an equivalence, but in a semantic setting.

Proposition 3.12 *Given a non-trivial lattice* L *and a weak* L*-entailment* $(S, \vdash)$*, for any* $\neg : S \to S$ *and* $\not\times : L \to L$ *such that* $\not\times 1 = 0$*, then* T *is* $(\neg, \not\times)$*-consistent implies that* T *is non-trivial.*

Proof Let $T \subseteq S$ be $(\neg, \not\times)$-consistent. By *reductio ad absurdum*, let us assume that T is trivial, i.e. that $(T \vdash \rho) = 1$ for each $\rho \in S$. For each $\rho \in S$ it follows that $(T \not\vdash \rho) = 0$, and by the $(\neg, \not\times)$-consistency hypothesis that $(T \vdash \neg\rho) = 0$, too. Since L is non-trivial, we have that $(T \vdash \neg\rho) \neq 1$, which yields a contradiction. □

The following is an immediate consequence of Definition 3.11.

Corollary 3.13 *In any weak* L*-entailment* $(S, \vdash)$*, for any* $\neg : S \to S$ *and any* $\not\times : L \to L$ *an complementation, any* $(\neg, \not\times)$*-consistent theory* $T \subseteq S$ *determines a square of opposition in* L *that generalizes the square* (1.1)



(3.3)

An example for L that gives rise to non-Boolean squares (3.3) is given by N5 above, with $\not\times a = o$ and $\neg a = e$.

3.3 Many-Valued Abstract Semantics

Now we show that many-valued consequence theoretic squares of opposition like (3.3) may arise from many-valued semantics. This means that now, besides the meta-level negation $\not\times$, the object level negation $\neg$ will be treated as a semantic entity, too. Let us extend Definition 1.2 to many-valued truth.

Definition 3.14 (Many-valued room) Given a set L, an L-*room* is a function $\models : \mathcal{M} \times S \to L$.

The following generalizes the famous Galois connection between syntax and semantics in institution theory [14].

Definition 3.15 [9] Given a complete lattice L and an L-room $\models\colon \mathcal{M} \times S \to L$, we let the following:

- For any $\mathcal{N} \subseteq \mathcal{M}$, $\mathcal{N}^*\colon S \to L$ is defined by $\mathcal{N}^*(\rho) = \bigwedge_{M\in\mathcal{N}}(M \models \rho)$; and
- For any $X\colon S \to L$, $X^* \subseteq \mathcal{M}$ is defined by $X^* = \{M \in \mathcal{M} \mid X \leq M^*\}$.

Corollary 3.16 [9] *The mapping* $(\_)^{**}\colon L^S \to L^S$ *is an* L*-consequence operator on* S.

Notation 3.17 For any complete lattice L and any L-room $\models\subseteq \mathcal{M} \times S$, let $(S, \models)$ denote the weak L-entailment determined by the L-consequence operator $(\_)^{**}$ on S according to Proposition 3.4.

The following gives the semantic concept of consistency, a consistent theory being one that has at least a model.

Definition 3.18 For any L-room $\models\subseteq \mathcal{M} \times S$, a $T \subseteq S$ is $\models$-*consistent* when $(\tilde{T})^* \neq \emptyset$.

The following generalizes Definition 1.7 to many-valued truth.

Definition 3.19 (Many-valued semantic negation) Given a function $\neg\colon L \to L$, an L-room $\models\subseteq \mathcal{M} \times S$ *admits* $\neg$ when there exists a function $\neg\colon S \to S$ such that for each $M \in \mathcal{M}$ and each $\rho \in S$

$$(M \models \neg\rho) = \neg(M \models \rho).$$

Proposition 3.20 *For any* $\neg\colon L \to L$ *and any* L*-room* $\models\subseteq \mathcal{M} \times S$ *that admits* $\neg$ *the following hold*:

1. *If* $T \subseteq S$ *is non-trivial for* $(S, \models)$ *then* T *is* $\models$*-consistent.*
2. *If* $1 \neq \neg 1$ *then* $\{\rho, \neg\rho\}$ *is* $\models$*-inconsistent.*
3. *If* $1 \neq \neg 1$ *then* $\{\rho, \neg\rho\}$ *is inconsistent in* $(S, \models)$.

Proof 1. Let $\rho \in S$ such that $(T \models \rho) \neq 1$. Since $(T \models \rho) = \bigwedge_{\tilde{T}\leq M^*}(M \models \rho)$ it follows that there exists $M \in \mathcal{M}$ such that $\tilde{T} \leq M^*$, hence $(\tilde{T})^* \neq \emptyset$.

2. By *reductio ad absurdum*, let us assume that $\{\rho, \neg\rho\}$ is $\models$-consistent. There exists $M \in \mathcal{M}$ such that $(M \models \rho) = 1$ and $(M \models \neg\rho) = 1$. Since the room admits $\neg$, it follows that $1 = (M \models \neg\rho) = \neg(M \models \rho) = \neg 1$, which contradicts $1 \neq \neg 1$.

3. Follows immediately from 1. and 2. by considering $T = \{\rho, \neg\rho\}$. □

Proposition 3.21 *Given a complete lattice* L*, functions* $\neg, \not\times\colon L \to L$*, and an* L*-room* $\models\subseteq \mathcal{M} \times S$ *such that*

- $\neg \leq \not\times$;
- $\not\times$ *is order-reversing*; *and*
- *the* L*-room admits* $\neg$,

then any $\models$*-consistent theory is* $(\neg, \not\times)$*-consistent in* $(S, \models)$.

Proof Let T be a $\models$-consistent theory and let $M \in (\tilde{T})^*$, which means $\tilde{T} \leq M^*$. The conclusion is established by the following calculation:

$$\begin{aligned}
(T \models \neg\rho) &= (\tilde{T})^{**}(\neg\rho) && \text{(by (3.1))} \\
&= \bigwedge_{\tilde{T} \leq N^*} (N \models \neg\rho) && \text{(by definition of } (S, \models)) \\
&\leq (M \models \neg\rho) && \text{(since } \tilde{T} \leq N^*) \\
&= \neg(M \models \rho) && \text{(because the } L\text{-room admits } \neg) \\
&\leq \not\times(M \models \rho) && \text{(since } \neg \leq \not\times) \\
&\leq \not\times \bigwedge_{\tilde{T} \leq N^*} (N \models \rho) && \text{(since } \tilde{T} \leq M^* \text{ and by monotonicity of } \not\times) \\
&= (T \not\models \rho) && \text{(by definition of } (S, \models)).
\end{aligned}$$

□

The following is an immediate consequence of Fact 3.9, Propositions 3.20, 3.21 and 3.12 and gives a set of sufficient conditions for the equivalence between many-valued non-triviality and consistency in an abstract semantic setup. It also constitutes a high generalization of Corollary 1.9.

Corollary 3.22 *Given a non-trivial complete lattice L, functions $\neg, \not\times : L \to L$, and an L-room $\models \subseteq \mathcal{M} \times S$ such that*

- $\neg \leq \not\times$;
- $\not\times 1 = 0$;
- $\not\times$ *is order-reversing*; *and*
- *the L-room admits $\neg$*,

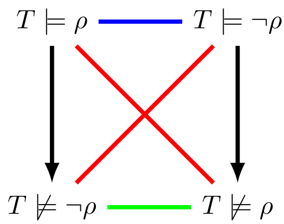
the following are equivalent for a theory $T \subseteq S$:

1. *$\tilde{T}$ is non-trivial for the semantic consequence operator $(\_)^{**}$*;
2. *T is non-trivial for $(S, \models)$*;
3. *T is $\models$-consistent*; *and*
4. *T is $(\neg, \not\times)$-consistent for $(S, \models)$*.

The binary case given by Corollary 1.9 is obtained from Corollary 3.22 by letting $L = \{0, 1\}$, and $\neg = \not\times$ the ordinary binary negation.

The following shows that in many-valued logics, semantically consistent theories give rise to squares of opposition which are in general non-Boolean.

Corollary 3.23 *If, in addition to the hypotheses of Proposition 3.21, we have that $\not\times$ is a complementation, then the following is an opposition square in L for each $\models$-consistent* *theory* T *and each* $\rho \in S$:



3.4 All Squares of Opposition Are Consequence Theoretic

The generality of the many-valued consequence theoretic concepts employed above allows for a simple representation of any square of opposition (in the general lattice theoretic sense of Definition 2.2) as a square of the form (3.3). This is based upon the following fact:

Fact 3.24 *Given a square of opposition* (a, e, i, o) *in a bounded lattice* L, *the following defines a weak* L*-entailment* $(S, \vdash)$:

- $S = \{\rho, \rho'\}$; *and*
-

$$\vdash \rho = a, \quad \{\rho'\} \vdash \rho = a, \quad \{\rho\} \vdash \rho = 1, \quad \{\rho, \rho'\} \vdash \rho = 1,$$
$$\vdash \rho' = e, \quad \{\rho\} \vdash \rho' = e, \quad \{\rho'\} \vdash \rho' = 1, \quad \{\rho, \rho'\} \vdash \rho' = 1.$$

Now it remains to note that we can define a function $\not< : L \to L$ such that $i = \not< e$ and $o = \not< a$ and define $\neg : S \to S$ by $\neg\rho = \rho'$ and $\neg\rho' = \rho$. Under these definitions, the square (a, e, i, o) is indeed of the form (3.3) by considering $T = \emptyset$.

4 Conclusions

We have provided a very general consequence-theoretic framework for the traditional square of opposition and have also presented an abstract Kripke semantics supporting the modal interpretation of the square of opposition. Next we have defined the square of opposition as a negation-free lattice theoretic structure and, within that context, performed a analysis of the relationships between its components. Moreover, the lattice theoretic hexagon of opposition came naturally as an extension of the square. We have generalized the consequence-theoretic square of opposition from the binary to many-valued truth, these being examples of non-Boolean squares, in general. This generalization is based upon a very abstract universal logic perspective on negation and consistency, in a many-valued context, and both at the consequence and model theoretic levels. Finally, we have shown that any lattice theoretic square of opposition may be represented as a many-valued consequence-theoretic one.

This work shows that the square of opposition as a general lattice theoretic structure represents a good meta-tool for a general universal logic study of negation and consistency.

Acknowledgements This work has been supported by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-II-ID-PCE-2011-3-0439.

References

1. Aristotle: De interpretatione and prior analytics. In: Barnes, J. (ed.) The Complete Works of Aristotle. Princeton University Press, Princeton (1984)
2. Béziau, J.-Y.: Classical negation can be expressed by one of its halves. Log. J. IGPL **7**(2), 151–154 (1999)
3. Béziau, J.-Y. (ed.): Universal Logic: An Anthology. Studies in Universal Logic. Springer, Basel (2012)
4. Béziau, J.-Y.: The meta-logical hexagon of opposition. Argumentos **5**(10), 111–122 (2013)
5. Béziau, J.-Y., Dale, J. (eds.): Around and Beyond the Square of Opposition. Springer, Basel (2012)
6. Blanché, R.: Structures Intellectulles. Vrin, Paris (1966)
7. Chakraborty, M.K.: Use of fuzzy set theory in introducing graded consequence in multiple valued logic. In: Gupta, M.M., Yamakawa, T. (eds.) Fuzzy Logic in Knowledge-Based Systems, Decision and Control, pp. 247–257. Elsevier Science, Amsterdam (1988)
8. Chakraborty, M.K.: Graded consequence: further studies. J. Appl. Non-Class. Log. **5**(2), 127–137 (1995)
9. Diaconescu, R.: Graded consequence: an institution theoretic study. Soft Comput. (2014). doi:10.1007/s00500-014-1231-y
10. Diaconescu, R., Stefaneas, P.: Ultraproducts and possible worlds semantics in institutions. Theor. Comput. Sci. **379**(1), 210–230 (2007)
11. Dutta, S., Basu, S., Chakraborty, M.K.: Many-valued logics, fuzzy logics and graded consequence: a comparative appraisal. In: Lodaya, K. (ed.) Logic and Its Applications. Lecture Notes in Artificial Intelligence, vol. 7750, pp. 197–209. Springer, Berlin (2013)
12. Eklund, P., Helgesson, R.: Monadic extensions of institutions. Fuzzy Sets Syst. **161**, 2354–2368 (2010)
13. Finger, M., Gabbay, D.M.: Adding a temporal dimension to a logic system. J. Log. Lang. Inf. **1**(3), 203–233 (1992)
14. Goguen, J., Burstall, R.: Institutions: abstract model theory for specification and programming. J. Assoc. Comput. Mach. **39**(1), 95–146 (1992)
15. Martins, M.-A., Madeira, A., Diaconescu, R., Barbosa, L.: Hybridization of institutions. In: Corradini, A., Klin, B., Cîrstea, C. (eds.) Algebra and Coalgebra in Computer Science. Lecture Notes in Computer Science, vol. 6859, pp. 283–297. Springer, Berlin (2011)
16. Pavelka, J.: On fuzzy logic I—many-valued rules of inference. Z. Math. Log. Grundl. Math. **25**, 45–52 (1979)
17. Priest, G., Routley, R., Norman, J. (eds.): Paraconsistent Logic: Essays on the Inconsistent. Philosophia Verlag (1990)
18. Scott, D.: Completeness and axiomatizability in many-valued logic. In: Henkin, L., et al. (eds.) Proceedings, Tarski Symposium, pp. 411–435. American Mathematical Society, Providence (1974)
19. Tarski, A.: On some fundamental concepts of metamathematics. In: Logic, Semantics, Metamathematics, pp. 30–37. Oxford University Press, Oxford (1956)

R. Diaconescu (✉)
Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania
e-mail: razvansdiaconescu@gmail.com

Universal Logic as a Science of Patterns

Brian R. Gaines

Abstract This article addresses Béziau's (Sorites 12:5–32, 2001) vision that universal logic should be capable of helping other fields of knowledge to build *the right logic for the right situation*, and that for some disciplines *mathematical abstract conceptualization* is more appropriate than *symbolic formalization*. Hertz's (Math. Ann. 87(3–4):246–269, 1922) diagrams of logical inference patterns are formalized and extended to present the universal logic conceptual framework as a comprehensible science of patterns. This facilitates those in other disciplines to develop, visualize and apply logical representation and inference structures that emerge from their problématique. A family of protologics is developed by resemantifying the sign for deduction, →, with inference patterns common to many logics, and specifying possible constraints on its use to represent the structural connectives and defeasible reasoning. Proof-theoretic, truth-theoretic, intensional and extensional protosemantics are derived that supervene on the inference patterns. Examples are given of applications problem areas in a range of other disciplines, including the representation of states of affairs, individuals and relations.

Keywords Universal logic · Inference patterns · Protologic · Protosemantics · Structural connectives · Paraconsistency · Default reasoning · Applied logic

Mathematics Subject Classification (2000) Primary 03B22 · Secondary 03A05

1 Introduction

Twenty years ago, Jean-Yves Béziau [9] proposed that the notion of *deduction* in any logical system should be studied within an integrative conceptual framework that presupposes no particular axioms. He termed this framework *universal logic* by analogy to Birkhoff's [23] *universal algebra*. This approach encompasses all logical systems and facilitates specifying each of their particular axioms whether they are common to many logics, or peculiar to a few.

Logical systems have proliferated since the 1920s when Hilbert, Frege and Russell extended the inference patterns of Aristotle's syllogistic [7, 33] to provide formal foundations for mathematics. Béziau's proposal situates a wide range of historical and ongoing studies of such systems within a coherent framework, and suggests significant directions for further research. His publications expounding [11, 13, 14] and illustrating [10, 12, 19] the approach, and the fora he has provided through editing books [15, 18], conference proceedings [21, 22], and the *Logica Universalis* journal, have inspired a research community collaborating within the universal logic paradigm.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_7

Béziau [14, p. 14] presents the study of logical structures as a mathematical discipline in its own right, a Bourbakian [25] *mother system* "having the same status as algebraic, topological and order structures". However, like other sub-disciplines of mathematics, it also provides foundational capabilities for many other disciplines. A universal logic conceptual framework contributes to these disciplines by providing techniques for tailoring logical systems to address the precise purposes for which they are required. It helps build "the right logic for the right situation" [11, p. 19], making it possible to avoid the wholesale import of inappropriate axioms of a logical system that may introduce artifacts by going beyond those deriving from a conceptual analysis of the problématique.

In order to support other disciplines, Béziau [11, p. 23] has suggested that universal logic might be presented as a formal but comprehensible conceptual framework that avoids unnecessary symbolism, that, for example, "the definitions philosophers need involve mathematical abstract conceptualization rather than symbolic formalization." However, most universal logic research has naturally adopted the symbolism of formal logic and this, like mathematical symbolism in general, can be a significant barrier to understanding [66].

This article addresses the question of whether the conceptual framework of universal logic may be formalized and presented in a way that minimizes the use of technical terminology and mathematical symbols. Its objective is to preserve formal rigor while providing a useful and comprehensible tool for those applying logic to problems of representation and inference in non-mathematical disciplines. It adopts the perspective that views mathematics as a *science of patterns* [95], and presents possible inference patterns in a logical system as two-dimensional configurations of an arrow symbolizing deduction.

2 Logic as a Science of Patterns

The objective of this section is to develop the conceptual framework of universal logic in a simple and comprehensible form as a science of patterns. The primitive notion of *deduction* is represented as a process of recognizing a pattern within a structure that licenses the addition of deletion of part of that structure. This process is itself represented by metastructures termed *inference patterns*, a collection of which will be said to constitute a *protologic* [79].

2.1 Foundations for a Universal Protologic

The genesis of what has come to be termed universal logic was in the philosophy of Hilbert that he derived by reflecting on his experience in proving his basis theorem [68] and rationally reconstructing Euclidean geometry [69], and the resultant controversy with Gordan [83, p. 18] and Frege [49, pp. 1–24] about his innovations in the logical foundations of mathematics.

Hilbert [70] evolved a new conceptual framework for *axiomatic thinking* that involved reconstructing a formalized discipline by abstraction to a minimal set of independent axioms, each having a meaningful interpretation in that discipline. His methodology introduced notions such as *logical existence* [49, p. 12] being equivalent to lack of contradiction, and *ideal elements* [71] being introduced in order to simplify the axioms even though they were not part of the original system.

Hilbert [70, p. 413] emphasized that the axiomatic foundations of logic underpinned those of other disciplines and themselves needed to be made secure. When he and his colleagues at Göttingen, such as Hertz, Bernays and Gentzen, worked on this problem it was natural for them to adopt his principles of axiomatic thinking and deconstruct logical deduction as a minimal collection of axioms, introducing ideal elements as necessary to simplify them, and focusing on freedom from contradiction and the complete reconstruction of expected inferential outcomes.

The objective was to provide logical foundations for mathematics and the axiomatic method rather than to characterize all possible logics. However, there were already two contending logics for mathematics, classical and intuitionistic [58], and it was natural for Hertz [67] in applying Hilbert's axiomatic thinking to logic to abstract as much as possible and consider inference patterns between arbitrary sentences. He even abstracted from the notion of inference itself, stating that there is no need to specify "what the symbol $\rightarrow$ linking the characters $a \rightarrow b$ or the word 'if' in the corresponding linguistic formulation means" [67, p. 247].

Bernays further extends Hertz's level of abstraction when he used it to exemplify Hilbert's philosophy of mathematics by reducing $\rightarrow$ to a *sign* rather than a *symbol*: "If the hypothetical relationship 'if A then B' is symbolically represented by $A \rightarrow B$, then the transition to the formal position is that we abstract from that the meaning of the symbol $\rightarrow$ and take the linkage by the 'sign' $\rightarrow$ itself as the primary consideration." [8, p. 333].

In her analyses of the evolution of written language to include non-phonetic technical material involving mathematical and logical symbology, Krämer [75] has characterized such extreme abstraction as complete *desemantification*. Dutilh Novaes [39] adopted this terminology and situated the desemantification of logic historically through her analysis of the use of the qualifier *formal* in the logical literature. She [40, §6.1.2] introduces the term *resemantification* to describe the process of reintroducing expected features of a desemantified system.

The following sections develop a universal protologic by commencing with deduction as a desemantified sign, $\rightarrow$, and incrementally resemantifying it by introducing common logical constraints as inference patterns represented by structures based on $\rightarrow$. The additional term *semantification* is used to distinguish various extra-logical interpretations that add meaning without changing the underlying logical system, including: those intrinsic to the logical system, such as truth values (Sect. 4.2), and intensional (Sect. 4.3) and extensional (Sect. 4.4) reconstruction.

2.1.1 Protosemantics

Whilst the protologic is itself a mathematical abstraction, it is intended to have practical applications and terms have been adopted to name abstract patterns that reflect those in the literature and seem natural to those patterns. However, these terms are strictly technical and none of their possible connotations beyond their formal definitions are used to draw inferences. That is, the protologic semantifies the terms and not the terms the protologic.

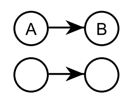
For example, $A \rightarrow B$ is read as 'A includes B' and two complementary technical terms are also introduced to provide an abstract protosemantics. The term *content* is used to reference that which might be included. That is, A → B may be read as 'the content of A includes the content of B.' This terminology is consistent with that of those who have proposed that logical entailment be explicated as *meaning* [41], *sense* [90] or *content* [26] *containment*, and enables one to assess when that notion is appropriate and when it is not.

The term *context*, is used to reference the effect of inclusion, of being within the scope of the meaning. That is, A → B may be read as 'the context of B encloses the context of A.' This terminology is consistent with Aristotle's [1, 1061b30] use of the term *qua* in his *Metaphysics* to introduce a context, for example, to consider physical objects 'qua moving' rather than 'qua bodies.' However, any connotations of the terms 'content' and 'context' derive explicitly from the constraints placed upon the use of →, rather than from *a priori* intuitions.

2.1.2 Arrows, Links and Graphs

Béziau [9, p. 85] introduces a minimal resemantification of a *deduction* sign by going back to the Latin roots of the term and interpreting it as *leading away* from premise to conclusion, that is, we may regard the arrow in A → B as a directed connective leading from the symbol 'A' at the tail to the symbol 'B' at the head. The resulting structure may be termed a *link* constituted by the triple (A, →, B). It may also be described as a *link out of* A constituted by the pair (→, B) or a *link in to* B constituted by the pair (A, →).

A link is naturally represented in graphical form as two labeled nodes with an arrow between them or, more generically, as an arrow between two anonymous nodes. It is assumed that labels constitute unique identifiers from some family of identifiers with an equality relation, and that an anonymous node has an implicit label unique to that node.



The resemantification involved in labeling nodes is the assumption that the entities linked by → may be identified, distinguished and equated. Node labels may be chosen to suggest possible connotations but these make no formal contribution to the logical structure represented by the links. They are logically meaningful only to the extent that the linkage structure represents such connotations.

Nodes with identical labels will be taken to represent the same node shown more than once, and may be merged to form a canonical graph-like structure with no duplication of nodes. This allows structures to be split into substructures, possibly overlapping, that can be merged to reconstitute the original structure. It also allows a structure to be merged with one of several other structures, each representing an alternative component, for example, a different 'ontology' or 'theory.'

Figure 1 shows a number of links merged into a graphical structure, S, with no duplication of nodes.

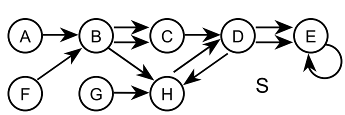


Fig. 1 Graphical structure specified by multiple links

Note that the graph representation adds no additional information to that of a linear representation as the multiset,

$$S \equiv [A \to B, B \to C, B \to C, C \to D, D \to E, D \to E, E \to E, F \to B, B \to H, G \to H, H \to D, D \to H]$$

It is only a more perspicuous representation of the structure represented by multiple links involving the same nodes.

Figure 2 shows the same links merged into three overlapping substructures, S1, S2 and S3, representing the same structure in a modular way.

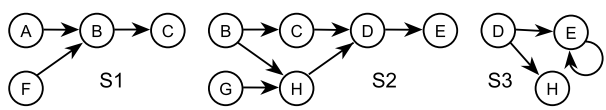


Fig. 2 Same structure specified modularly as three substructures

The linear representations are shown below and S is their multiset sum:

$$S1 \equiv [A \to B, B \to C, F \to B]$$

$$S2 \equiv [B \to C, C \to D, D \to E, B \to H, H \to D, G \to H]$$

$$S3 \equiv [D \to E, E \to E, D \to H]$$

$$S \equiv S1 \uplus S2 \uplus S3$$

The resemantification of the arrow sign so far is sufficient to support the basic structures of graph theory. However, as evident in the examples, there are no constraints to preclude parallel arrows between the same nodes or loops from a node to itself. Hence, in graph-theoretic terms, the examples given are not strictly *directed graphs* but rather *nets* [65, pp. 4–7] or *directed pseudographs* [4, p. 4].

There is also, as yet, no support for the inferential processes of logic. However, in the continuing resemantification inference pattern are introduced such that $\to$ represents a partial order underlying a logical system (Sect. 2.5).

2.1.3 Inference Patterns and Invariance Under Logical Interpretations

While it is convenient to use the terminology of graph theory to describe the nets that represent collections of links, logical theory focuses on the *dynamics of change* in such collections. This is not a primary concern of graph theory [64].

The logical dynamics of nets will be captured in terms of *inference patterns* in which an abstract subnet of a particular form is recognized as licensing the addition or deletion of one or more links while leaving any 'logical interpretation' of the net invariant. These notions are formalized in the following sections.

A *logical interpretation* of a net is defined to be an inference-preserving conservative translation [29] of the net into statements of a logical system supporting some notion of inference. To state that it is *invariant* under a change to the net is to assert that, within that logical system, the translation of the changed net will have the same inferences as that of the original net. Two nets with the same logical interpretations will be termed *logically equivalent*.

This is a constraint upon appropriate translations, that the target logical system must implement within its own framework the inference patterns of the protologic, and this must be verified for each translation. It is also intended that the inference patterns in the protologic can be understood in their own right and that the graphical language can be used to represent the form of knowledge structures and the dynamics of inference.

An inference pattern may be seen as an *analytic invariant* of a net in that it can dynamically expand a net by adding logical inferences implicit in its links, and hence can also delete them as being superfluous, contracting the net, possibly to a minimal form. Additions or deletions that maintain the net logically invariant will be termed *conservative expansions* or *conservative contractions*, respectively.

2.2 Inclusion Inference Pattern

A significant example of resemantification through a logical inference pattern is that *transitivity* of → which is common to its usage in most logical systems and has been taken to be a characteristic feature of 'a logic' [101]. Figure 3 shows Hertz's [67, Fig. 1] diagram to specify the transitivity of → on the left, and, in the center, its representation by an *inclusion inference pattern* in a net.

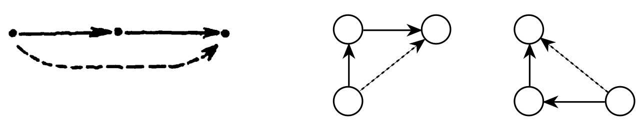


Fig. 3 Inclusion inference pattern

A metalogical distinction has been made in the inference pattern by showing the pattern-defining arrows as solid lines and the inferred arrow as a dotted line. The inference pattern indicates that if the pattern-defining links are found then an inferred link may be added, or any existing inferable link may be deleted, without changing the logical interpretation of the net of which the pattern is part.

The same pattern is drawn differently on the right to show that the inclusion inference pattern might be visualized either as the copying, or inheritance, of a horizontal inclusion link downwards from the head to the tail of a vertical inclusion link, or of a reverse horizontal inclusion link upwards from the tail to the head of a vertical inclusion link.

If x, y and z are links, S a multiset of links, and ≡ indicates identical logical interpretations, an equivalent linear representation might be:

$$\text{for any x, y, z, S,} \quad [S, x \to y, y \to z] \equiv [S, x \to y, y \to z, x \to z]$$

One advantage of using the two-dimensional structure of the page to provide a non-linear graphical presentation is that, as Shönfinkel [99, p. 17] has noted, the introduction of variable names is distracting because they serve merely to link multiple occurrences of the same logical entity. This linkage may be specified formally in a net without requiring an artificial label for a generic node.

However, the graphic representation has introduced no additional constructs beyond those of a conventional linear representation. It is a formal specification with no dependence on visual intuitions that avoids a proliferation of symbols.

2.2.1 Cycles and Node Equivalence

A cycle of mutual inclusion between two nodes, as shown in Fig. 4 left, may be treated as a single symmetric link with the directional arrow heads omitted as shown in the center.

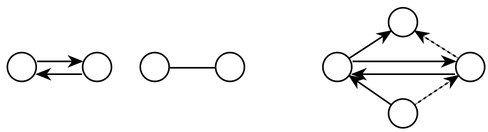


Fig. 4 Cycles and node equivalence within cycle

As shown on the right, such a cycle induces an equivalence between nodes because the inclusion inference pattern implies that any nodes which includes, or is included in, one node includes, or is included in, the other node.

This generalizes to larger cycles because the inclusion inference pattern implies that all nodes in a longer cycle are also linked pairwise in simple cycles, and hence all nodes in a cycle are equivalent (Fig. 5)—Hertz's [67, Definition 10] *web*.

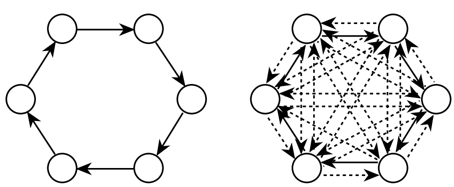


Fig. 5 Inference of node equivalence within a longer cycle

It will be noted in introducing other arrows and their link types that, as they are defined in terms of →, the node equivalence induced by a cycle generalizes to commonality of links of any type. Any nodes in a cycle have the same incoming and outgoing links of all types and are equivalent in this respect, and their labels may be regarded as aliases for one another under this equivalence.

2.2.2 Equivalence and Loop Inference Patterns

Another commonly expected logical constraint on → is that should it be *reflexive* so that every node has an associated loop (as illustrated by node E in Figs. 1 and 2). Rather than being introduced as an *ad hoc* assumption, reflexivity can be derived from a more specific requirement that nodes equivalent to any node may be added or deleted without change of logical interpretation. This corresponds to the *equivalence inference pattern* shown on the left of Fig. 6, which may be explained in terms of nodes always being able to have more than one label.

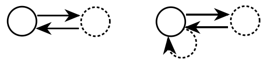


Fig. 6 Equivalence inference pattern and derived loop

As shown on the right, application of the inclusion inference pattern to the equivalence inference pattern implies that the node has an arrow to itself, a *loop*. These results may be used to derive the *loop inference pattern* corresponding to reflexivity through the inference sequence shown in Fig. 7.

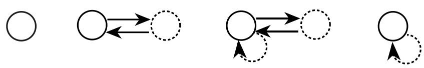


Fig. 7 Derivation of loop inference pattern

From left to right: the equivalence inference pattern is used to add an equivalent node to an isolated node; the inclusion inference pattern is used to infer that the original node has a loop; and the equivalence inference pattern is used to delete the equivalent node resulting in the loop inference pattern on the right.

Reflexivity is significant in the theoretical development of logics but irrelevant to practical applications where, whatever an inclusion link represents, it is unlikely to be useful to infer that a node includes, or is included in, itself—the epitome of a circular argument.

2.3 Repetition Inference Pattern

Another significant example of resemantification through a logical inference pattern is that for adding or deleting parallel arrows between nodes. For many logical interpretations multiple links of the same type in the same direction between two nodes do not affect the logical interpretation of a net. Their addition or deletion is a conservative expansion or contraction.

This could be implemented by requiring a collection of links to be a *set* rather than a multiset or by requiring the graph to represent a *relation* rather than a net. However, in some logical systems such as linear logic [59], the repetition of a statement may be significant, and the lack of significance in other logics is best represented explicitly as an *repetition inference pattern* (Fig. 8).

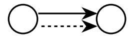


Fig. 8 Inclusion repetition inference pattern

2.4 Protoconjunctions and Protodisjunctions

When a node includes, or is included by, two or more other nodes, it is common in interpretations of nets in other logical formalisms to shorten the representation by collecting the labels of the other nodes separated by conjunction or disjunction symbols, respectively. For example, a description logic [3] translation of the nets in Fig. 9 might be:

$$A \sqsubseteq (B \sqcap C) \qquad (A \sqcup D) \sqsubseteq B$$

Fig. 9 Protoconjunction and protodisjunction

However, the nets might equally well be translated without the introduction of conjunction and disjunction symbols as:

$$A \sqsubseteq B \qquad A \sqsubseteq C \qquad D \sqsubseteq B$$

The use of conjunction and disjunction symbols in the first translation, introduced only to collect terms rather than compose them to define an ideal node, will be characterized as specifying a *protoconjunction* or *protodisjunction*. Protoconjunctions and protodisjunctions satisfying certain extremal conditions to define an additional ideal node as a composition of links will be termed *structural conjunctions* (Sect. 5.1) and *structural disjunctions* (Sect. 5.2), respectively. Similar considerations lead to a distinction between *protonegations* and *structural negations* (Sect. 2.8.2).

The significance of these distinctions is that the semantics, ontological commitments and inferential complexity of the protoconnectives is substantially simpler than that of the structural ones, and that the representation of many significant generic knowledge schemata requires only the protoconnectives [51]. However, the linguistic usage of 'and,' 'or' and 'not' does not give a clear indication of the type of connective intended and neither do most logical symbolisms.

2.5 The Resemantified → as a Preorder or Partial Order

Given any net, if one conservatively contracts it by merging it to canonical form, using the inclusion inference pattern to equivalence the nodes in a cycle, the equivalence inference pattern to delete all but one node in the cycle, and the loop inference pattern to delete any loops, the resultant net is a *directed graph* [4, §1.2].

It can be further contracted by using the inclusion inference pattern to delete all links that can be inferred from it resulting in the *transitive reduction* [4, p. 177] of the cycle-free graph as the *minimal canonical form* of the original net (unique up to the node label chosen to represent all the nodes in a cycle). Figure 10 shows S_{min}, the minimal canonical form of the net specified in Fig. 2.

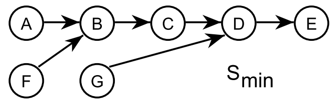


Fig. 10 Net from Fig. 2 reduced to a minimal canonical form

The transitive, reflexive relation defined by the repetition, inclusion and equivalence inference patterns indicates that, if → complies with these patterns, it specifies a *preorder* on the nodes. If the node equivalence of Sect. 2.2.1 is taken to define equality of nodes then the preorder becomes a *partial order*.

A preorder has been taken by Straßurger [101] to answer the question "what is a logic?" As the remaining constructions in this article are defined in terms of the preorder it might seem that the preceding sections have already resemantified → adequately to specify a logical system. The sign → has become the symbol for an *abstract copula* [37, p. 104] capturing the essential features of deduction.

However, the preorder alone offers limited representation and reasoning capabilities. The additional metalogical definitions that follow significantly extend the representation and inference capabilities of the protologic. Because they are always available as constructions within the protologic (as constraints on the usage of →), one could argue that they are inherent in the order relation. One could also argue that their definition is a significant additional resemantification of the sign → leading to a richer notion of what it is to be a logic.

2.6 Definition of an Exclusion Arrow and Link

In representing logical inference, what is most obviously missing in the structures discussed so far is the notion that they may be structurally unsound or *inconsistent*. Every net based on the inclusion link, and every net derived from it using the various inference patterns, is a legitimate logical structure. Nodes and links may be added to any net without logical constraint. There is support for the notion that derived links are logically necessary but none for the notion that some potential links may be logically impossible.

The *exclusion arrow* ⇸ supports the addition of a different type of link to a net that constrains what further links may be added to it. An *exclusion link* is defined in terms of the inclusion arrow:

Definition An *exclusion link* may be constituted with ⇸ iff any node that includes the tail is excluded from also including the head.

The meta-inference pattern for this definition is shown in Fig. 11. A red cross is used as a metalogical symbol to indicate that an inclusion link is prohibited.

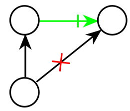


Fig. 11 Exclusion link definition

The tail node of an exclusion link is said to *exclude* the head node, and the head node to be *excluded by* the tail node. The head node may also be termed *opposite to*, *contrary to* or *incompatible with* the tail node.

2.6.1 Exclusion Inference Patterns

Three inference patterns may be derived from the metalogical definition of the exclusion link. First, it complies with a repetition pattern, the *exclusion repetition inference pattern* (Fig. 12 left).

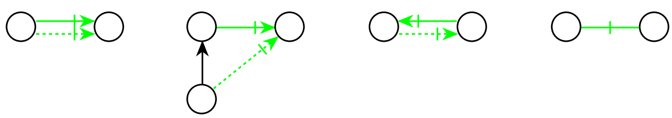


Fig. 12 Exclusion inference patterns

Second, because any node that includes the lower node also includes that to which it is linked (inclusion pattern), it cannot have a link to the node on the right. Hence the lower node satisfies the exclusion link definition and has an exclusion link to the node on the right. The resulting pattern is the *exclusion inference pattern* (Fig. 12 second from left).

Third, any node that includes the right hand node cannot also include the left hand node since the exclusion link definition would be contravened. Hence there is a also an exclusion link from right to left, resulting in the *exclusion symmetry inference pattern* (Fig. 12 second from right). This pattern marks a major difference in the inferential dynamics of the two link types, that the exclusion link is symmetric and can be represented by the single undirected link on the right.

The undirected link has repetition and exclusion inference patterns derived from those of the exclusion link (Fig. 13 left). The right two nets show a *reverse exclusion inference pattern* that also follows from the exclusion inference and symmetry patterns. The variant on the far right shows how this may also be seen as a *reverse inclusion inference pattern* such that an exclusion link to the tail of an inclusion link may be inferred from one to the head.

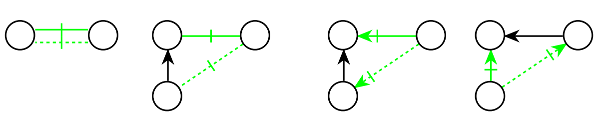


Fig. 13 Exclusion link and reverse exclusion/inclusion patterns

The exclusion inference pattern implies that if two nodes are in a cycle then any node which excludes, or is excluded by one node, excludes, or is excluded by the other. This generalizes the equivalence between nodes in a cycle defined in Sect. 2.2.1 to exclusion as well as inclusion links.

From the perspective of meaning containment, an exclusion link between two nodes signifies that the contents of the two nodes are *incompatible* and their contexts are *disjoint*. That is, some part of the content of one node cannot be included with some part of the content of the other node, and no node may be enclosed in the context of both the nodes.

From an inferential perspective the directed exclusion link can always be replaced by the undirected one with no change in logical interpretation; exclusion is always mutual exclusion. However, from a semantic perspective the metalogical distinction between the *specification* of the net and the *inferences* ensuing from that specification may be significant. There may, for example, be significance to the notion that the specified direction of the arrows depicts how a new node *depends* on existing nodes. Hence the symmetry of the exclusion link is, for some purposes, best represented as an inference rather than an intrinsic feature.

2.6.2 Inconsistency Inference Pattern

The definition of an exclusion link in Sect. 2.6 introduces a metalogical constraint that an inclusion link is not allowed between two nodes. If this constraint is violated in the net, either by specification or by inference, the outcome will be that two nodes are connected by both inclusion and exclusion links as shown in Fig. 14 (the exclusion link is shown in symmetric form to encompass all cases of parallel inclusion and exclusion links).

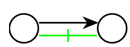


Fig. 14 Inconsistent link

Such a pair of links will be considered as constituting a single link combining conflicting link types and termed an *inconsistent link* from the tail of the inclusion link to its head. The tail node will be termed an *inconsistent node*, and the net of which it is part an *inconsistent net*.

The inconsistency inference pattern shown in Fig. 15 derives from inclusion and exclusion inference patterns. If one node has an inconsistent link to another then any node that includes it also has an inconsistent link to it and to that node, as will any down a path to it. Since any node is down a path from itself, any node with an inconsistent link also has an inconsistent link to itself.

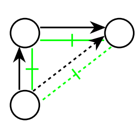


Fig. 15 Inconsistency inference pattern

An inconsistent link in a net, whether specified or inferred, implies that the exclusion link definition has been violated by the specification of a node that includes both the tail and the head of an exclusion link. It derives from the specification of net and indicates a structural inconsistency in that specification.

2.6.3 Paraconsistency

In translations of the protologic to some logical systems, inconsistency cannot be represented and is treated simply as an error in the specified logical structure. In a *paraconsistent* system [20, 91], meaningful inference will generally still be possible, and the protologic supports both possibilities. Whether the inconsistency corresponds to an intrinsically inconsistent entity or to an erroneous specification is an extra-logical issue that does not affect the inference patterns involved in representing inconsistency.

An inconsistent link in a net is *localized* in that it can only propagate through the inclusion and exclusion patterns to a node which includes the tail of its inclusion link. Inconsistency of some links in a net does not 'explode' to inconsistency of all links in the net. Thus, the logical inference schemata that have been defined are paraconsistent in their containment of inconsistency. It is possible to reason normally both about the parts of the net which are not inconsistent and about the propagation of the inconsistencies themselves. The principle of non-contradiction continues to apply, but, if the specification of a net violates it, the adverse impact of the violation is localized to a sub-net.

For example, the net shown in Fig. 16 illustrates the inferences when the definition of the exclusion link specified between nodes A and B is violated by node C which is specified to include both. C is inferred to have inconsistent links to A and B. These propagate to node D which is specified to include C. However, the exclusion link between A and E propagates without inconsistency to C and D, as does the inclusion link from B to F. An inconsistent entity can still exhibit normal patterns of inference for aspects unrelated to its inconsistency, and inconsistency itself propagates in a meaningful way.

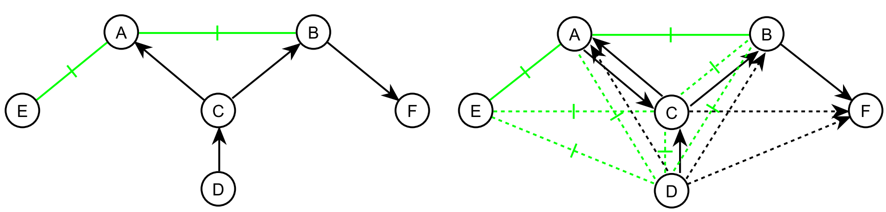


Fig. 16 Localization of inconsistency

2.7 Generic Inference Patterns for the Kernel Protologic

The protologic based on the resemantification of → with the inference patterns defined so far will be termed the *kernel protologic*. It is constituted through inference patterns constraining $\rightarrow$ to represent the partial order expected of derivation relations together with a further constraint $\nrightarrow$ restricting the placement of $\rightarrow$ to represent inconsistency.

The kernel protologic is extremely simple, yet logically powerful enough to illustrate the dynamics of logical systems that differentiate them from other mathematical structures (Sect. 3), to provide protosemantics underlying a wide range of semantic interpretations of many logics (Sect. 4), and to support the representation of a wide range of common knowledge representation schemata, such as determinables, graded scales, taxonomies and frames [51].

Structural connectives may be defined through constraints on nodes defined in terms of the basic inference patterns, but only very restricted forms of them are needed for much practical reasoning (Sect. 5). Nonmonotonic inference may be represented through a preference relation between nodes (Sect. 6).

The inference patterns for the kernel protologic have similarities that allow them to be condensed to a generic form. For example, the exclusion inference pattern mimics the inclusion inference pattern. If one shows them together (Fig. 17), Fahlman's [44] process of *virtual copying* is apparent. Both inclusion and exclusion links out of the top center node are 'copied' by inference to the lower node so that the links do not need to be shown explicitly but can be treated as 'implicit,' 'virtual' or 'inherited.'

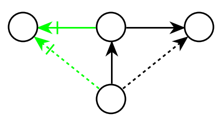


Fig. 17 'Virtual copying' of links

The inference patterns for $\rightarrow$ and $\nrightarrow$ are common to six link types, inclusion, equivalence, exclusion (in both directions and symmetric form), and inconsistent. These links may be treated as parametrized instances of a generic link, symbolized by $\dashrightarrow$, allowing the inference patterns to be represented generically (Fig. 18).

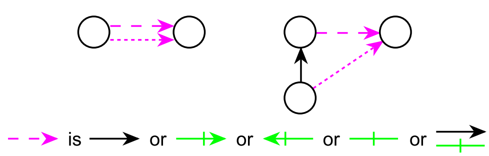


Fig. 18 Generic inference patterns

From an algebraic perspective, the inference patterns for $\dashrightarrow$ signify: left, that the instances represent relations; right, that the relations are idempotent when residuating $\rightarrow$ on the right. That is, in the Boolean algebra of binary relations, $\rightarrow\backslash\dashrightarrow\,=\,\dashrightarrow$ for the six substitution instances of $\dashrightarrow$ that are shown. This is the algebraic basis of notions of 'copying' and 'inheritance' and foundational to studies of logics as residuated lattices [53].

2.8 *Definition of a Coexclusion Arrow and Link*

One may interpret the inference pattern for an inclusion link (Fig. 3 center) as requiring that any node with an inclusion link to the tail also has an inclusion link to the head, and for an exclusion link (Fig. 12 second left) that it has an exclusion link to the head. There are also two obvious complementary link types where an exclusion link to the tail node requires either an inclusion or exclusion link to the head node. The latter is already available as the converse of an inclusion link (Fig. 13 right) but the former is a new link type that will be termed a *coexclusion link* and represented by an arrow with two bars $\nrightarrow\!\!\!\!\nrightarrow$.[1]

The formal definition of a coexclusion link is:

Definition A *coexclusion link* may be constituted with $\nrightarrow\!\!\!\!\nrightarrow$ iff any node that excludes the head must include the tail.

This definition is illustrated in graphical form in Fig. 19:

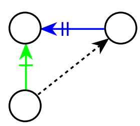


Fig. 19 Coexclusion link definition

From the perspective of meaning containment, a coexclusion link specifies that if the content of the head is excluded then that of the tail is included, and if a node is not within the context of the head then it is within the context of the tail. One might also interpret this definition as specifying that the content of the tail of a coexclusion link is included in the content of the tail of any exclusion link with the same head, and that the tail of a coexclusion link provides a context for the tail of such an exclusion link.

2.8.1 Coexclusion Inference Patterns

Figure 20 shows derived inference patterns for coexclusion links. From left to right, coexclusion links have a repetition pattern, are symmetric allowing the arrow heads to be dropped (subject to semantic considerations) to provide an equivalent line form, and have a downward inheritance pattern similar to that of the reverse inclusion link (Fig. 3).

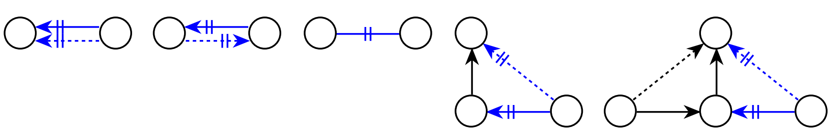


Fig. 20 Coexclusion inference patterns

Figure 21 shows a generic representation of the coexclusion inference patterns.

[1] *Coexclusion* is a neologism reflecting the duality between exclusion and coexclusion links. In the ancient logical literature, the terms *subpares* (Apuleius) and *subcontarias* (Boethius) have been used for a coexclusion relation in the metalogic of the syllogistic.

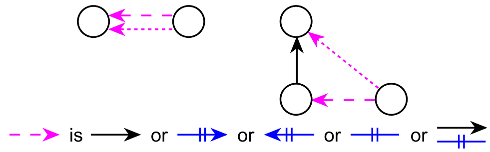


Fig. 21 Generic inference patterns for inclusion and coexclusion

From an algebraic perspective, the inference patterns for $\dashleftarrow$ signify: left, that the instances represent relations; right, that the relations are idempotent when residuating $\rightarrow$ on the left. That is, in the Boolean algebra of binary relations, $\dashleftarrow / \rightarrow = \dashleftarrow$ for all the substitution instances of $\dashleftarrow$.

2.8.2 Negation Inference Patterns

Exclusion and coexclusion links may coexist between the same nodes without inconsistency, and the combination of exclusion and coexclusion links acts as a structural negation in the protologic. Figure 22 shows some of the possibilities, all of which are logically equivalent but may be semantically distinct in indicating different origins of the negation.

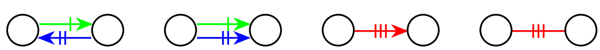


Fig. 22 Structural negation as paired exclusion and coexclusion

The negation aspects of the combination are apparent in the inference patterns shown in Fig. 23. On the left it can be seen that an inclusion link to one node results in an exclusion link to the other, and *vice versa*.

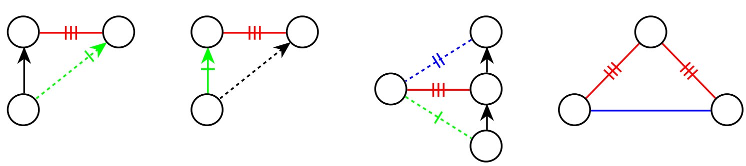


Fig. 23 Structural negation inference patterns

In the center, it can be seen that a node which includes one node excludes the other, and one that is included coexcludes the other. The lower exclusion inference captures Brandon's [27, p. 126] definition that a contradiction is that which is entailed by any contrary, and Dunn [38, p. 10] and Brady's [26, p. 20] that it is a disjunction of all possible contraries. The upper coexclusion inference provides dual definitions: that a negation entails any coexclusion and is the conjunction of all possible coexclusions. The inference pattern may also be seen as splitting a negation into its exclusion and coexclusion components.

On the right is shown the inference pattern for double negation, that if two nodes have negation links to the same node then they are equivalent. The inferred inclusion links derive directly from the coexclusion definition of Fig. 19. This pattern also shows that the negation of a node is unique modulo equivalence.

The exclusion and coexclusion link types may be seen as a *protonegations*: that the inclusion of some content is incompatible with the inclusion of some other; or that the exclusion of some content requires the inclusion of some other.

The latter is a less natural constraint than the former and may be seen as the source of the *semantic fragmentation* [45] that led to Plato and Aristotle's critiques of a bare negation used as if it generated a meaningful concept. For example, that, whilst 'being Greek' is a meaningful concept, use of the term 'barbarian' for 'not being Greek' does not define a concept having a coherent meaning [89, 262d].

2.8.3 Coexclusion Squares and Hexagons of Oppositions

Figure 24 left shows a square net with coexclusion links on the diagonals and an exclusion link on the top side such that inclusion links may be inferred on the left and right sides and a coexclusion link on the bottom side. Similar structures have been studied in the logical literature, both ancient [78] and modern [21, 22], as *squares of oppositions*.

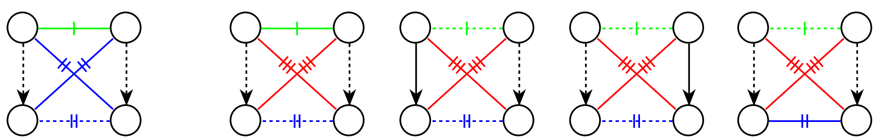


Fig. 24 Coexclusion squares of oppositions

The inferred implications on the left and right of the square are derived directly from the coexclusion definition. The inferred coexclusion link is derived because any node with an exclusion link to one of its nodes will also have an exclusion link to the node above (exclusion inference pattern) and hence an inclusion link to the other node (coexclusion inference pattern) and hence satisfies the coexclusion definition.

In the net on the left, having only coexclusion links as diagonals, the additional inclusion and coexclusion links may be derived from these and the exclusion link at the top. However, if this link is omitted, it cannot be derived by specifying one of the links previously derived.

However, if exclusion links are added to the coexclusion links on the diagonals to make them negations then specifying the link on *any* side allows those on the other three sides to be derived (Fig. 24 right). These are the classic *squares of oppositions* that are common to many logical systems [21].

The net on the left demonstrates that the main phenomena of interest in squares of oppositions derive from the co-exclusion component of the negations on the diagonals, that a node that excludes a node at one end of the diagonal must include that at the other end.

Blanché [24] extended the square of oppositions with nodes to a hexagon of oppositions by adding the disjunction of the top nodes and the conjunction of the bottom ones. His construction has proved significant in a universal logic framework in representing the relations between major constructs in a wide range of logical systems [17, 19], and is interesting to see how it may be represented in the kernel protologic with coexclusion.

Figure 25 left shows the basic square extended to a hexagon with a protoconjunction at the bottom and a protodisjunction at the top, and that one may be inferred from the other. On the right this is extended to a full hexagon of oppositions by adding an exclusion link between the top two nodes of the square and inferring the other links on the sides of the square. As already shown in Fig. 24, any of the four links on the sides of the square could be added to infer the others.

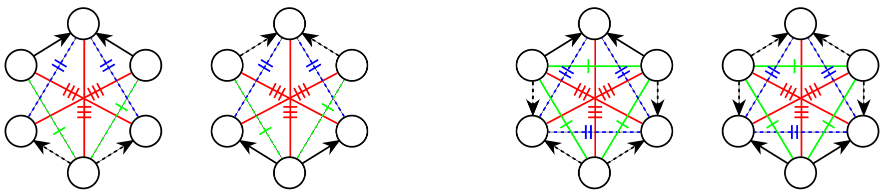


Fig. 25 Inferences in a square extended to hexagon

Thus the logical derivations involved in the Blanché hexagon may be factored through the inclusion, exclusion and coexclusion links involved in their derivations, and only protodisjunction and protoconjunction nodes are required for the top and bottom nodes that he added, not the ideal nodes of structural definitions (Sect. 5).

3 Dynamics of Logical Structures

As noted in Sect. 2.1.3, although logical structures may be represented in graph-theoretic form, logic goes beyond graph theory in its emphasis on the *dynamics of change* in those structures. This encompasses not only the addition and deletion of links in nets through inference patterns but also various ways of restructuring nets, splitting and merging of nets, the specification of additional nodes and links, and the consequences for the logical interpretations of the nets involved. It also involves considerations of effective means of communicating the dynamics to those developing and using knowledge structures represented in nets.

3.1 Minimal and Maximal Canonical Forms

In has been noted (Sect. 2.5) that a net with inclusion links can be reduced to a minimal canonical form having the same logical interpretation. The exclusion and coexclusion inference patterns support a similar deletion of exclusion and coexclusion links that may be regarded as superfluous because they can be inferred from others. However, the symmetry of these links means that the direction in which they are specified is inferentially irrelevant and it is appropriate to treat the symmetric form as canonical in deriving a unique canonical form. Similarly, the repetition inference pattern for all links means that it is appropriate to treat reduction to a single link as canonical for the sake of unicity.

These considerations lead to a minimal canonical form for a net with that is unique up to the choice of node label used to represent all the nodes in a cycle, and has the same logical interpretation as the net from which it was derived. The inclusion, exclusion and coexclusion inference patterns may be used to expand this minimal canonical form to a unique maximal one. Figure 26 illustrates the minimal and maximal canonical forms of a specified net of inclusion and exclusion links.

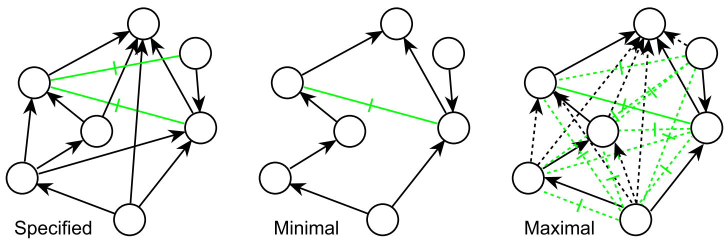


Fig. 26 Example of minimal and maximal canonical forms

The inference patterns of Sect. 2 and their use to expand or contract the net as discussed in this section exemplify net dynamics through the addition or deletion of links. The following sections extend this to expansion and contraction of nets through the addition or deletion of other nets (including single nodes). Since some expansions do not change the logical interpretation of the net, this allows for further expansion beyond the maximal canonical form to a net that is logically equivalent modulo added nodes.

3.2 Consistent and Conservative Imports and Merges

Nets are dynamic, not only through inference patterns that do not change their logical interpretation, but also though the addition of links that are not implicit in the inference patterns and hence change the interpretation, and through the addition of nodes and associated links that may have a variety of effects.

When nodes are added they may already be named and/or linked, and the general case may be considered as one where one net is *imported* by another, either through merging nodes with the same names, or by linking nodes in one net with nodes in the other, or both.

One major consideration is whether the resultant net has additional inconsistent links. An import is termed *consistent* if it does not. Usually the focus is on the import of a consistent net by a consistent net to produce a consistent net, but it can be appropriate to consider situations where the importing or imported net is inconsistent.

Another major consideration is whether the import is *conservative* in that no additional links may be inferred between the nodes in the imported net. This is of concern, for example, in many computational applications where the imported net is intended to be a generic module representing a 'library,' 'ontology' or 'theory.' The intent is to generate additional inferences in the importing net but not in that which is imported [62].

A conservative import may impact the dynamics of the importing net by making inconsistent potential links that could have been added to that net, that is, constraining it more strongly. An import that imposes no such constraints is termed *ultra-conservative*.

An import of one net by another may also be regarded conversely as an import by the second net of the first. If both imports are conservative, or ultra-conservative, it may be termed a conservative, or ultra-conservative, *merger* of the nets.

Since the inference patterns are such that only inclusion or coexclusion links from nodes in one net to nodes in another can generate new links within the first net, importing a net such that there are no inclusion or coexclusion links from the imported net to the importing one is always a conservative import, but generally not ultra-conservative.

Conservative imports are reversible through deletion of the imported net provided this deletion is from the maximal canonical form. This restriction is necessary because the purpose of some imports is to simplify the net by reducing the number of links in the minimal canonical form and, hence, the links removed from the original net must be restored before the imported net is removed (for example, in the *factoring* expansion pattern of Sect. 3.4).

3.3 Single Node and Link Expansion Patterns

Adding an additional anonymous node as an *ideal element* with up to one link to a node in a net is an ultra-conservative merger, and the single node and link expansion patterns of Fig. 27 are always available.

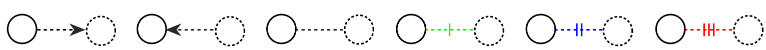


Fig. 27 Single node and link expansion patterns

Thus, nets are extensible by an anonymous node with a single link without logical constraint. However, extra-logical or metalogical constraints may be applied such as a node being specified to be a bottom or top node.

Expansion patterns based on ultra-conservative mergers provide a symmetric contraction pattern, similar to those of the inference patterns, in that the anonymous node may be removed without affecting the logical interpretation of the original net. The addition of a labeled node is conservative if there are no nodes with that label in the net, but generally not conservative if this is not so unless the link is one that may already be inferred.

3.4 Factoring Expansion Patterns

The minimal canonical form (Sect. 3.1) minimizes the number of links whilst keeping the number of nodes constant (treating nodes in a cycle as a single node). There are other forms of restructuring that can further reduce the number of links at the expense of adding additional nodes. For example, one can factor several links with a common tail or head and add an additional node to group them.

Figure 28 left shows a net having 8 nodes and 12 links, with the links at the top having common tails and those at the bottom having common heads. In the center is shown a net with 10 nodes and 10 links where two anonymous nodes have been interpolated to factor out the commonality. The two nets have the same logical interpretation with respect to the nodes in the original net. The import of the anonymous nodes is ultra-conservative, and the original links are implicit, or virtual, in that they may be inferred from the inference patterns.

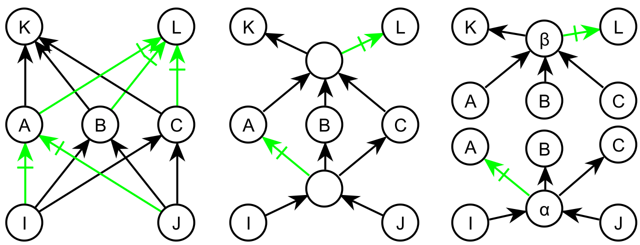


Fig. 28 Example of factoring a net

On the right the factored net has been split into modules and the added nodes have been named. The *interface* nodes between the modules have been duplicated to separate the net into two simpler components.

The factoring pattern is analogous to introducing the middle term of a syllogism, and Vaihinger [104, pp. 212–213] discusses it as an illustration of the introduction in logic of *fictions*, abstract notions that are treated 'as if' they were real. Hertz [67, pp. 248–249] uses this pattern to simplify systems of axioms and notes that the interpolated node may be regarded as an *ideal element*.

Factoring clusters some collection of links in a way that reduces the arrow crossings making common structures more apparent. In an application the resulting sub-structures may well suggest labels for the anonymous nodes that have been created as significant entities in the domain of the application. It generally does not lead to a unique canonical form as there may be several possible factorings having the same logical interpretation as the original net.

The increased perspicuity that can result from factoring is not an artifact of the graphical presentation of the net. If one represents the structures in Fig. 28 as multisets of links then the more modular representation of the net after factoring is also apparent:

$$[\mathrm{I} \nrightarrow \mathrm{A}, \mathrm{I} \rightarrow \mathrm{B}, \mathrm{I} \rightarrow \mathrm{C}, \mathrm{J} \nrightarrow \mathrm{A}, \mathrm{J} \rightarrow \mathrm{B}, \mathrm{J} \rightarrow \mathrm{C}, \mathrm{A} \rightarrow \mathrm{K}, \mathrm{B} \rightarrow \mathrm{K},$$
$$\mathrm{C} \rightarrow \mathrm{K}, \mathrm{A} \nrightarrow \mathrm{L}, \mathrm{B} \nrightarrow \mathrm{L}, \mathrm{C} \nrightarrow \mathrm{L}]$$

$$\big[[\mathrm{I} \rightarrow \alpha, \mathrm{J} \rightarrow \alpha], [\alpha \nrightarrow \mathrm{A}, \alpha \rightarrow \mathrm{B}, \alpha \rightarrow \mathrm{C}],$$
$$[\mathrm{A} \rightarrow \beta, \mathrm{B} \rightarrow \beta, \mathrm{C} \rightarrow \beta], [\beta \rightarrow \mathrm{K}, \beta \nrightarrow \mathrm{L}]\big]$$

$$\big[\big[[\mathrm{I} \rightarrow \alpha, \mathrm{J} \rightarrow \alpha], [\alpha \nrightarrow \mathrm{A}, \alpha \rightarrow \mathrm{B}, \alpha \rightarrow \mathrm{C}]\big],$$
$$\big[[\mathrm{A} \rightarrow \beta, \mathrm{B} \rightarrow \beta, \mathrm{C} \rightarrow \beta], [\beta \rightarrow \mathrm{K}, \beta \nrightarrow \mathrm{L}]\big]\big]$$

The description logic translation makes the *protodisjunctions* and *protoconjunctions* represented graphically by the interpolated nodes more apparent:

$$(\mathrm{I} \sqcup \mathrm{J}) \sqsubseteq \alpha \sqsubseteq (\neg \mathrm{A} \sqcap \mathrm{B} \sqcap \mathrm{C}) \qquad (\mathrm{A} \sqcup \mathrm{B} \sqcup \mathrm{C}) \sqsubseteq \beta \sqsubseteq (\mathrm{K} \sqcap \neg \mathrm{L})$$

From a meaning containment perspective the factoring is innocuous because the logical interpretation of the original net is unchanged; the links between the original nodes are unchanged. The ideal nodes which have been introduced derive content only through their links. Their lack of intrinsic content may be represented logically by making the protoconjunction or protodisjunction structural (Sect. 5), *defining* the ideal nodes in terms of their links. If this is not done and they are named to introduce additional content then extra-logical criteria are involved.

3.5 Existential Status of Added Nodes

Vaihinger's analysis raises the issue of the existential status of added nodes. One may distinguish at least five aspects of the notion of *existence* in the protologic: being a node (*conceptual existence, daseinfrei* [81, p. 51]); being a consistent node (*logical existence*, [49, p. 12]); being specified as an essentially top node representing a maximally generic entity (*categorical existence*); being specified as an essentially bottom node representing a single entity (*singular existence*); and being a bottom node representing an entity within a specified universe of discourse (*situated existence*), such as a state of affairs, phenomenon, event, experience or individual, in some situation, world or time interval specified in logical terms.

The semantification involved in naming ideal elements is not a resemantification of the underlying logical system. It is the extra-logical semantification provided by attaching domain-specific meanings to abstract patterns within that system that represent concepts in the domain.

3.6 Adding a Bottom Node, Probes

A simple but important structure is a *probe*, a bottom node that is added to a net but not regarded as part of it and has an inclusion or exclusion link to some of the nodes constituting the net. Two probes will be termed *distinct* if they have different sets of links.

A probe is *compatible* if it is in maximal canonical form, having all the links that may be inferred from the net. A probe is *admissible* if it is compatible and has no inconsistent links other than those resulting from a link to an inconsistent node; that is, it has no unnecessary inconsistent links. If the net itself is consistent then an admissible probe is also consistent.

Figure 29 illustrates these distinctions for a simple net with a single link between two nodes: P1 is consistent, admissible and compatible; P2 is incompatible because the inferable inconsistent links are not shown; P3 is inconsistent but still incompatible because only one inconsistent link is shown; P4 is compatible but not admissible since the inconsistent links do not derive from an inconsistent node; P5 is admissible because its inconsistent links derive from one in the net.

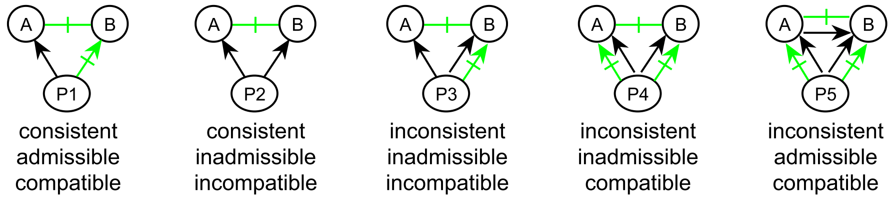


Fig. 29 Examples of different forms of probe

A *full probe* is one with a link to every node in the original net, otherwise it is a *partial probe*. If there are n nodes with distinct labels in the net then there will be 2^n distinct full probes having different combinations of link types, but not all of them may be consistent. A set of all possible distinct admissible probes is said to *saturate* a net.

A probe is not unique and there is no way of specifying it to be so within the protologic so far defined. Unicity can be specified in the metalogic by defining it relative to its links as being the *greatest node* in the preorder that has those links. The general form of this additional constraint is analyzed in Sect. 5.1 and, if pushed down to the protologic, is shown to provide a *structural conjunction* [73, Chap. 13].

3.7 *Admissible Probes Characterizing Possible Links Between Two Nodes*

The four possible probes for the two nodes A and B are:

P0: $A \nleftarrow\ \nrightarrow B$ P1: $A \leftarrow\ \rightarrow B$ P2: $A \nleftarrow\ \rightarrow B$ P3: $A \leftarrow\ \nrightarrow B$

There are 16 possible combinations of consistent probes and these may be used to distinguish between the 16 nets having two nodes linked by all possible types of link or link combination. Figure 30 shows the admissible probes (including the inconsistent probe Pi) for all possible links between A and B. It can be seen that each net has a distinct characteristic set of consistent probes, and hence that such a set may be used to identify the link between the two nodes.

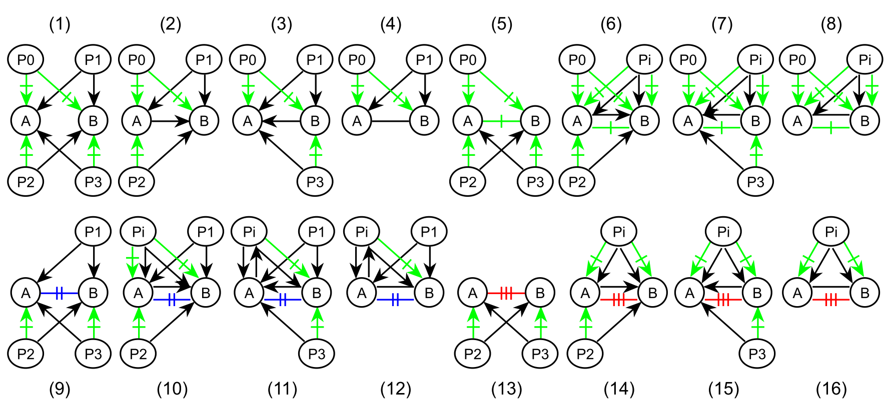


Fig. 30 Admissible probes for sixteen link types

More generally, one can reconstruct the nets from functions, π and $\underline{\pi}$, associating with each node the set of admissible probes that include or exclude that node, respectively. There is an inclusion link from A to B if the set of probes with inclusion links to A, πA, is a subset of that with inclusion links with B, πB, and the set of probes with exclusion links to B, $\underline{\pi}$B, is a subset of that with exclusion links to A, $\underline{\pi}$A, that is $\pi A \subseteq \pi B$ and $\underline{\pi} B \subseteq \underline{\pi} A$. There is an exclusion link from A to B if $\pi A \subseteq \underline{\pi} B$ and $\pi B \subseteq \underline{\pi} A$, and a coexclusion link if $\underline{\pi} A \subseteq \pi B$ and $\underline{\pi} B \subseteq \pi A$.

The conjunction of the two tests is necessary to allow for inconsistent nodes that would otherwise lead to spurious links because a probe has an exclusion link to every exclusion inconsistent node and an inclusion link to every coexclusion inconsistent node. Unless both nodes are inconsistent, the conjunction of the two tests ensures a correct reconstruction.

As shown in Table 1, for the 16 nets of Fig. 30 the links are correctly reconstructed from the relations between the sets of consistent probes. The final example for net (16) is a pseudo inference pattern; all completely inconsistent nodes are both equivalent to and contradictory to one another, playing no useful role in inference processes.

Table 1 Reconstructing links from probe sets

	πA	πB	$\underline{\pi}$A	$\underline{\pi}$B	$\pi A \subseteq \pi B$ $\underline{\pi} B \subseteq \underline{\pi} A$ $A \rightarrow B$	$\pi B \subseteq \pi A$ $\underline{\pi} A \subseteq \underline{\pi} B$ $B \rightarrow A$	$\pi A \subseteq \underline{\pi} B$ $\pi B \subseteq \underline{\pi} A$ $A \dashv B$	$\underline{\pi} A \subseteq \pi B$ $\underline{\pi} B \subseteq \pi A$ $A \dashv\vdash B$
(1)	{1, 3}	{1, 2}	{0, 2}	{0, 3}	false	false	false	false
(2)	{1}	{1, 2}	{0, 2}	{0}	true	false	false	false
(3)	{1, 3}	{1}	{0}	{0, 3}	false	true	false	false
(4)	{1}	{1}	{0}	{0}	true	true	false	false
(5)	{3}	{2}	{0, 2}	{0, 3}	false	false	true	false
(6)	{i}	{i, 2}	{0, i, 2}	{0, i}	true	false	true	false
(7)	{i, 3}	{i}	{0, i}	{0, i, 3}	false	true	true	false
(8)	{i}	{i}	{0, i}	{0, i}	true	true	true	false
(9)	{1, 3}	{1, 2}	{2}	{3}	false	false	false	true
(10)	{i, 1}	{i, 1, 2}	{i, 2}	{i}	true	false	false	true
(11)	{i, 1, 3}	{i, 1}	{i}	{i, 3}	false	true	false	true
(12)	{i, 1}	{i, 1}	{i}	{i}	true	true	false	true
(13)	{3}	{2}	{2}	{3}	false	false	true	true
(14)	{i}	{i, 2}	{i, 2}	{i}	true	false	true	true
(15)	{i, 3}	{i}	{i}	{i, 3}	false	true	true	true
(16)	{i}	{i}	{i}	{i}	true	true	true	true

In Sect. 4.4, it is shown that this technique scales up to nets with any number of nodes and provides extensional semantics for the paraconsistent protologic.

4 Protosemantics

The notion of *semantics* in the logical literature is based on *translations* between a logical system and some other systems, either formal or informal. If one side of the translation is regarded as better founded than the other then it may be used to provide 'foundations' for the other. If not, then both sides may provide insights into aspects of the other. If one system is naturalistic and informal then the more formal system may be seen as providing an *explicatum* of the less formal. Conversely, the naturalistic system may be seen as providing a test of the *adequacy* of its representation by the formal one.

For a universal protologic one would expect there to be a very wide variety of significant translations between it and many other systems, providing protosemantics underlying each of the wide range of approaches to semantics for various logical systems. Some of these have already been mentioned such as the algebraic interpretations in terms of graph theory, residuated relations and preorders. Others have been suggested in terms of potential formalizations of notions of *content* and *context*. The possibility of the reconstruction of nets from probes, bottom nodes that may be regarded as representing states of affairs acting as *protoindividuals*, provides extensional semantics.

This section addresses basic semantic interpretations for the kernel protologic that are common to many logical systems, and later sections examine how they extrapolate to further patterns of inference as they are introduced. Examples of meaningful knowledge structures represented in the protologic are provided in Sect. 7.

4.1 Proof-Theoretic Semantics

Gentzen [57, § II 5.1.3] provided proof-theoretic foundations for logical systems when he extended Hertz's [67] approach to the definition of logical systems by defining natural deduction inference schema comprising paired rules for the introduction and elimination of logical constants. He remarks that the introduction rules could be seen as definitions of the constants and the elimination rules as consequences of those definitions.

The protologic has a foundational proof-theoretic semantics though its specification in terms of *inference patterns*, each of which licenses the addition and deletion of a link (including some to an additional node). The patterns form complementary pairs where the addition of a link defines a feature of a logical constant, such as transitivity, and deletion is in harmony with addition, because the link deleted remains 'virtually' present, always available to be reintroduced.

This leads to two levels of *meaning* for the protologic connectives: at the logical level, it has operational semantics in terms of the inference patterns; at the metalogical level, it has conceptual semantics in terms of phenomena induced by the operations such as the *idempotence*, *transitivity*, *reflexivity*, and *symmetry*.

4.2 Truth-Theoretic Semantics

The graphical representation of inference used so far is to indicate inferred links and nodes with dotted lines. However, as Fig. 26 right illustrates, in larger nets the inferred links may become difficult to discriminate. The graphic language provides an alternative way of displaying the same information by allowing one or more nodes to be marked and then marking the other nodes to show what kind of specified or inferred link the marked nodes have to them.

A vertical mark placed in a node results in all those it includes being marked with a vertical bar, and those it excludes, or exclude it, by a horizontal bar. A horizontal mark placed in a node results in all those nodes that include it being marked by a horizontal bar, and all those that it coexcludes, or coexclude by it, being marked with a vertical bar.

Figure 31 shows a net in which the center node labeled 'A' has different links to five other nodes, and the effect of marking that node with a vertical mark, a horizontal mark, or both.

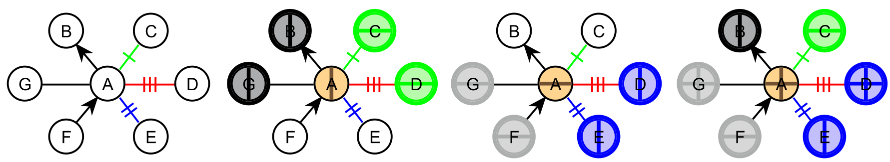


Fig. 31 Propagating marks from one node to linked nodes

The marks are colored similarly to the links to indicate the nature of the source: brown for an asserted mark; black for a vertical bar inferred from an inclusion link; blue for a vertical bar inferred from a coexclusion link; grey for a horizontal bar from an inclusion link; and green for a horizontal bar from an exclusion link (with grey having priority). It is apparent from Fig. 31 that propagation of a vertical mark through negation is through its exclusion component, and a horizontal mark through its coexclusion component.

One can recover a familiar logical vocabulary by terming a vertical mark logically *true* and a horizontal one logically *false*. The inference patterns shown can then be interpreted as: 'if A is true then B and G are true, C and D are false;' 'if A is false then D and E are true, F and G are false;' 'if A is inconsistent (both true and false) then B and E true, C and F are false, D and G inconsistent.'

When a node is marked false all those nodes included in its context are also inferred to be marked false, and this may be viewed as the context determining their *relevance*. One might prefer to say that the node has become *irrelevant* because its context is inapplicable rather than that it is simply false. The grey color of a horizontal bar for falsity distinguishes this type of inference from the other two possibilities of the node being asserted to be false (brown), or being inferred to be so from a node with an exclusion link being inferred false (green).

This explication of logical truth in the protologic is simply an alternative way of representing the inference patterns in a net, a *deflationary* account according to Gupta's definition [63, p. 57]. The complexity of more profound notions of truth [76] beyond those of logic is epistemological, the consideration of issues of the extra-logical justifications for making assertions.

The paraconsistency of the protologic corresponds to it not being bivalent and supporting inference from nodes that are incoherent in being marked both 'true' and 'false.' The inference patterns of Fig. 31 are those of Belnap's [6] *useful four-valued logic* based on 'told true' and 'told false' as independent assertions.

4.3 Intensional Semantics

The set of links having the same tail node provides an *intensional* characterization that node in terms of those of other nodes. A node includes itself, and this link is part of its intension. This corresponds to nodes in the protologic so far developed being *primitive*, constrained by their links to other nodes but not defined by them.

Glashoff [60] has developed an intensional semantics for the categorical syllogistic in which a term, A, is characterized by an *intensional interpretation* comprising a pair of disjoint subsets of a set $\mathcal{O}$, (s(A), $\sigma(A)$), such that the inclusion and exclusion links may be recovered as:

$$\mathrm{A} \rightarrow \mathrm{B} \quad \text{iff } \mathrm{s(A)} \supseteq \mathrm{s(B)} \text{ and } \sigma(\mathrm{A}) \supseteq \sigma(\mathrm{B})$$
$$\mathrm{A} \dashv\vdash \mathrm{B} \quad \text{iff } \mathrm{s(A)} \cap \sigma(\mathrm{B}) \neq \varnothing \text{ or } \sigma(\mathrm{A}) \cap \mathrm{s(B)} \neq \varnothing$$

He does not exemplify particular interpretations of members of $\mathcal{O}$, but it is reasonable to suppose each member is a 'merkmal,' 'mark,' 'feature,' 'property,' 'quality,' 'characteristic' or a similar term used technically or colloquially for a component of content or meaning, or *atom of intensionality*.

The pair of sets may be reduced to a single set by using a syntactic marker, such as an overbar, to distinguish each member of σ(A) from those of s(A), and merging the two sets to represent the *content* of A, c(A). If content is defined to be *inconsistent* if one pair of members has the same label both marked and unmarked, then one can recover the links as:

$$\mathrm{A} \rightarrow \mathrm{B} \quad \text{iff } \mathrm{c(A)} \supseteq \mathrm{c(B)}$$
$$\mathrm{A} \dashv\vdash \mathrm{B} \quad \text{iff } \mathrm{c(A)} \cap \mathrm{c(B)} \text{ is inconsistent}$$

which nicely characterizes content inclusion as meaning containment, and content exclusion as meaning incompatibility.

Glashoff's construction provides the basis for the mereological explication of the inclusion of content in the kernel protologic, complementary to Euler diagrams as a mereological explication of context.

An intensional partition of a net in the kernel protologic (without coexclusion links) is the set of subgraphs each constituted by a single node together with the two sets of nodes at the head of the inclusion and exclusion links of which it is the tail (corresponding to Glassoff's [60, §1.2.2] sets of positive and negative terms). If X is a node in a net,

$$\mathrm{s(X)} = \{\mathrm{Y} : \mathrm{X} \rightarrow \mathrm{Y}\}$$
$$\sigma(\mathrm{X}) = \{\mathrm{Y} : \mathrm{X} \nrightarrow \mathrm{Y}\}$$

defines an intensional partition where all links are included, and the recovery of the net is a simple merger of the node sub-graphs equivalent to

$$\mathrm{A} \rightarrow \mathrm{B} \quad \text{iff } \mathrm{B} \in \mathrm{s(A)}$$
$$\mathrm{A} \dashv\vdash \mathrm{B} \quad \text{iff } \mathrm{B} \in \sigma(\mathrm{A})$$

When the net characterized is in maximal canonical form, Glashoff's recovery of the connectives may be seen as a consequence of the inclusion inference pattern and the exclusion link definition. More complex intensional partitions satisfying these definitions may be developed by following Glashoff and treating node labels as designating non-empty sets which conform with the inclusion and exclusion inference patterns and whose members represent atoms of intensionality.

Such sets may be subsumed within the protologic by representing their members as additional nodes with links that have not been explicitly specified as part of the net. The acceptance that node labels may designate implicit links provides an intensional semantics in terms of the explicit and implicit structure of a net that explicates notions of meaning and content within the protologic without introducing additional constructions. In particular, it recognizes that nodes may have unspecified common content.

Coexclusion adds a third link type to the kernel protologic and the intensional partition of a net may be extended to encompass this by defining a third set, $\omega(\mathrm{X})$, characterizing the coexclusion links of node X such that

$$\omega(\mathrm{X}) = \{\mathrm{Y} : \mathrm{X} \not\twoheadrightarrow \mathrm{Y}\}$$

The net can be recovered by the merger of the node sub-graphs, equivalent to adding the recovery of coexclusion links through

$$\mathrm{A} \not\twoheadrightarrow \mathrm{B} \quad \text{iff } \mathrm{B} \in \omega(\mathrm{A})$$

4.4 *Extensional Semantics*

Extensional semantics in which a logical structure is characterized through the set-theoretic relations between the sets of individuals providing models that are consistent with that structure play a central role in modern logic. They can be studied in the protologic without introducing an additional construct to represent an 'individual' by noting that the only feature required in a suitable entity is that it be impredicable or noninstantiable [61], that is, specified to be a bottom node in the protologic.

Thus, the question becomes one of whether a sufficient collection of bottom nodes may be used to reconstruct the inclusion and exclusion links between the non-bottom nodes in a net. Probes were introduced in Sect. 3.6 as bottom nodes added to investigate a net from outside in order to infer its internal structure as constituted by the links between its nodes.

Extensional semantics for syllogistic systems with essentially the same connectives as inclusion and exclusion links have been developed [34] and the techniques and results are applicable to the protologic. The standard result is that, for a consistent net in the kernel protologic, a set of consistent full probes that *saturates* the net by including all possible distinct examples of such probes may be used to reconstruct the inclusion and exclusion relations in the net [80]. If $\pi\mathrm{A}$ is the set of probes that include node A, then

$$\mathrm{A} \rightarrow \mathrm{B} \quad \text{iff } \pi\mathrm{A} \subseteq \pi\mathrm{B}$$
$$\mathrm{A} \mathrel{-\!\!\!+} \mathrm{B} \quad \text{iff } \pi\mathrm{A} \cap \pi\mathrm{B} = \varnothing$$

The algorithm is basically that of Sect. 3.6 illustrated in Table 1 since $\pi\mathrm{A} \cap \pi\mathrm{B} = \varnothing$ is equivalent to $\pi\mathrm{A} \subseteq \underline{\pi}\mathrm{B}$ if the net is consistent (since $\underline{\pi}$ is then the complement of π relative to the set of admissible probes).

Coexclusion relations may be reconstructed in terms dual to those for exclusion relations: A $\mathrel{-\!\!\!+\!\!\!+}$ B iff $\underline{\pi}A \cap \underline{\pi}B = \varnothing$, which is equivalent to A $\mathrel{-\!\!\!+\!\!\!+}$ B iff $\underline{\pi}A \subseteq \pi B$ if the net is consistent.

If the net is inconsistent, there will be one or nodes where the associated set of consistent nodes is empty and the algorithm would infer that such nodes include and exclude all others, that is, the structure of the links leading to inconsistency would be lost. As noted in Sect. 3.6, the algorithm can be extended to recover the full structure from the set of admissible probes some of which may be inconsistent.

The inference of A → B from the subset relation carries over, but that of A $\mathrel{-\!\!\!+}$ B and A $\mathrel{-\!\!\!+\!\!\!+}$ B from the disjoint relations does not; inconsistency arises because a node has been specified that violates the exclusion or coexclusion link constraints. However, these relations can be inferred through a slight extension to the definitions (noting that $\underline{\pi}$, the set of probes that exclude a node, is no longer the complement of π in inconsistent nets). The extended reconstructions are:

$$
\begin{array}{ll}
A \rightarrow B & \text{iff } \pi A \subseteq \pi B \text{ and } \underline{\pi}B \subseteq \underline{\pi}A \\
A \mathrel{-\!\!\!+} B & \text{iff } \pi A \subseteq \underline{\pi}B \text{ and } \pi B \subseteq \underline{\pi}A \\
A \mathrel{-\!\!\!+\!\!\!+} B & \text{iff } \underline{\pi}A \subseteq \pi B \text{ and } \underline{\pi}B \subseteq \pi A
\end{array}
$$

For consistent nets, where $\underline{\pi}$ is the complement of π, these definitions reduces to

$$
\begin{array}{ll}
A \rightarrow B & \text{iff } \pi A \subseteq \pi B \\
A \mathrel{-\!\!\!+} B & \text{iff } \pi A \cap \pi B = \varnothing \\
A \mathrel{-\!\!\!+\!\!\!+} B & \text{iff } \underline{\pi}A \cap \underline{\pi}B = \varnothing.
\end{array}
$$

As an illustration of the reconstruction algorithm, consider the set of admissible probes (Table 2) for the net of Fig. 16, the last two probes of which are inconsistent, one having a positive link to C and another to D. Table 3 shows the π and $\underline{\pi}$ sets of probes associated with each node.

Table 2 Admissible probes for the net of Fig. 16

	A	B	C	D	E	F
1	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$
2	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\rightarrow$	$\nrightarrow$
3	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\rightarrow$
4	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\rightarrow$	$\rightarrow$
5	$\nrightarrow$	$\rightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\rightarrow$
6	$\nrightarrow$	$\rightarrow$	$\nrightarrow$	$\nrightarrow$	$\rightarrow$	$\rightarrow$
7	$\rightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$
8	$\rightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\nrightarrow$	$\rightarrow$
9	$\nrightarrow \rightarrow$	$\nrightarrow \rightarrow$	$\nrightarrow \rightarrow$	$\nrightarrow$	$\nrightarrow$	$\rightarrow$
10	$\nrightarrow \rightarrow$	$\nrightarrow \rightarrow$	$\nrightarrow \rightarrow$	$\nrightarrow \rightarrow$	$\nrightarrow$	$\rightarrow$

Table 3 Admissible probes associated with each node

	π	$\underline{\pi}$
A	$\{7, 8, 9, 10\}$	$\{1, 2, 3, 4, 5, 6, 9, 10\}$
B	$\{5, 6, 9, 10\}$	$\{1, 2, 3, 4, 7, 8, 9, 10\}$
C	$\{9, 10\}$	$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$
D	$\{10\}$	$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$
E	$\{2, 4, 6\}$	$\{1, 3, 5, 7, 8, 9, 10\}$
F	$\{3, 4, 5, 6, 8, 9, 10\}$	$\{1, 2, 7\}$

The subset relations for π and $\underline{\pi}$ reconstruct the maximal canonical form of the net of Fig. 16 as shown in Table 4.

Table 4 Extensional reconstruction of links for the net of Fig. 16

A	$(-\!\!\!\!+, \mathrm{B}), (\leftarrow -\!\!\!\!+, \mathrm{C}), (\leftarrow -\!\!\!\!+, \mathrm{D}), (-\!\!\!\!+, \mathrm{E})$
B	$(-\!\!\!\!+, \mathrm{A}), (\leftarrow -\!\!\!\!+, \mathrm{C}), (\leftarrow -\!\!\!\!+, \mathrm{D}), (\rightarrow, \mathrm{F})$
C	$(\rightarrow -\!\!\!\!+, \mathrm{A}), (\rightarrow -\!\!\!\!+, \mathrm{B}), (-\!\!\!\!+, \mathrm{C}), (\leftarrow -\!\!\!\!+, \mathrm{D}), (-\!\!\!\!+, \mathrm{E}), (\rightarrow, \mathrm{F})$
D	$(\rightarrow -\!\!\!\!+, \mathrm{A}), (\rightarrow -\!\!\!\!+, \mathrm{B}), (\rightarrow -\!\!\!\!+, \mathrm{C}), (-\!\!\!\!+, \mathrm{D}), (-\!\!\!\!+, \mathrm{E}), (\rightarrow, \mathrm{F})$
E	$(-\!\!\!\!+, \mathrm{A}), (-\!\!\!\!+, \mathrm{C}), (-\!\!\!\!+, \mathrm{D})$
F	$(\leftarrow, \mathrm{B}), (\leftarrow, \mathrm{C}), (\leftarrow, \mathrm{D})$

5 Structural Connectives

The focus of prior sections has been on logical structures based on the basic inclusion and exclusion connectives in order to demonstrate the representational and inference capabilities of the foundational logical connectives. This contrasts with specifications of logical systems that take structural connectives as primitive.

Structural connectives may be introduced in the protologic as *ideal elements* [7, 71] represented by the addition of nodes that are extremal relative to a set of links in the order relation associated with $\rightarrow$. Koslow [73, 74] uses this approach to generalize Gentzen's introduction and elimination rules for logical connectives. He considers a general implication relation subject to the normal axioms for a consequence operator [14] and defines the usual structural connectives as extremal structures in the order relation of implication/consequence.

5.1 Structural Conjunction

The inclusion and exclusion links out of a node may be viewed as specifying a *proto-conjunction* (Sect. 2.4) in that, if a probe includes that node then, from the inclusion and exclusion inference patterns, it will also include or exclude all the linked nodes. In the terminology of Sect. 4.2, if the node is marked *true* all the nodes that it includes will be marked *true* and those it excludes *false*.

However, the converse does not apply. One cannot infer that the node is true if some or all of the nodes to which it has outgoing links are marked appropriately. One can, however, define a constraint that specializes a node to be a full structural conjunction through an inference pattern that specifies this.

Definition A node is the *conjunction* of a set of outgoing links iff any node with the same links includes it.

Figure 32 is a graphical representation of this definition. The outgoing links defining the conjunctive node at the bottom left are distinguished by conjunctive variants of the inclusion and exclusion arrows that have a heavier head. Any node, such as that at the bottom right, that has the same set of outgoing links may be inferred to include the conjunctive node.

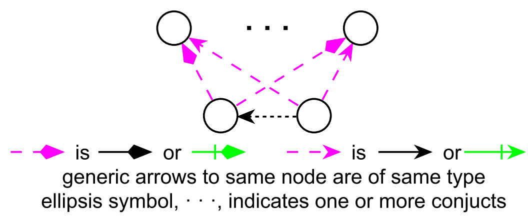


Fig. 32 Structural conjunction inference pattern

If the bottom node on the right is regarded as a probe then the inference pattern implies that any probe that marks the nodes to which the conjunctive node has inclusion links *true*, and those to which is has exclusion links *false*, also marks node the conjunctive node as *true*. These are the converse truth conditions defining a full structural conjunction.

The inference pattern of Fig. 32 parallels Koslow's [73, §13.1] approach to the definition of conjunction but takes into account exclusion as well as inclusion links. It specifies that, in terms of the order relation associated with →, the conjunctive node is *maximal* among all nodes having the outgoing links specified by the heavier arrows.

Arbitrary conjunctive nodes may be freely added as ideal elements to any net but they will not always be consistent; for example, if inclusion links are specified to nodes between which there is an exclusion link. In terms of content containment, these ideal nodes are *defined* by their conjunctive links and have no other content. However, if additional links are added such that they include or exclude additional content then the nodes become *rules* imposing meaning constraints upon the net that reflect extra-logical normative or empirical contingencies.

5.2 Structural Disjunction

The inclusion and coexclusion links into a node may be viewed as specifying a *protodisjunction* (Sect. 2.4) in that, if a probe excludes that node then, from the exclusion and coexclusion inference patterns, it will also exclude or include all the linked nodes. In the terminology of Sect. 4.2, if the node is marked *false* all the nodes that it includes will be marked *false* and those it coexcludes *true*.

However, the converse does not apply. One cannot infer that the node is false if all the nodes from which it has incoming links are marked appropriately. One can, however, define a constraint that specializes the node itself through an inference pattern that allows this to be inferred.

Definition A node is the *disjunction* of a set of incoming links iff it includes any node with the same links.

Figure 33 is a graphical representation of this definition. The incoming links defining the disjunctive node at the top left are distinguished by disjunctive variants of the inclusion and coexclusion arrows that have a double head. Any node, such as that on at the top right, that has the same set of incoming links may be inferred to be included in the disjunctive node.

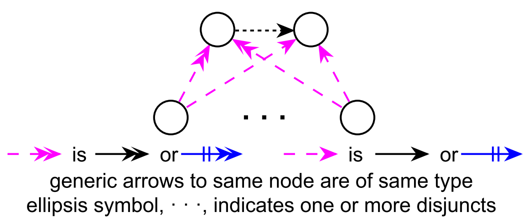


Fig. 33 Structural disjunction inference pattern

If a probe has exclusion links to the tail nodes of the inclusion links, and inclusion links to the tail nodes of the coexclusion links, defining the disjunctive node then it may, without inconsistency, have an exclusion link to a an arbitrary node with inclusion links from these nodes such as that at the top right. Such a node also has an inclusion link from the disjunctive node so that an exclusion link may be inferred from the probe to the disjunctive node. That is, if the tail nodes of the incoming links defining the disjunctive node are all marked to propagate false then the disjunctive node is false. This is the converse condition defining a full structural disjunction.

The definition of structural disjunction specifies that, in terms of the order relation associated with →, the disjunctive node is *minimal* among all nodes having the incoming links specified by the double arrows. This differs from Koslow's [73, §13.1] definition which specifies the maximality of a different construction, but the two definitions are equivalent.

An important application of structural disjunction is to represent *abduction* as inference to those *abducibles* consistent with a state of affairs [47]. If the disjuncts in Fig. 33 are bottom nodes representing possible states of affairs, such as Millikan's [82] *substance templets*, asserting the disjunction true for a particular state of affairs represents the abductive hypothesis that one or more of those templets must fit that state. Disjunctive inference then derives the consequences of this hypothesis.

5.2.1 Material Implication and Other Sentential Calculus Formulae

The inclusion link, $\rightarrow$, is appropriate to represent *entailment*, necessary, rather than material, implication. It specifies relations between terms prescribing their proper usage, conventions of language rather contingencies of a world. In addition, the implicative conditional is not represented in a form subject to inference; it may only be *used* in the protologic, not *mentioned*.

The disjunctive inference pattern may be used to represent a *material implication* as a logical structure in which the implication is itself represented as a node subject to inference. Figure 34 left illustrates the simplest case. If node M is asserted true then at least one of its disjuncts may be inferred to be true. Thus $\neg A \vee B$ is true, the classical interpretation of $A \supset B$ as material implication. However, the conditioning on M means that what is represented is $M \supset (A \supset B)$.

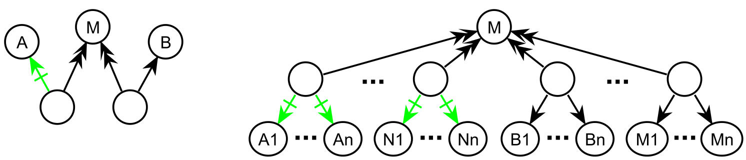


Fig. 34 Representing material implication in the protologic

This non-truth-functional representation in terms of implication rather than equivalence avoids the so-called ‘paradoxes of material implication’ [42, §2.3]. The only situation in which the truth value of M is determined by those of A and B is that it is inferred false if A is asserted true and B false, that is when the material implication does not hold.

Figure 34 right shows a more general form of material implication where there are multiple disjuncts of nodes with multiple outgoing links. This represents:

$$M \supset \big(((A1 \vee \cdots \vee An) \wedge \cdots \wedge (N1 \vee \cdots \vee Nn)\big) \supset \big((B1 \wedge \cdots \wedge Bn) \vee \cdots \vee (M1 \wedge \cdots \wedge Mn)\big)\big)$$

Thus a disjunctive node together with inclusion/exclusion links may be used to represent a material implication between complex clauses with premises in conjunctive normal form and conclusions in disjunctive normal form.

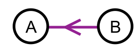
6 Defeasible Inference

Paraconsistency in a logical system is desirable in that an inconsistency is localized and does not undermine reasoning outside that locale. However, such containment alone is inadequate to support reasoning *within* the logical system about the sources and consequences of the inconsistency. These aspects of the inconsistency remain at the metalogical level.

Reasoning about inconsistency can be pushed down to the logical level by introducing nodes representing potential inconsistencies as anomalies, and playing no role unless nodes representing normal states of affairs do not apply. This may be implemented through a *preference relation* between possible states of affairs represented as disjuncts of a disjunction such that, when that disjunction is asserted true, a disjunct is inferred to be false if any preferred disjunct is not false.

If an anomalous state of affairs is represented as a disjunct that is less preferred than the other disjuncts, it will be inferred false and play no role unless all the other disjuncts are false (non-applicable). Then, being the only non-false node, it will be inferred to be true (applicable).

6.1 *Preferential Defeasible Inference*



The preference relation allows a partial order to be specified between the nodes of a structural disjunction. It is represented textually by the symbol, $<$, with A $<$ B specifying that node B is preferred to node A, and in a net by a preference link as shown on the right.

The associated inference patterns are shown in Fig. 35: left, transitivity of preference; right, exclusion if a preferred node is included or not linked (that is, it is not excluded).

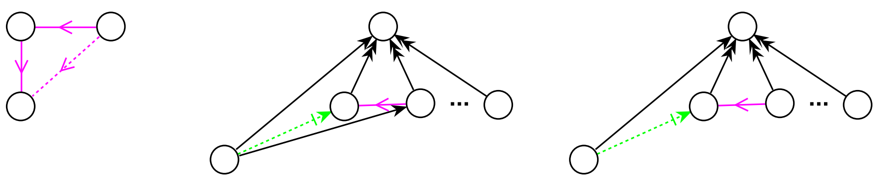


Fig. 35 Preference inference patterns

In terms of contextual semantics, if the context of a node is part of that of a node that has sub-contexts then it is assumed to be disjoint with that of a less preferred sub-context if it is not disjoint with that a more preferred one.

These inference patterns are only applicable once all other inference patterns have been applied. They, and ensuing inferences, are *defeasible* because they are based on the preference relation. The source of the defeasibility is apparent in the inference pattern on the far right which licenses an inference to be made when the truth value of a node is unspecified.

6.1.1 Normal Defaults

The preference inference patterns may be used to represent the *normal defaults* [77, 94] used in applications of nonmonotonic reasoning, that a proposition is true unless there are grounds for it being false. Figure 36 from left to right shows nets representing, from left to right: :A, that A is normally true; A:B, that A normally materially implies B; A:B and ¬B:¬A, Lehman's [77, §6] preferred representation of a default material conditional that also incorporates default *modus tollendo tollens*; and a more complex combination of defaults such that: A is normally true; B false; A normally implies C; but A and B normally implies not C.

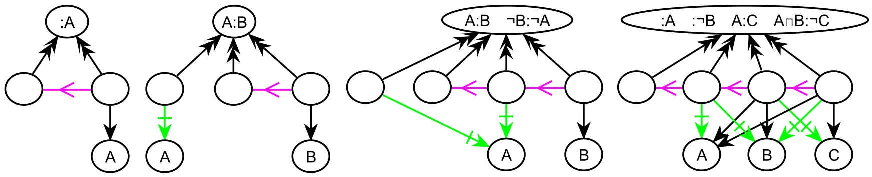


Fig. 36 Normal defaults

Figure 37 from left to right illustrates default reasoning: (1) when no truth values are asserted for A and B, no inferences are made about them (the nodes inferred false by the preference inference patterns are colored mauve); (2) when A is asserted true, B is defeasibly inferred true; (3) if B is also asserted false, this is accepted and an anomaly is inferred (anomalous node is colored red as an inference from the disjunctive definition); (4) if B is asserted false, A is defeasibly inferred false (if the inference is overridden then (3) results again).

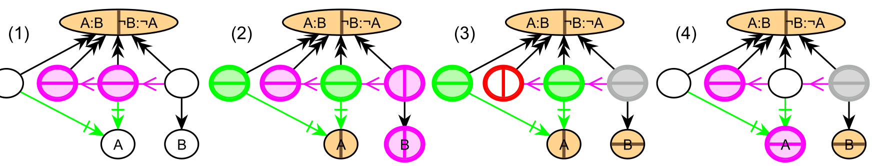


Fig. 37 Inference from normal defaults

This implementation of defeasible reasoning encompasses the test cases in the literature collected in [77], and more complex ones such as Stalnaker's [100] combinations of classical and default inference. It also encompasses abductive, or case-based, reasoning where the disjuncts represent states of affairs templets to be fitted to a particular state of affairs.

7 Some Illustrative Applications

A major objective of this article has been to address Béziau's [11] proposal that universal logic should be able to support other fields of knowledge to build *the right logic for the right situation*. This section provides some brief illustrative example of how the protologics developed above can be applied in other disciplines.

7.1 Merging Biological Taxonomies

Reasoning about taxonomies is important to biological science, for example, in analyses of the consequences of merging taxonomies of the same species from different sources. There are literature studies of the use of powerful theorem provers to reason about taxonomic merging [103], but the same results may be obtained in the kernel protologic in a more perspicuous form.

Figure 38 presents an example of two simple taxonomies, their proposed merging through the articulations shown, and queries arising about the outcome.

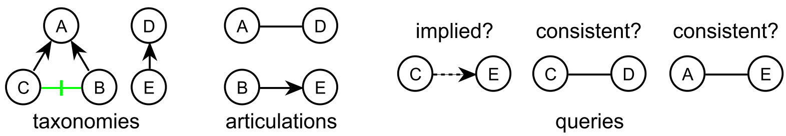


Fig. 38 Some taxonomy articulation queries [103, p. 198]

The three nets on the left of Fig. 39 show the inferred links corresponding to the to the three queries: there is no inferred link from C to E; equivalencing C and D makes the net inconsistent; but doing so for A and E does not.

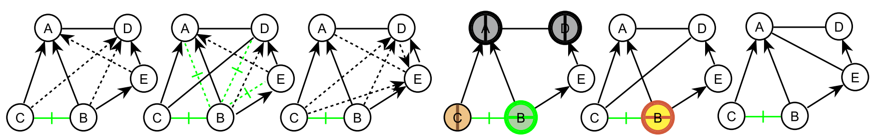


Fig. 39 Inferred responses to the taxonomy articulation queries

The nets on the right provide the same information in terms of true–false marking: if C is marked true there is no propagation to E; if C is equivalenced to D then node D is marked false in all consistent truth-assignments, that is, inconsistent; if A is equivalenced to E there are no such inferences.

The study cited represents the taxonomic data by similar diagrams but treats inferences from them as informal and resolves the queries through logical symbolism and a first-order logic theorem prover. However, the diagrams themselves can be used to provide formally well-founded responses to the queries.

7.2 *An Expert System*

Figure 40 shows a net implementing a simple expert system [32] that prescribes hard or soft contact lenses for a client having the attributes the four determinables shown at the top left leading to one of three prescriptions shown at the top right.

The solution is specified in terms of conjunctions specifying rules and exceptions: a client whose tear production is normal should be prescribed a hard lens if astigmatic and a soft lens if not; however, there is an exception to the soft prescription if the patient is presbyopic and myopic, and to the hard if hypermetropic and old. The inferences are shown when the state of affairs representing a client at the bottom of Fig. 40 is activated by marking it true.

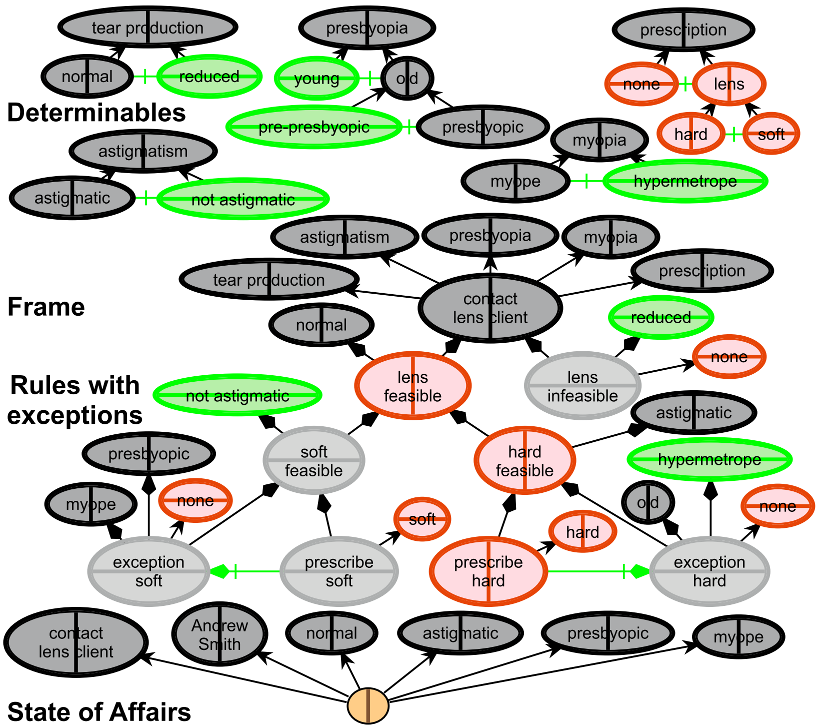


Fig. 40 Inference from rules with exceptions in an expert system

7.3 Defining an Art Object

The problem of defining what it is to be an *art object* has been a continuing issue for the philosophy of art community since ancient times [30, 35]. It is generally accepted that a definition in terms of necessary and sufficient conditions is not appropriate [106], and a variety of other classificatory techniques have been investigated such as *family resemblances* [28] and *cluster concepts* [55].

Gaut [56] has defined the notion of an *art object* as a cluster concept, and one may argue [51] that the definitional aspects of his model lie in the frame constituted by the eleven determinables that he uses to characterize art objects. Figure 41 illustrates the use of the kernel protologic and default reasoning to analyze aspects of his classificatory structure.

At the top are shown: left, four of Gaut's determinables; right, the frame for an art object that represents the relevance of these determinables. Beneath these are shown templets for five types of art of art object including, on the left, that for an 'ideal art object' that includes all the positive determinates. These together capture the essence of Gaut's analysis of art as a cluster concept.

The defaults at the bottom capture expectations when one is told that some entity is an art object: that, by default, it should have the characteristics of an ideal art object. The markings illustrate the reasoning when it is asserted that something is an art object but not created with the intent that it be one, such as Duchamps' *readymades* [36]. The default

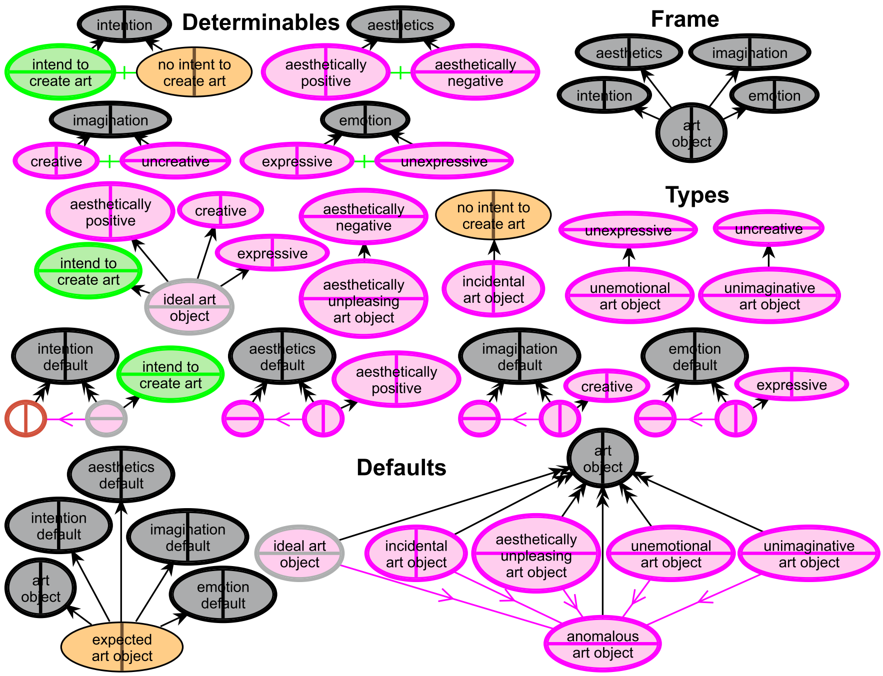


Fig. 41 Default reasoning about art objects

inferences are that it has the other characteristics of an 'ideal art object' but is only an 'incidental art object.'

A similar structure of defaults may be used to represent logical aspects of fiction such as Ryan's [98] *principle of minimal departure* from expectations corresponding to those of a world appropriate to the fictional genre.

7.4 Representing States of Affairs, Individuals and Relations

The notions of *states of affairs* [2, 72] and their constituent *individuals* and *relations* are important to most applications of logical systems but their appropriate definition raises many deep philosophical issues [43] leading to a range of different approaches to their representation. A universal logic framework needs to be able to represent these adequately whilst making no commitment to any in particular.

The protologic as developed above does not require the definition of individuals as primitive entities for any of the technical purposes for which they are normally required. Set-theoretic extensional semantics have been provided in Sect. 4.4 based on *probes* simply defined as bottom nodes in Sect. 3.6. Bottom nodes satisfy Gracia's impredicability criterion for an individual and could be termed *protoindividuals* but they need not satisfy Strawson's [102, p. 214] criteria that they be distinct and reference some other entity. Extensional semantics are a technical feature intrinsic to the protologic involving no other connotations or denotations.

Figure 42 is a simple illustration of the representation of a state of affairs referencing an individual based on the well-known example that Ramsey [92] provided to counter Russell's [97] argument that there was an essential logical distinction between universals and particulars. The anonymous bottom node represents a templet for a state of affairs in which an individual named 'Socrates' exhibits behavior that can be characterized as 'wise.'

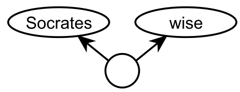


Fig. 42 Representation of an individual

The nature of the Russell–Ramsey debate may be analyzed in terms of Fig. 42. Ramsey's position corresponds to the anonymous node being symmetrically 'some Socrates' entity and 'some wise' entity. However, in its role of representing reference to the state of an individual, it may be distinguished from the other nodes in being an impredicative bottom node corresponding to the focus in Russell's article on singulars as states of affairs.

The bottom node satisfies Gracia's criterion for an individual if it is intended to represent a particular situation. If it is intended as a generic *substance templet* [82] that applies to particular states of Socrates then it does not, but those instances of which it is predicated will satisfy the criterion. That is, impredicability is not an absolute distinction but rather a pragmatic feature of use.

Strawson's criterion of reidentification captures the essence of the intent behind the use of proper names, to track what is imputed to be the 'same' individual in different situations. That is, reference to an individual is intended to be a *rigid designator* and this is also a pragmatic rather than logical distinction [5].

The bottom node will satisfy Strawson's criterion of distinctness if the term 'Socrates' is always used to identify what is intended to be a single individual, but, if it is applicable to several individuals, then additional identifying terms may be required to track the intended individual. That is, proper names may require contextual information to disambiguate their intended application [87, p. 295].

As Engel [43, §3.3.2] notes, Ramsey's argument concerns the logical, rather than ontological, status of particulars and universals. The distinction between a node being used to reference an individual and to ascribe a property to one still needs to be representable, but not necessarily as a logical primitive.

Castañeda [31] models individuation as requiring an *individuator operator* providing indexical access. One can model the individuator operator as one that infers which bottom nodes would make the specified indexical nodes true, and this is implemented in the protologic through the form of abductive inference discussed in Sect. 5.2. That is, given the assignment of truth values to the indexical nodes, one considers the disjunction of all the bottom nodes representing states of affairs that have not been inferred to be false and draws the inferences common to them.

This corresponds to Perry's [88, 93] analysis of proper names as functioning to index *mental files* of information about an individual, it may be that a combination of name and context is required to disambiguate the file access. One may generalize this to any combination of terms, including names, being used to provide a context sufficiently constrained to identify a particular individual, as in Orilia's [84] *contextual descriptivist* account of singular reference.

7.4.1 Venus as Morning Star and Evening Star

To illustrate the representation of states of affairs, Fig. 43 represents the two perceived states of Venus that Frege [48] used to exemplify his distinction between *sense* and *reference*. The planet appears as a bright celestial object near the rising or setting sun at dawn or dusk. Early Greek astronomers regarded the phenomena as arising from different celestial bodies and named them differently [105, n. 23]. As shown, abductive inference over the bottom nodes marks the common features as defeasibly true.

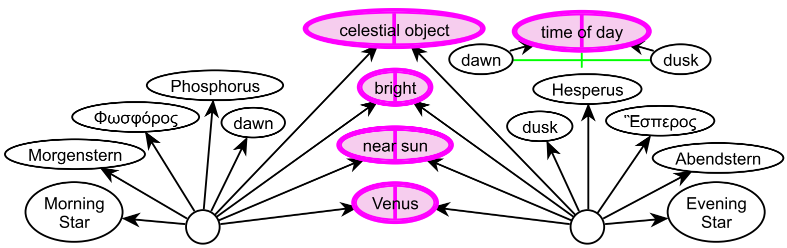


Fig. 43 Representation of perceived states of Venus

The net illustrates the major issues. Many names may be used to reference the same physical entity, and someone may not be aware that they do not refer to different ones; removing the node for Venus illustrates this. Someone may know some of the names and that they have a common referent but not know those names in another language; removing some state-specific names illustrates this.

Abductive inference after marking the nodes representing dawn or dusk as true leads to all the morning star names and Venus being marked true but none of the evening star names, and *vice versa*. That is, it is reasonable to have different names that reference the same entity in different contexts.

7.4.2 Representing Relations and Structural Universals

While the representation of individuals requires no additional constructs in the protologic, that of relations requires distinguishing the states of affair that are related. One version implemented in the protologic is based on Orilia's [85] representation of *neutral relations* [46] in terms of *onto-thematic roles*.

As shown in Fig. 44, relations are represented by nodes having *relational links* represented by arrows having distinctive tails that serve to group relational links of the same type. There may be any number of types of relational links. Nodes with only one type of relational link represent symmetric relations, and those with two or more asymmetric relations. On the left are inference patterns through which relational links interact with inclusion links.

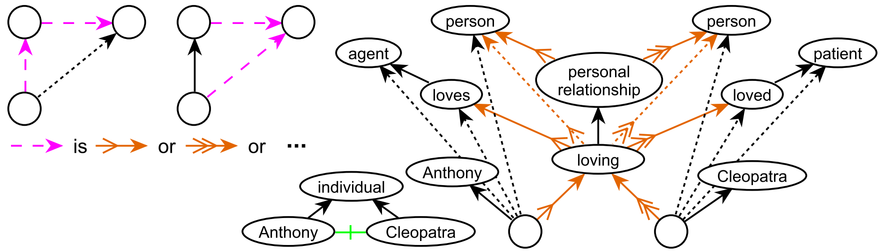


Fig. 44 Relation inference patterns and example relation

At the bottom center 'Anthony' and 'Cleopatra' are typed as two distinct individuals, with the node 'individual' making Castañeda's individuator operator available as needed. On the right, the state of affairs where there is a loving relationship between them is represented as a personal relationship having two thematic roles, 'loves' as *agent* and 'loved' as *patient*. Inferences from the patterns are shown with dotted lines.

This representation avoids many of the known issues of representing relations [46, 54, 85]. There is a single relation and the notion of *converse* is a matter of its linguistic expression. Subordination of relations is represented. The relational arrows have no ontological role but serve only to link the constraints on each component of a relation. The components are distinguished by their thematic roles, not their positions or link types. The roles themselves are simply nodes that may be part of a net representing their logical interrelationships. The generic linguistic roles mark those that include them as relational roles, providing similar functionality to an individuator operator. Additional types of relational arrows may be added to represent other roles such as 'instrument.'

Figure 45 exemplifies Armstrong's [2] *structural universals* in terms of a templet of the states of affairs and their relationships in a water molecule.

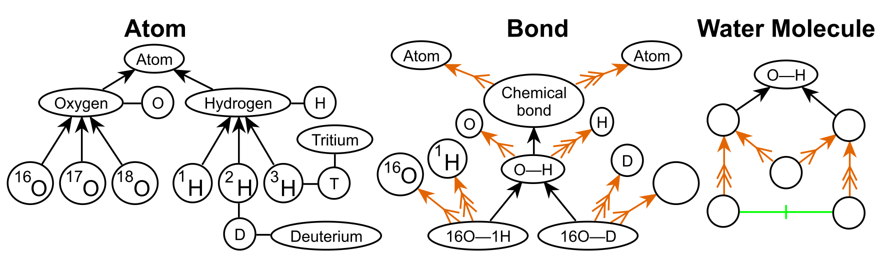


Fig. 45 Relational structure of a water molecule

From left to right: the determinable, atom, is represented with oxygen, hydrogen and their isotopes as determinates; the hydroxyl chemical bond is represented together with some of its forms; the structure constituting a water molecule is represented as a structural universal having two hydroxyl bonds with a common oxygen atom and two distinct hydrogen atoms [86].

8 Conclusions

Over the past two decades, Béziau's [9] notion of *universal logic* as an integrative conceptual framework for all logical systems that presupposes no particular axioms has provided a focal point for a wide range of historical and ongoing studies of foundational issues in logic and the nature and relationship of logical systems, as did Birkhoff's [23] notion of *universal algebra* for foundational studies of algebras.

Béziau also envisioned [11] that the universal logic conceptual framework could help many fields of knowledge build *the right logic for the right situation*, noting that for some disciplines *mathematical abstract conceptualization* is more appropriate than *symbolic formalization*. This idea has so far had less impact although Béziau has exemplified it in his own research, for example, on the diverse interpretations of the square of oppositions and the special issue of *Logica Universalis* on 'Is logic universal?' [16] raising a range of cross-disciplinary issues [52].

One should not expect a significant impact of the conceptual framework of universal logic on other disciplines to develop rapidly because the diffusion of knowledge and techniques between different fields of knowledge is known to be intrinsically slow [96]. In particular, the symbolic formalism of the normal expositions of logic may be an impediment to diffusion and it might help expedite wider adoption if the conceptual framework of universal logic could be presented in a way that is more accessible whilst remaining formally sound.

This article has provided an alternative formalism that avoids mathematical symbolism by extending Hertz's [67] original graphic presentation of the principles of logical deduction to encompass the sequent calculus that he, Gentzen [57], and others developed. The approach is based on past research on the formal representation of semantic networks as a practical tool applicable to the development, representation and application of knowledge structures in diverse applications [50, 51].

One side-effect of the techniques used to avoid mathematical symbolism is that the knowledge structures used in various disciplines become represented in a graphical form that often mimics the informal diagrams used in those disciplines yet allows formal logical inference patterns to be applied.

The conceptual and computational tools illustrated in this article have proved useful in several disciplines, and should, hopefully, contribute to the wider adoption of Béziau's visionary ideas. Universal logic can provide a *mathematical abstract conceptualization* that is logically sound, readily comprehensible and practically useful in the clarification, communication and evaluation of ideas, theories and associated controversies in many fields of knowledge. Facilitating this is a significant and challenging research objective for the universal logic community.

References

1. Aristotle: Metaphysics (trans. William David Ross). In: Barnes, J. (ed.) The Complete Works of Aristotle, pp. 1552–1728. Princeton University Press, Princeton (1984)
2. Armstrong, D.M.: A World of States of Affairs. Cambridge Studies in Philosophy. Cambridge University Press, Cambridge (1997)
3. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.: The Description Logic Handbook. Cambridge University Press, Cambridge (2003)

4. Bang-Jensen, J., Gutin, G.: Digraphs: Theory, Algorithms, and Applications. Springer, Berlin (2009)
5. Baumann, P.: Are proper names rigid designators? Axiomathes **20**, 333–346 (2010)
6. Belnap, N.D.: A useful four-valued logic. In: Dunn, J.M., Epstein, G., Wolf, R.G. (eds.) Modern Uses of Multiple-Valued Logic, pp. 8–37. Reidel, Dordrecht (1977)
7. Bernays, P.: Probleme der theoretischen Logik. Unterr.bl. Math. Nat.wiss. **33**, 369–377 (1927)
8. Bernays, P.: Die Philosphie der Mathematik und die Hilbertische Beweistheorie. Bl. dtsch. Philos. **4**, 326–367 (1930–1931)
9. Béziau, J.-Y.: Universal logic. In: Childers, T., Majer, O. (eds.) Logica'94: Proceedings of the 8th International Symposium, pp. 73–93. Czech Academy of Sciences, Prague (1994)
10. Béziau, J.-Y.: La véritable portée du théorème de Lindenbaum–Asser. Log. Anal. **42**(167–168), 341–359 (1999)
11. Béziau, J.-Y.: From paraconsistent logic to universal logic. Sorites **12**, 5–32 (2001)
12. Béziau, J.-Y.: Sequents and bivaluations. Log. Anal. **44**(176), 373–394 (2001)
13. Béziau, J.-Y.: 13 questions about universal logic. Bull. Sect. Log. **35**(2–3), 133–150 (2006)
14. Béziau, J.-Y.: From consequence operator to universal logic: a survey of general abstract logic. In: Béziau, J.-Y. (ed.) Logica Universalis, pp. 3–17. Birkhäuser, Basel (2007)
15. Béziau, J.-Y.: Logica Universalis: Towards a General Theory of Logic. Birkhäuser, Basel (2007)
16. Béziau, J.-Y.: Preface: is logic universal? Log. Univers. **4**(2), 161–162 (2010)
17. Béziau, J.-Y.: The power of the hexagon. Log. Univers. **6**(1–2), 1–43 (2012)
18. Béziau, J.-Y.: Universal Logic: An Anthology. Birkhäuser, Basel (2012)
19. Béziau, J.-Y.: The metalogical hexagon of opposition. Argumentos **10**, 111–122 (2013)
20. Béziau, J.-Y., Carnielli, W.A., Gabbay, D.M. (eds.): Handbook of Paraconsistency. College publications, London (2007)
21. Béziau, J.-Y., Dale, J.: Around and Beyond the Square of Opposition. Studies in Universal Logic. Birkhäuser, Basel (2012)
22. Béziau, J.-Y., Payette, G.: The Square of Opposition: A General Framework for Cognition. Lang, Bern (2012)
23. Birkhoff, G.: Universal algebra. In: Proceedings of the First Canadian Mathematical Congress 1945, pp. 310–326. University of Toronto Press, Toronto (1946)
24. Blanché, R.: Sur l'opposition des concepts. Theoria **19**(3), 89–130 (1953)
25. Bourbaki, N.: The architecture of mathematics. Am. Math. Mon. **57**(4), 221–232 (1950)
26. Brady, R.T.: Universal Logic. CSLI, Stanford (2006)
27. Brandom, R.: Between Saying and Doing: Towards an Analytic Pragmatism. Oxford University Press, Oxford (2008)
28. Campbell, K.: Family resemblance predicates. Am. Philos. Q. **2**(3), 238–244 (1965)
29. Carnielli, W.A., Coniglio, M.E., D'Ottaviano, I.M.L.: New dimensions on translations between logics. Log. Univers. **3**, 1–18 (2009)
30. Carroll, N.: Theories of Art Today. University of Wisconsin Press, Madison (2000)
31. Castañeda, H.-N.: Individuation and non-identity: a new look. Am. Philos. Q. **12**(2), 131–140 (1975)
32. Cendrowska, J.: An algorithm for inducing modular rules. Int. J. Man-Mach. Stud. **27**(4), 349–370 (1987)
33. Clark, J.T.: Conventional Logic and Modern Logic: A Prelude to Transition. Philosophical Studies of the American Catholic Philosophical Association. Woodstock College Press, Woodstock (1952)
34. Corcoran, J.: Completeness of an ancient logic. J. Symb. Log. **37**(4), 696–702 (1972)
35. Davies, S.: Definitions of Art. Cornell University Press, Ithaca (1991)
36. de Duve, T.: Pictorial Nominalism on Marcel Duchamp's Passage from Painting to the Readymade. University of Minnesota Press, Minneapolis (1991)
37. De Morgan, A.: On the symbols of logic, the theory of the syllogism, and in particular of the copula, and the application of the theory of probabilities to questions of argument and authority. Trans. Camb. Philos. Soc. **9**, 79–127 (1856)
38. Dunn, J.M.: Generalized ortho negation. In: Wansing, H. (ed.) Negation: A Notion in Focus, pp. 3–26. de Gruyter, Berlin (1996)
39. Dutilh Novaes, C.: The different ways in which logic is (said to be) formal. Hist. Philos. Logic **32**(4), 303–332 (2011)

40. Dutilh Novaes, C.: Formal Languages in Logic: A Philosophical and Cognitive Analysis. Cambridge University Press, Cambridge (2012)
41. Eberle, R.A.: Goodman on likeness and difference of meaning. Erkenntnis **12**, 3–16 (1978)
42. Edgington, D.: On conditionals. Mind **104**(414), 235–329 (1995)
43. Engel, P.: The Norm of Truth: An Introduction to the Philosophy of Logic. University of Toronto Press, Toronto (1991)
44. Fahlman, S.E.: NETL, a System for Representing and Using Real-World Knowledge. MIT Press, Cambridge (1979)
45. Ferejohn, M.T.: Plato and Aristotle on negative predication and semantic fragmentation. Arch. Gesch. Philos. **71**(3), 257–282 (1989)
46. Fine, K.: Neutral relations. Philos. Rev. **109**(1), 1–33 (2000)
47. Flach, P.A., Kakas, A.C.: Abductive and inductive reasoning: background and issues. In: Flach, P.A., Kakas, A.C. (eds.) Abduction and Induction: Essays on Their Relation and Integration, pp. 1–27. Kluwer, Dordrecht (2000)
48. Frege, G.: Über Sinn und Bedeutung. Z. Philos. philos. Krit. **100**, 25–50 (1892)
49. Frege, G.: Gottlob Freges Briefwechsel mit D. Hilbert, E. Husserl, B. Russell, sowie ausgewählte Einzelbriefe Freges. Meiner, Hamburg (1980)
50. Gaines, B.R.: An interactive visual language for term subsumption visual languages. In: IJCAI91: Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, pp. 817–823. Morgan Kaufmann, San Mateo (1991)
51. Gaines, B.R.: Designing visual languages for description logics. J. Log. Lang. Inf. **18**(2), 217–250 (2009)
52. Gaines, B.R.: Human rationality challenges universal logic. Log. Univers. **4**(2), 163–205 (2010)
53. Galatos, N., Jipsen, P., Kowalsi, T., Ono, H.: Residuated Lattices: An Algebraic Glimpse at Substructural Logics. Elsevier, Amsterdam (2007)
54. Gaskin, R., Hill, D.J.: On neutral relations. Dialectica **66**(1), 167–186 (2012)
55. Gasking, D.: Clusters. Australas. J. Philos. **38**(1), 1–36 (1960)
56. Gaut, B.: "Art" as a cluster concept. In: Carroll, N. (ed.) Theories of Art Today, pp. 25–44. University of Wisconsin Press, Madison (2000)
57. Gentzen, G.: Untersuchungen über das logische Schließen. I. Math. Z. **39**(1), 176–210 (1935)
58. Gentzen, G.: Über das verhältnis zwischen intuitionistischer und klassischer Arithmetik. Arch. Math. Log. Grundl.forsch. **16**(3–4), 119–132 (1974) (orig. 1933)
59. Girard, J.-Y.: Linear logic. Theor. Comput. Sci. **50**(1), 1–101 (1987)
60. Glashoff, K.: An intensional Leibniz semantics for Aristotelian logic. Rev. Symb. Log. **3**(02), 262–272 (2010)
61. Gracia, J.J.E.: Individuals as instances. Rev. Metaphys. **37**(1), 37–59 (1983)
62. Grau, B.C., Horrocks, I., Kazakov, Y., Sattler, U.: Modular reuse of ontologies: theory and practice. J. Artif. Intell. Res. **31**, 273–318 (2008)
63. Gupta, A.: A critique of deflationism. Philos. Top. **21**(2), 57–81 (1993)
64. Harary, F., Gupta, G.: Dynamic graph models. Math. Comput. Model. **25**(7), 79–87 (1997)
65. Harary, F., Norman, R.Z., Cartwright, D.: Structural Models: An Introduction to the Theory of Directed Graphs. Wiley, Chichester (1965)
66. Heeffer, A.: Algebraic symbolism as a conceptual barrier in learning mathematics. In: Proceedings 6th East Asia Regional Conference on Mathematics Education (EARCOME 6), Phuket, Thailand, vol. 6, pp. 2–10 (2013)
67. Hertz, P.: Über Axiomensysteme für beliebige Satzsysteme I: Sätze ersten Grades. (Über die Axiomensysteme von der kleinsten Satzzahl und den Begriff des Idealen Elementes). Math. Ann. **87**(3–4), 246–269 (1922). (English trans. Javier Legris [18, 11–29])
68. Hilbert, D.: Über die Theorie der algebraischen Formen. Math. Ann. **36**(4), 473–534 (1890)
69. Hilbert, D.: Grundlagen der Geometrie. Teubner, Leipzig (1899)
70. Hilbert, D.: Axiomatisches Denken. Math. Ann. **78**(1–4), 405–415 (1917)
71. Hilbert, D.: Über das Unendliche. Math. Ann. **95**(1), 161–190 (1926)
72. Jago, M.: Setting the facts straight. J. Philos. Log. **40**(1), 33–54 (2011)
73. Koslow, A.: A Structuralist Theory of Logic. Cambridge University Press, Cambridge (1992)
74. Koslow, A.: Structuralist logic: implications, inferences, and consequences. Log. Univers. **1**(1), 167–181 (2007)

75. Krämer, S.: Operative Schriften als Geistestechnik. Zur Vorgeschichte der Informatik. In: Schefe, P., Hastedt, H., Keilg, Y. (eds.) Informatik und Philosophie, pp. 69–84. Spektrum, Heidelberg (1993)
76. Künne, W.: Conceptions of Truth. Clarendon Press, Oxford (2003)
77. Lehmann, D.: Another perspective on default reasoning. Ann. Math. Artif. Intell. **15**, 61–82 (1995)
78. Londey, D., Johanson, C.: The Logic of Apuleius: Including a Complete Latin Text and English Translation of the Peri Hermeneias of Apuleius of Madaura. Brill, Leiden (1987)
79. Lorenzen, P.: Protologik. Kant-Stud. **47**(1–4), 350 (1956)
80. Martin, J.M.: Aristotle's natural deduction reconsidered. Hist. Philos. Logic **18**(1), 1–15 (1997)
81. Meinong, A.: Über die Stellung der Gegenstandstheorie im System der Wissenschaften. Voigtländer, Leipzig (1907)
82. Millikan, R.G.: On Clear and Confused Ideas: An Essay About Substance Concepts. Cambridge University Press, Cambridge (2000)
83. Noether, M.: Paul Gordan. Math. Ann. **75**, 1–41 (1914)
84. Orilia, F.: Singular Reference: A Descriptivist Perspective. Springer, Dordrecht (2010)
85. Orilia, F.: Relational order and onto-thematic roles. Metaphysica **12**(1), 1–18 (2011)
86. Pagès, J.: Structural universals and formal relations. Synthese **131**(2), 215–221 (2002)
87. Pelczar, M., Rainsbury, J.: The indexical character of names. Synthese **114**(2), 293–317 (1998)
88. Perry, J.: A problem about continued belief. Pac. Philos. Q. **61**, 317–332 (1980)
89. Plato: The Statesman. Harvard University Press, Cambridge (1952) (Harold North Fowler, trans.)
90. Priest, G.: Sense, entailment and modus ponens. J. Philos. Log. **9**, 415–435 (1980)
91. Priest, G., Routley, R.: A preliminary history of paraconsistent and dialethic approaches. In: Priest, G., Routley, R., Norman, J., Arruda, A.I. (eds.) Paraconsistent Logic: Essays on the Inconsistent, pp. 3–75. Philosophia, München (1989)
92. Ramsey, F.P.: Universals. Mind **34**(136), 401–417 (1925)
93. Récanati, F.: Mental Files. Oxford University Press, Oxford (2013)
94. Reiter, R.: A logic for default reasoning. Artif. Intell. **13**(1–2), 81–132 (1980)
95. Resnik, M.D.: Mathematics as a Science of Patterns. Clarendon Press, Oxford (1997)
96. Rogers, E.M.: Diffusion of Innovations, 5th edn. Free Press, New York (2003)
97. Russell, B.: On the relations of universals and particulars. Proc. Aristot. Soc. **12**, 1–24 (1911)
98. Ryan, M.-L.: Fiction, non-factuals, and the principle of minimal departure. Poetics **9**(4), 403–422 (1980)
99. Schönfinkel, M.: Über die Bausteine der mathematischen Logik. Math. Ann. **92**(3–4), 305–316 (1924)
100. Stalnaker, R.: What is a nonmonotonic consequence relation? Fundam. Inform. **21**, 7–21 (1994)
101. Straßburger, L.: What is a logic, and what is a proof? In: Béziau, J.-Y. (ed.) Logica Universalis, pp. 135–152. Birkhäuser, Basel (2007)
102. Strawson, P.F.: Strawson, Individuals, an Essay in Descriptive Metaphysics. Methuen, London (1959)
103. Thau, D., Ludäscher, B.: Reasoning about taxonomies in first-order logic. Ecol. Inform. **2**(3), 195–209 (2007)
104. Vaihinger, H.: Die Philosophie des Als Ob: System der theoretischen, praktischen und religiösen Fiktionen der Menschheit auf Grund eines idealistischen Positivismus. Meiner, Leipzig (1922)
105. van der Sluijs, M.A.: Who are the "attendants of Helios"? J. Am. Orient. Soc. **129**(2), 169–177 (2009)
106. Weitz, M.: The Opening Mind: A Philosophical Study of Humanistic Concepts. University of Chicago Press, Chicago (1977)

B.R. Gaines (✉)
University of Victoria, Victoria, Canada
e-mail: gaines@uvic.ca

B.R. Gaines
University of Calgary, Calgary, Canada
e-mail: gaines@ucalgary.ca

Hexagonal Logic of the Field $\mathbb{F}_8$ as a Boolean Logic with Three Involutive Modalities

René Guitart

Abstract We consider the Post–Malcev full iterative algebra $\mathbb{P}_8$ of all functions of all finite arities on a set $\underline{8}$ with 8 elements, e.g. on the Galois field $\mathbb{F}_8$. We prove that $\mathbb{P}_8$ is generated by the logical operations of a canonical boolean structure on $\mathbb{F}_8 = \mathbb{F}_2^3$, plus three involutive $\mathbb{F}_2$-linear transvections A, B, C, related by circular relations and generating the group $\mathrm{GL}_3(\mathbb{F}_2)$. It is known that $\mathrm{GL}_3(\mathbb{F}_2) = \mathrm{PSL}_2(\mathbb{F}_7) = \mathrm{G}_{168}$ is the unique simple group of order 168, which is the group of automorphisms of the Fano plane. Also we obtain that $\mathbb{P}_8$ is generated by its boolean logic plus the three cross product operations $R\times$, $S\times$, $I\times$.

Especially, our result could be understood as a *hexagonal logic*, a natural setting to study the logic of functions on a hexagon; precisely, it is a *hexagonal presentation of the logic of functions on a cube with a selected diagonal*.

Keywords Hexagon of opposition · Borromean object · Specular logic · Boolean algebra · Modality · Many-valued logics · Finite fields · Fano plane

Mathematics Subject Classification (2000) 00A06 · 03B45 · 03B50 · 03G05 · 06Exx · 06E25 · 06E30 · 11Txx

1 Introduction: How and Why to Generate Functions on $\underline{8}$?

Our concrete result is formulated in elementary terms in Theorem 8.1. Its geometrical formulation is in Proposition 7.7, and other variations are given by Propositions 7.4 (for A, B, C), 6.11 and 6.12 (for r, s, i and r^{-1}, s^{-1}, s^{-1}), and 5.6 and 5.7 (for the calculus of avatars); these variations have their own interest *and* are steps in the proof of Theorem 8.1.

The question is to generate the system of all the functions of all finite arities on a set $\underline{8} = \{\underline{0}, \underline{1}, \underline{2}, \underline{3}, \underline{4}, \underline{5}, \underline{6}, \underline{7}\}$ with 8 elements.

The method is to emphasize the hexagon character of the situation, at the level of the data ($\mathbb{F}_8$), as well as at the level of functions elements of ($\mathbb{P}_8$); then we explore arithmetic and geometry around the Fano plane with 7 points and the Galois field $\mathbb{F}_8$ with 8 elements; and we mainly look at polynomial equations, computation of cross products, and the geometry of the cube $\mathbb{F}_2^3$. The advantage of the method is to obtain a solution in which a hexagonal symmetry is assumed among elements: the hexagonal symmetry among the data is reproduced among the solutions.

In friendly homage to Jean-Yves Béziau, on the occasion of his 50th birthday.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_8

In Sect. 2, the beginning of the paper provides motivations, in the context of splitting paradoxes in analysis of discourses. We explain what is meant by a hexagon seen as a cube with a selected diagonal, and hexagonal functions.

In Sect. 3, we represent $\underline{8}$ by the Galois' field $\mathbb{F}_8$, we develop the analysis of the field $\mathbb{F}_8$ and its hexagonal generation, in relation with the Fano plane. We introduce the Galois' field structure and arithmetical calculus with the roots R, S, I of $X^3 + X^2 + 1 = 0$.

In Sect. 4 the geometry of $\mathbb{F}_8$ as an $\mathbb{F}_2$-vectorial space is developed, with scalar, cross and mixed products, its canonical boolean logic, and the galoisian data of the Frobenius map.

In Sect. 5, with the help of results of Sect. 3 and 4, we obtain the representation of $\mathbb{P}_8$ by sums of conjunctions of avatars. This would be enough in order to solve some logical paradoxes, in the style of [8] with $\mathbb{F}_4$.

In Sect. 6, we prove that $\mathbb{F}_8$ admits a unique strictly auto-dual basis, showing that the canonical logic on $\mathbb{F}_8$ is really canonical. With the result of Sect. 5 and the analysis of $\mathrm{GL}_3(\mathbb{F}_2)$ from [7], this is used to prove the R, S, I generation of $\mathbb{P}_8$.

In Sect. 7, we arrive at the A, B, C presentation by three involutive transvections (plus the canonical boolean structure), and we end by the $R^\times, S^\times, I^\times$ presentation.

In the conclusive Sect. 8, we stress the decoration of a hexagon by functions of $\mathbb{P}_8$, and—as announced at the beginning of this introduction—we reach Theorem 8.1.

Warning on notations We will see in Proposition 5.5 that any function $f : \mathbb{F}_8 \to \mathbb{F}_8$ could be represented by a polynomial P, and especially it is true for linear functions. But a linear f is also representable by matrices M relative to the canonical basis κ, and in principle we must not confuse f, P and M. This is important for product and composition. If we write NM, we mean the *composition* of the matrix N applied to M, and gf means $g \circ f$ the composition of g applied to f, and especially f^2 means $f \circ f$, the composition of f with f, defined by $f^2(u) = f(f(u))$. But by QP we mean the *product* of Q and P, P^2 means the square of P defined by $P^2(u) = (P(u))^2$. The reader will make the distinction according to the context.

Convention If necessary, in order to avoid confusion, if M is a matrix of f and P a polynomial of the same f, we introduce

$$P_f = P = \underline{M}, \qquad M_f = M = \hat{P},$$

in such a way that $\underline{N}\,\underline{M}$ is the product polynomial QP, and $\hat{Q}\hat{P}$ is the composition of matrices NM. This is useful, especially in Sects. 6 and 7.

2 $\mathbb{F}_8$ and the Oriented Hexagon

2.1 The Square and the Splitting of Paradoxes in $\mathbb{F}_4$

In [6–8], we have introduced the use of a Galois field of characteristic 2 as a logical tool to analyze paradoxical sentences. It was a continuation of [5], as an arithmetical version of the idea of point of view and *speculation*. It was related to the picture of a hexagon and ideas of *borromean object* and *boolean manifold*. We gave explicit results in the 2-dim case $\mathbb{F}_4$, using two facts: the shape of the system of various boolean structures on $\mathbb{F}_4$—alias the system of $\mathbb{F}_2$-basis on $\mathbb{F}_4$—and the existence of the Frobenius map $(-)^2 : \mathbb{F}_4 \to \mathbb{F}_4 : x \mapsto x^2$, which, in this logical context, is considered as a new type of modality, expressing a *galoisian indiscernibility*.

The field $\mathbb{F}_4$ with four elements is the arithmetic of a square with a fixed oriented diagonal $0 \to 1$, and two other indiscernible corners α and ω, $\mathbb{F}_4 = \{0, 1, \alpha, \omega\}$, with $\alpha + \beta = \alpha\beta = 1$, in such a way that the polynomial $X^2 + X + 1$ (paradoxical, i.e. without roots in $\mathbb{F}_2$) splits in $\mathbb{F}_4$. The logic of the square, i.e. the organization of the system $\mathbb{P}_4$ of functions $\mathbb{F}_4^k \to \mathbb{F}_4$, was presented in [8]; in this 'logic', several hexagonal pictures appear, which allow understanding this logic as a 'borromean logic', or as a 'boolean manifold'. This logic was shown to be useful for the splitting of paradoxes and analysis of paradoxical sentences.

2.2 The Cube and the Question of the Hexagonal Symmetry of $\mathbb{P}_8$

2.2.1 Splitting of Paradoxes in $\mathbb{F}_8$

Now we consider the field $\mathbb{F}_8$. This case is interesting because $\mathbb{F}_8 = \mathbb{F}_{2^3}$ is the smallest case of a 3-dim space over a field (it is the first *cubic field*), and consequently in this space we can imagine knots, borromeam links, etc., analogous to curves in $\mathbb{R}^3$.

The field $\mathbb{F}_8$ with 8 elements is the arithmetic supported by a cube with a fixed diagonal $0 \to 1$, that is to say, by an oriented hexagon (see Sect. 2.3). It will be generated by 3 elements R, S, I, roots of $X^3 + X^2 + 1$ (paradoxical, i.e. without roots in $\mathbb{F}_2$). We denote it by (cf. Proposition 3.6):

$$\mathbb{F}_8 = \{0, 1, R, S, I, R', S', I'\}.$$

Of course, the work done with $\mathbb{F}_4$ could be repeated with the Galois' field $\mathbb{F}_8$; proceeding in this way again, we would split some paradoxes, observe various hexagonal pictures, and borromean objects, etc.

2.2.2 Hexagonal Symmetry of $\mathbb{P}_8$

But here this question of splitting paradoxes (2.2.1) is not our direct aim. Rather we would like to study *the logic* of *the hexagon itself*, i.e. the system $\mathbb{P}_6$ of functions on 6 elements with a special central symmetry. For that we have to understand the structure of $\mathbb{P}_8$ and mainly its *hexagonal symmetry*.

A first tool is the existence on $\mathbb{F}_8$ of a *canonical* boolean logic, in which *false* $= 0$ and *true* $= 1$; so the use of these boolean functions in the presentation of the system $\mathbb{P}_8$ implicitly emphasizes the axis 0–1 in the cube, and reduces symmetries of the cube to symmetries of the hexagon (see Sect. 2.3).

As $6 = 8 - 2 = 2^3 - 2^1$, it would be possible to consider a *hexagonal function* (see Sect. 2.3) as a function on $\mathbb{F}_8$ which never takes the value 0, and with value 1 if and only if one of the variables is 0 or 1; so $\mathbb{P}_6 \subset \mathbb{P}_8 = \mathbb{P}(\mathbb{F}_8)$, and we can describe $\mathbb{P}_6$ with the help of the nice structure of $\mathbb{F}_8$ which is both a field *and* a boolean algebra.

A second tool is the group of colineations of the Fano plane. The Fano plane is almost the same thing that a hexagon (cf. Remarks 2.2 and 3.5). On the one hand, this group is $\mathrm{PSL}_2(\mathbb{F}_7)$, and, on the other hand, it is $\mathrm{GL}_3(\mathbb{F}_2)$ (cf. [4]). The ternary symmetry of the hexagon will be taken into account by generating this group with the 3 transvections A, B, C with circular relations (Proposition 7.1).

2.3 *Hexagonal Functions*

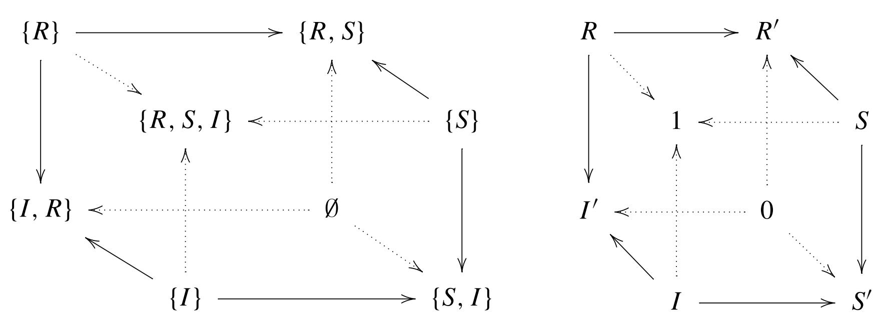
2.3.1 The Hexagon in Logic

We want to increase the value of a mixed logico-geometrical discipline, in which a given diagram stipulates a system of possible points of view acting on propositions as modalities. It is easy to deduce this plan from the Sesmat's approach. The very special case of a hexagonal diagram [2, 3, 16] is very fundamental, and a nice convincing modern re-examination and development is given by Jean-Yves Béziau [1]. In [7, 8], turning around the idea of hexagon, we prefer to work with the notion of *borromean object*; but the first aspect of a borromean object is its hexagonal appearance, and the second aspect is the exactness of diagonals, in some sense equivalent to the description of opposite corners as complements (as in a logical hexagon of oppositions).

2.3.2 From the Cube to the Hexagon, and to Hexagonal Functions

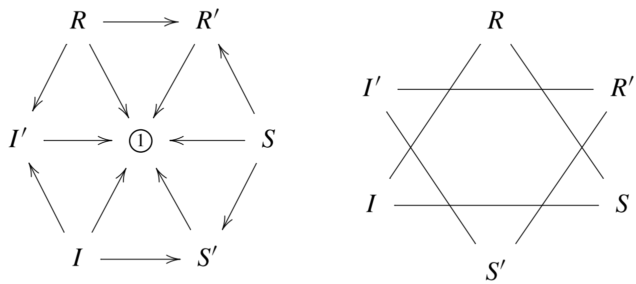
Jacques Lacan introduced the *RSI logic* for psychoanalysis, with $R =$ Real, $S =$ Symbolic, $I =$ Imaginary, explaining that discourses work under these three modalities linked in a borromean way, as the Father, the Son and the Holly Spirit, the three hypostases of God in Christian Trinity, are linked. In homage to Lacan, here we will use these three letters in our presentations and computations.

The geometrical cube K_3 is the shape of a boolean logical cube $\mathcal{P}(\{R, S, I\})$, drawn as follows, with $R' = \{R, S\}$, $S' = \{S, I\}$, $I' = \{I, R\}$, $1 = \{R, S, I\}$, $0 = \emptyset$.



So our basic picture will be the view of the cube orthogonal to its axis 0–1, i.e. the *hexagon*:

Definition 2.1 The *hexagon* is the picture



The set of elements of this hexagon which are different from 0 and 1 is denoted by $\mathcal{H} = \{R, S, I, R', S', I'\}$, and often—abusively—the hexagon itself also will be denoted by $\mathcal{H}$.

A *hexagonal function* is a function $h : \mathcal{H}^k \to \mathcal{H}$, with $k \in \mathbb{N}$, and the set of these functions is denoted by $\mathbb{P}(\mathcal{H}) = \bigcup_{n\in\mathbb{N}} \mathcal{H}^{\mathcal{H}^k}$; it is the Post–Malcev algebra $\mathbb{P}_6$ on 6 elements, as defined in [13] or [10].

Remark 2.2 1. Let us remark that, with these notations, the *complement* R^c of R is not R' but

$$R^c = S', \qquad S^c = I', \qquad I^c = R'.$$

In fact, in our future computations, R' will be the *inverse* of R:

$$R' = R^{-1}, \qquad S' = S^{-1}, \qquad I' = I^{-1}.$$

2. In Proposition 3.4, we will recover this hexagon as the Fano plane, and in fact automorphisms of this $\mathbb{F}_2$-projective plane will play a decisive part in our analysis.

Remark 2.3 As explained in Sect. 2.2.2, we can consider

$$\mathbb{P}(\mathcal{H}) = \bigcup_{n\in\mathbb{N}} \mathcal{H}^{\mathcal{H}^k} \subset \bigcup_{n\in\mathbb{N}} \mathbb{F}_8^{\mathbb{F}_8{}^k} = \mathbb{P}(\mathbb{F}_8),$$

and so $\mathbb{P}_6$ appears as a sub-Post–Malcev algebra of $\mathbb{P}_8$. But what we do here is only to give a presentation of $\mathbb{P}_8$ which 'respects' the presence of $\mathbb{P}_6$ in $\mathbb{P}_8$; we don't claim that our generators A, B, C or the canonical logical functions are in fact in $\mathbb{P}_6$; and we don't claim that our 'hexagonal functions', i.e. functions on the hexagon (elements of $\mathbb{P}_6$), are the functions respecting the 'geometry of the hexagon'. We just say that if someone wants to describe these last 'geometrical' functions among the functions of $\mathbb{P}_6$ then, after a precise determination of what he means by the 'geometry of the hexagon', he could use our presentation of $\mathbb{P}_6$ in $\mathbb{P}_8$ with the logical functions and the A, B, C as a natural setting because this A, B, C presentation shows the hexagonal symmetry of $\mathbb{P}_8$.

The use of elements external to $\mathbb{P}_6$ in order to present elements of $\mathbb{P}_6$ becomes natural if we want a maximal logical component in our analysis because the cardinality of the set $\mathcal{H}$ is not a power of 2, but rather it could be embedded in a boolean algebra of cardinality 8, as well as the set of its elements which are different from false and truth. In this way, the hexagonal symmetry of $\mathbb{P}_8$ could act on $\mathbb{P}_6$.

But in fact, reversing the problem, we precisely claim that any 'geometry of the hexagon' will be determined by the data of any sub-algebra of $\mathbb{P}_6$ compatible with the symmetry of the A, B, C presentation (cf. Proposition 7.4). So our exclusive purpose will be to clearly understand the hexagonal presentation of $\mathbb{P}_8$, which mixes arithmetical, geometrical and logical aspects.

3 Circular Presentation of Arithmetic in $\mathbb{F}_8$ with R, S, I

In this section, we describe the finite field with 8 elements, and we emphasize its presentation as an $\mathbb{F}_2$-algebra with a circular presentation. This presentation will be useful in the sequel, for the study of geometry and of logic. In this section, we use this presentation to study third degree equations in $\mathbb{F}_8$, reducible to linear and second degree, and so with geometrical interpretations. We obtain the reduced third degree paradoxes in $\mathbb{F}_8$.

3.1 Splitting $X^3 + X^2 + 1 = 0$ and $X^3 + X + 1 = 0$

Proposition 3.1 *Let $\mathbb{F}_2 = (\{0, 1\}, +, \times)$ be the field of integers modulo 2. The polynomials $X^3 + X^2 + 1$ and $X^3 + X + 1$ are reciprocal, i.e. exchangeable by $X \mapsto X^{-1}$, the roots of the first are linearly independent, but the roots of the second are linearly dependent. They are the two irreducible polynomials of degree 3 over $\mathbb{F}_2$, the fields $\mathbb{F}_2[X]/(X^3 + X^2 + 1)$ and $\mathbb{F}_2[X]/(X^3 + X + 1)$ are isomorphic, with 8 elements. Both are realization of a splitting field of $X^8 - X$ over $\mathbb{F}_2$, i.e. the smallest extension of $\mathbb{F}_2$ in which $X^8 - X$ splits in a product of linear factors*:

$$X^8 - X = X(X-1)\big(X^3 + X^2 + 1\big)\big(X^3 + X + 1\big).$$

Furthermore the 'squaring' Frobenius map $(-)^2 : x \mapsto x^2$ is $\mathbb{F}_2$-linear, i.e.

$$(x + y)^2 = x^2 + y^2.$$

Proof The sum of the roots of the second polynomial is 0; and the sum of the roots of the first is 1. The squaring $(-)^2$ is linear because we are in characteristic 2. Clearly, $X^3 + X^2 + 1$ and $X^3 + X + 1$ have no root in $\mathbb{F}_2$, and, as they are of degree 3, they are irreducible, and the quotients rings are fields. The decomposition for $X^8 - X$ is proved by expansion of the right side. Then, for every element x in the field (whatever copy of it is chosen) we have $x^8 - x = 0$, and for elements $x \neq 0$ we have $x^7 = 1$, i.e. $x^{-1} = x^6$. The determination up to isomorphism of finite fields, one exactly for each cardinal p^n, for p prime and n integer, is known since 1893 [14]: it works here for $p = 2$, $n = 3$, and $2^3 = 8$. For recent classic manuals, see [12, 15]. □

Proposition 3.2 *In the field with 8 elements defined in Proposition 3.1, an element a is a root of $X^3 + X^2 + 1 = 0$ if and only if $b = a + 1$ is a root of $X^3 + X + 1 = 0$; then every element is a power of a and a power of b, with the correspondence*:

$$a = b^3, \qquad a^2 = b^6, \qquad a^3 = b^2, \qquad a^4 = b^5, \qquad a^5 = b, \qquad a^6 = b^4;$$

$$b = a^5, \qquad b^2 = a^3, \qquad b^3 = a, \qquad b^4 = a^6, \qquad b^5 = a^4, \qquad b^6 = a^2.$$

The roots of $X^3 + X^2 + 1$ *are* a, a^2 *and* a^4, *the powers of* a^2, *which are exchangeable by the powers of* $(-)^2$, *the roots of* $X^3 + X + 1$ *are* b, b^2 *and* b^4, *the powers of* b^2, *which are exchangeable by the powers of* $(-)^2$.

Proof It is easily checked. At first, $(a+1)^3+(a+1)+1=0$. From $a^3=a^2+1$ we obtain $a^4=a^2+a+1$, $a^5=a+1$, $a^6=a^2+a$, $a^7=1$, and with $b=a+1$ we obtain $b=a^5$, $b^2=a^3$, $b^3=a$, etc., as announced. For the distribution of roots, we verify for a^2 that $0=a^4(a^3+a^2+1)=1+a^6+a^4=(a^2)^3+(a^2)^2+1$, and for a^4 with $a^5=b=a+1$ we get $0=a^5+a+1=a^{12}+a^8+1=(a^4)^3+(a^4)^2+1$. □

Proposition 3.3 *We consider* $\mathbb{F}_2[Y]/(Y^3+Y^2+1)=\mathbb{F}_2(\alpha)$, *with* α *an abstract root of* Y^3+Y^2+1, *e.g.* $\alpha=Y$, *and* $\mathbb{F}_2[Z]/(Z^3+Z+1)=\mathbb{F}_2(\beta)$ *with* β *an abstract root of* Z^3+Z+1, *e.g.* $\beta=Z$. *An explicit isomorphism* A *between these two fields and its inverse* B *are given by*

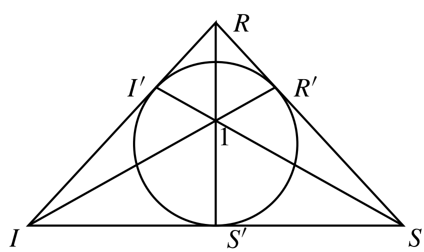
$$A(\alpha^n)=\beta^{3n}, \qquad B(\beta^n)=\alpha^{5n}.$$

Proof It results from Proposition 3.2, where both $\mathbb{F}_2(\alpha)$ and $\mathbb{F}_2(\beta)$ are realized as "the" splitting field of $X^8-X=0$, and as $\mathbb{F}_2(a)$ and $\mathbb{F}_2(b)$, with $\alpha=a$, and $\beta=b$: at this level, the maps A and B become the identity map. As $b=a+1$ and $a=b+1$, we get $B(\beta)=\alpha+1$, $A(\alpha)=\beta+1$, etc. □

3.2 *The Circular Presentation by* R, S, I

Proposition 3.4 *The projective plane over* $\mathbb{F}_2$*—the Fano plane—is constructible with* 7 *points named* $1, R, S, I, R', S', I'$, *as on the picture, where* 7 *'straight lines' are drawn named*:

$$R^\perp=\{S,S',I\}, \qquad S^\perp=\{I,I',R\}, \qquad I^\perp=\{R,R',S\},$$
$$R'^\perp=\{R',1,I\}, \qquad S'^\perp=\{S',1,R\}, \qquad I'^\perp=\{I',1,S\},$$
$$1^\perp=\{R',S',I'\}.$$



Remark 3.5 This plane is important for us, especially through its group of projective automorphisms, which is $\mathrm{G}_{168}=\mathrm{PSL}_2(\mathbb{F}_7)=\mathrm{GL}_3(\mathbb{F}_2)$. The involutions A, B, C that we will introduce in Proposition 7.1 are geometrical maps (colineations) on this plane.

Proposition 3.6 *A concrete model isomorphic to the fields in Proposition 3.1 is*

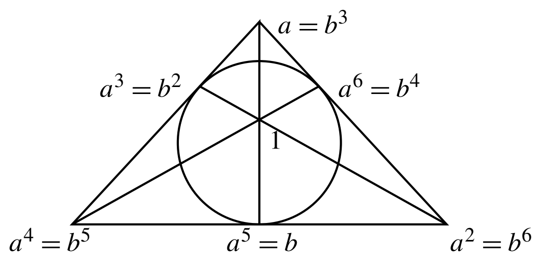
$$\mathbb{F}_8 = \{0, 1, R, S, I, R', S', I'\},$$

where the addition is described by $x + x = 0$, when $x \in \mathbb{F}_8$, and by lines in the Fano plane (Proposition 3.4): if $x, y \in \mathbb{F}_8$, $x \neq y$, then $x + y = z$ is equivalent to $\{x, y, z\}$ being one of the 7 lines in the Fano plane. So we can use the Fano plane as an addition table in $\mathbb{F}_8$, and simultaneously, by the use of exponent, as a multiplication table. This allows us to grasp $\mathbb{F}_8$ in a flash.

Proof The field $\mathbb{F}_8$ can be described in multiplicative terms, as in Proposition 3.2, with the a or the b terms. Multiplying $a^3 + a^2 + 1 = 0$ by powers of a, we obtain:

$$a^2 + a^5 + a^4 = 0; \qquad a^4 + a^3 + a = 0; \qquad a + a^6 + a^2 = 0;$$
$$a^6 + 1 + a^4 = 0; \qquad a^5 + 1 + a = 0; \qquad a^3 + 1 + a^2 = 0;$$
$$a^6 + a^5 + a^3 = 0.$$

This allows us to replace: $R = a$, $S = a^2$, $I = a^4$, $R' = a^6$, $S' = a^5$, $I' = a^3$.



□

Proposition 3.7 *As a field of characteristic 2, the field $\mathbb{F}_8$ could be presented in a circular symmetric way—and we shall name this* the R, S, I presentation—*with the following relations among the elements*:

$$RSI = 1, \qquad RS + SI + IR = 0, \qquad R + S + I = 1,$$
$$R^2 = S, \qquad S^2 = I, \qquad I^2 = R.$$

Proof The first line of equations expresses that R, S, I are the three roots of $X^3 + X^2 + 1$, and the second line explains how the Frobenius squaring transforms them. We have to prove that this completely determines all calculations in $\mathbb{F}_8$ as a field of characteristic 2. We introduce

$$R' := I + 1 = R + S, \qquad S' := R + 1 = S + I, \qquad I' := S + 1 = I + R.$$

We have $R(S + I) = SI$, $R^2(S + I) = RSI = 1$, $S(S + I) = 1$, and so $SI = I + 1$, $SI = R'$. And then $RSI = 1$ means that $RR' = 1$. Similarly, we have $IR = S'$, $SS' = 1$, $RS = I'$, $II' = 1$. Also $S + I = S' = R + 1$, etc.

Products $u \cdot u' := uu'$ and sums $u + u'$ are given by the tables:

$\cdot$	R	S	I	R'	S'	I'
R	S	I'	S'	1	R'	I
S	I'	I	R'	R	1	S'
I	S'	R'	R	I'	S	1
R'	1	R	I'	S'	I	S
S'	R'	1	S	I	I'	R
I'	I	S'	1	S	R	R'

$+$	R	S	I	R'	S'	I'	1
R	0	R'	I'	S	R'	I	S'
S	R'	0	S'	R	I	1	I'
I	I'	S'	0	1	S	R	R'
R'	S	R	1	0	I'	S'	I
S'	R'	I	S	I'	0	R'	R
I'	I	1	R	S'	R'	0	S
1	S'	I'	R'	I	R	S	0

□

Remark 3.8 It is easy to check that

$$R'S'I' = 1, \qquad R'S' + S'I' + I'R' = 1, \qquad R' + S' + I' = 0,$$

i.e. that R', S' and I' are the roots of $X^3 + X + 1$, and that

$$I'^2 = R', \qquad R'^2 = S', \qquad S'^2 = I'.$$

With R', S', I' and these last relations another circular presentation is possible—complementary to the one with the R, S, I. But we do prefer the R, S, I because of their linear independence (Proposition 4.1).

Remark 3.9 The conditions $R^2 = S,\ \ S^2 = I,\ \ I^2 = R$ are 'almost' not necessary. In fact, the other conditions imply that $R^2 = S$ or $R^2 = I$, i.e. $(R^2 + S)(R^2 + I) = 0$, or $R^4 + R^2(S+I)+SI = 0$, $R^4+R(RS+RI)+SI = 0$, $R^4+RSI+SI = 0$, $R^4+1+R^{-1} = 0$, or $R^5+R+1 = 0$. And this formula is equivalent to $R^5(R^3+R^2+1) = 0$, since we know a priori that $R^7 = 1$, from the general theory, as the degree of $X^3 + X^2 + 1$ is 3 and as $8 = 2^3$.

3.3 Linear and Polynomial Equations in $\mathbb{F}_8$, Paradoxes in $\mathbb{F}_8$

In $\mathbb{F}_8 = \mathbb{F}_2^3$, we can discuss and solve $\mathbb{F}_2$-linear equations, and in the field $\mathbb{F}_8$ some polynomial equations in one variable with degree 1, 2, 3. We show that reduced third degree equations are paradoxical, i.e. without solutions when they correspond to bijective linear maps.

Proposition 3.10 *$\mathbb{F}_8$ is an $\mathbb{F}_2$-linear space, and (R, S, I) is a basis of it (see Proposition 4.1).*

Each $\mathbb{F}_2$-linear map $f = \mathbb{F}_8 \to \mathbb{F}_8$, with $f(R) = e_1$, $f(S) = e_2$, $f(I) = e_3$, is given by a unique expression

$$f(u) = au^4 + bu^2 + cu.$$

The discussion of the equation $f(u) = u'$ comes back to the discussion of a system in $\mathbb{F}_8^3$ with parameter and unknown in $\mathbb{F}_2$.

Proof Clearly, $m(-)$, with $m \in \mathbb{F}_8$ and $(-)^2$ are linear, and so $f(u) = au^4 + bu^2 + cu$ is linear. Conversely, given an $\mathbb{F}_2$-linear map $f : \mathbb{F}_8 \to \mathbb{F}_8$, we can find $a, b, c \in \mathbb{F}_8$ such that $f(u) = au^4 + bu^2 + cu$, i.e. solution of the system

$$aR^4 + bR^2 + cR = f(R); \qquad aS^4 + bS^2 + cS = f(S); \qquad aI^4 + bI^2 + cI = f(I),$$

because this is an $\mathbb{F}_8$-linear system, where $\mathbb{F}_8$ is a commutative field, has determinant 1, and the solutions can be expressed via the determinant; we find

$$a = e_2R + e_3S + e_1I;$$
$$b = e_3R + e_1S + e_2I;$$
$$c = e_1R + e_2S + e_3I.$$

Let us remark that in any case, with $u = xR + yS + zI$, $x, y, z \in \{0, 1\}$, $u' = x'R + y'S + z'I$, $x', y', z' \in \{0, 1\}$, the linear equation

$$au^4 + bu^2 + cu = u',$$

with $a, b, c \in \mathbb{F}_8$ is equivalent to the following system, with $x', y', z' \in \{0, 1\}$:

$$cx + ay + bz = x'; \qquad bx + cy + az = y'; \qquad ax + by + cz = z',$$

with determinant $(a + b + c)(a^2 + b^2 + c^2 + ab + bc + ca)$, and there is a unique solution if and only if: $a + b + c = 1$ and $a^2 + b^2 + ab + a + b + 1 \neq 0$, and the solution $(x_1, y_1, z_1) \in \mathbb{F}_8^3$ given by Cramer's formulas is in fact in $\mathbb{F}_2^3$, i.e. is such that $x_1^2 = x_1$, $y_1^2 = y_1$, $z_1^2 = z_1$. □

Proposition 3.11 *Each $\mathbb{F}_2$-linear form on $\mathbb{F}_8$, $f = \mathbb{F}_8 \to \mathbb{F}_2$ is given by the scalar product with a vector c (as introduced in Proposition 4.1):*

$$f(u) = \langle c, u\rangle := \mathrm{tr}(cu) = c^4u^4 + c^2u^2 + cu.$$

Proof The map $\langle c, -\rangle$ is linear. For a linear form we need $f(u) \in \{0, 1\}$, i.e. for every u, $f(u)^2 = u$, i.e. $b^2u^4 + c^2u^2 + a^2u = au^4 + bu^2 + cu$, i.e. $b = c^2$, $a = c^4$. □

Proposition 3.12 *With the notations of Proposition 3.10, let $f : \mathbb{F}_8 \to \mathbb{F}_8$ be an $\mathbb{F}_2$-linear map given by*

$$f(u) = au^4 + bu^2 + cu.$$

This f is bijective if and only if

$$\Delta := a^7 + b^7 + c^7 + abc\big(a^3b + b^3c + c^3a\big) \neq 0.$$

If f is bijective, its inverse f^{-1} is given by:

$$f^{-1}(v) = lv^4 + mv^2 + nv,$$

with l, m, n polynomial functions of e_1, e_2, e_3, and rational functions of a, b, c.

Proof f is bijective if and only if (e_1, e_2, e_3) is a basis, i.e. if and only if $[e_1, e_2, e_3] = 1$ (cf. Proposition 4.4). Then f^{-1} is the map sending e_1, e_2, e_3 to R, S, I, i.e. it is given by $f^{-1}(v) = lv^4 + mv^2 + nv$, with the system of equations:

$$le_1^4 + me_1^2 + ne_1 = R; \qquad le_2^4 + me_2^2 + ne_2 = S; \qquad le_3^4 + me_3^2 + ne_3 = I,$$

with determinant 1, and the solutions are given by the Cramer's formulas:

$$\begin{aligned} l &= \left(e_2^2e_3 + e_2e_3^2\right)R + \left(e_3^2e_1 + e_3e_1^2\right)S + \left(e_1^2e_2 + e_1e_2^2\right)I, \\ m &= \left(e_2^4e_3 + e_2e_3^4\right)R + \left(e_3^4e_1 + e_3e_1^4\right)S + \left(e_1^4e_2 + e_1e_2^4\right)I, \\ n &= \left(e_2^4e_3^2 + e_2^2e_3^4\right)R + \left(e_3^4e_1^2 + e_3^2e_1^4\right)S + \left(e_1^4e_2^2 + e_1^2e_2^4\right)I. \end{aligned}$$

Also, writing $f^{-1}(f(u)) = u$, we have the system

$$lc^4 + mb^2 + na = 0; \qquad la^4 + mc^2 + nb = 0; \qquad lb^4 + ma^2 + nc = 1.$$

This system has a unique solution if its determinant Δ is $\neq 0$, and the solution is given by Cramer's formulas:

$$l = \frac{b^3 + c^2a}{\Delta}; \qquad m = \frac{a^5 + bc^4}{\Delta}; \qquad n = \frac{c^6 + a^4b^2}{\Delta}. \qquad \square$$

Proposition 3.13 *In $\mathbb{F}_8$, we consider the second degree equation*

$$ax^2 + bx + c = 0, \quad \textit{with } a \neq 0.$$

1. *If $a = 0$ and $b \neq 0$ there is a unique solution $x_1 = b^6c$.*
2. *If $a \neq 0$ and $b = 0$ there is a unique solution $x_1 = a^3c^4$.*
3. *If $a, b \neq 0$ and $c = 0$, there are two solutions $x_1 = 0$ and $x_2 = a^6b$.*
4. *If $a, b, c \neq 0$, there are solutions if and only if $a^3b^3c^3 + ac + b^2 = 0$, and these solutions are*:

$$x_1 = a^5b^3c^6, \qquad x_2 = ab^4c^2.$$

Proof The standard method with $\Delta = b^2 - 4ac$ and division by $2a$ is not available in characteristic 2. Rather we introduce $x = \frac{b}{a}y$, and the equation becomes $y^2 + y = \frac{ac}{b^2} := d$, $d \neq 0$.

For any y, we have $(y^2 + y)^4 + (y^2 + y)^2 + (y^2 + y) = 0$, and the necessary condition: $d^4 + d^2 + d = 0$, i.e. $d^3 = d + 1$, and the announced condition.

In this case, $d^{12} = d^4 + 1$, one solution is $y_1 = d^{-1} = d^6$, the other is $y_2 = y_1 + 1$, that is to say, $x_1 = a^5b^3c^6$, $x_2 = x_1 + \frac{b}{a} = ab^4c^2$.

Propositions 4.5 and 4.6 give a geometric interpretation of the second degree equation. $\square$

Proposition 3.14 *In $\mathbb{F}_8$, the third degree equation*

$$ax^3 + bx^2 + cx + d = 0$$

can be solved as follows:

1. *If* $a = 0$, *discussion and solution are given in Propositions* 3.13.
2. *If* $a \neq 0$, *with* $x = u + a^6 b$ *and multiplying by* u, *we obtain a linear equation as in Proposition* 3.10:

$$au^4 + \big(a^6 b^2 + c\big)u^2 + \big(a^6 bc + d\big)u = 0.$$

Example

1. Generally, the equation

$$Ru^4 + Su^2 + Iu = u'$$

   has no solution. Applying Proposition 3.10, we obtain that this equation has a solution if and only $u' = 0$ or 1, and the solution is $u = 0$ or 1.
2. The equation

$$Ru^2 + Su + I = 0$$

   has no solution, as we see by Proposition 3.13.
3. The equation

$$u^2 + Su + 1 = 0$$

   admits 2 solutions, R and R'.

Proposition 3.15 *If* $f(u) = au^4 + bu^2 + cu$ *is a bijective* $\mathbb{F}_2$*-linear map on* $\mathbb{F}_8$, *with* $a \neq 0$, *then the polynomial* $au^3 + bu + c = 0$ *has no root in* $\mathbb{F}_8$, *and so it expresses a paradox in* $\mathbb{F}_8$. *In this way, we exactly get the different reduced* (= *without the degree* 2 *term*) *third degree equations without solutions in* $\mathbb{F}_8$, *of which there are fewer than* 168.

Proof f is bijective if and only if $f(u) = 0$ has only 0 as a solution. The second point results from the fact that $\mathrm{GL}_2(\mathbb{F}_2)$ is of cardinality 168. □

4 Construction of $\mathbb{F}_8$ from Its 3-Dimensional Vectorial Geometry and Its 3-Circular Boolean Logic

We develop the 3 dimensional geometry of $\mathbb{F}_8$, starting from the circular presentation with R, S, I, and we introduce the '*canonical*' boolean algebra on $\mathbb{F}_8$. Then we re-construct the multiplication of the field, starting from the geometry (scalar, cross, and mixed products), or from the logic (conjunction and negation), with the help of the Frobenius squaring $(-)^2$. This allows us to prove that any function on $\mathbb{F}_8$ is a boolean combination of '*avatars*' of its variables. From the logical point of view, the squaring u^2 of u is a kind of modality, as well as the 6 associated avatars $u^{(i)}$, with $2 \leq i \leq 7$. These avatar-like operations $(-)^{(i)}$ are organized in a commutative monoïd $\mathbb{A}$. So the full logic of $\mathbb{F}_8$ appears as a boolean logic with a 3-circular automorphism, as well as with modalities coming from the action of $\mathbb{A}$.

4.1 Geometrical Tools in $\mathbb{F}_8$

4.1.1 Scalar, Cross and Mixed Products

Proposition 4.1 *$\mathbb{F}_2$ is a sub-field of the field $\mathbb{F}_8$, and so $\mathbb{F}_8$ is an $\mathbb{F}_2$-algebra of dimension* 3. *A basis is given by $\kappa = (R, S, I)$. Given $u = xR + yS + zI$ and $u' = x'R + y'S + z'I$, with $x, y, z, x', y', z' \in \mathbb{F}_2$, we define the* trace *and the* scalar product *by*

$$\mathrm{tr}(u) = x + y + z, \quad \langle u, u' \rangle = xx' + yy' + zz'.$$

We have $\mathrm{tr}(u) = \langle u, u \rangle$, *and the values*:

$$\mathrm{tr}(u) = 0 \quad \Leftrightarrow \quad u \in \{0, R', S', I'\}, \quad \mathrm{tr}(u) = 1 \quad \Leftrightarrow \quad u \in \{1, R, S, I\}.$$

We introduce the cross product (*also named crossed product or vector product*) *and the* mixed product (*also named scalar triple product*) *by*:

$$u \times u' = (yz' - zy')R + (zx' - xz')S + (xy' - yx')I,$$
$$[u, u', u''] = \langle u, u' \times u'' \rangle = \langle u \times u', u'' \rangle.$$

Then (u, u', u'') is a basis of $\mathbb{F}_8$ over $\mathbb{F}_2$ if and only if $[u, u', u''] = 1$. Furthermore, we have (double cross product *formula*):

$$u \times (u' \times u'') = \langle u, u'' \rangle u' - \langle u, u' \rangle u''.$$

Proof There is no linear relation between R, S and I: $R, S, I \neq 0$, $R + S = R' \neq 0$, $S + I = S' \neq 0$, $I + R = I' \neq 0$, $R + S + I = 1 \neq 0$. For $\mathrm{tr}(u) = \langle u, u \rangle$ we have $x^2 = x$, $y^2 = y$, $z^2 = z$. Clearly, tr is linear, $\langle , \rangle$ and $\times$ are bilinear, both symmetrical (we are in characteristic 2). It is elementary to check that if $(x, y, z) \neq (0, 0, 0)$ and $(x', y', z') \neq (0, 0, 0)$ then $u \times u' = 0$ if and only if $u = u'$. We have $[u, u', u''] \neq 1$ exactly if $[u, u', u''] = 0$, and this means that $u = 0$, or $u \neq 0$ and $u = u'$, or $u \neq 0$ and $u = u''$, or $u \neq 0$ and $u = u' + u''$. Another proof results from the formula for $[u, u', u'']$ in Proposition 4.3. The last formula could be verified directly. □

Proposition 4.2 *As a linear space of characteristic* 2 *equipped with a bilinear* (*anti*)*symmetric multiplication $\times$, $\mathbb{F}_8$ could be presented in a circular symmetric way—and we shall name this* the R, S, I linear multiplicative presentation—*by the following relations among the elements*:

$$R \times S = I, \qquad S \times I = R, \qquad I \times R = S.$$

Then, because of the linear relations

$$R' = R + S, \qquad S' = S + I, \qquad I' = I + R, \qquad 1 = R + S + I,$$

the cross product $\times$ is given by the table:

$\times$	R	S	I	R'	S'	I'	1
R	0	I	S	I	S'	S	S'
S	I	0	R	I	R	I'	I'
I	S	R	0	R'	R	S	R'
R'	I	I	R'	0	1	1	R'
S'	S'	R	R	1	0	1	S'
I'	S	I'	S	1	1	0	I'
1	S'	I'	R'	R'	S'	I'	0

4.1.2 Geometrical Operations from the Operations of the Field

Proposition 4.3 *The space $\mathbb{F}_8$ yields the $\mathbb{F}_2$-linear map of* squaring $(-)^2$ (*Frobenius map*). *For any u, the elements u, u^2, u^4 are said to be* conjugate; *they are*:

$$u = xR + yS + zI, \qquad u^2 = zR + xS + yI, \qquad u^4 = yR + zS + xI,$$

and their sum is the trace *of u*:

$$\mathrm{tr}(u) = u + u^2 + u^4 = x + y + z.$$

The scalar product is the trace of the product:

$$\langle u, u'\rangle = \mathrm{tr}(uu') = uu' + (uu')^2 + (uu')^4 = xx' + yy' + zz',$$

and if $u = xR + yS + zI$, then $x = \mathrm{tr}(uR)$, $y = \mathrm{tr}(uS)$, $z = \mathrm{tr}(uI)$.

The cross product and mixed product from Proposition 4.1 are given by

$$u \times u' = \big(uu'(u + u')\big)^2,$$
$$[u, u', u''] = uu'u''(u + u')(u' + u'')(u'' + u)(u + u' + u'').$$

So, all the vectorial analysis in $\mathbb{F}_8$ according to the tools in Proposition 4.1 is expressible in terms of polynomial functions with coefficients in $\mathbb{F}_2$.

Proof We have $\mathrm{tr}(uR) = xR + x^2S + x^4I$, $\mathrm{tr}(uR)R = u^2 + u^4 + u^4R + (u + u^2)S$; with similar formulas for $\mathrm{tr}(uS)S$ and $\mathrm{tr}(uI)I$, the sum of the three is u.

We compute $(yz' - y'z)R = (yz' + y'z)R$ as $(\mathrm{tr}(uS)\,\mathrm{tr}(u'I) + \mathrm{tr}(u'S)\,\mathrm{tr}(uI))R$, and we obtain

$$(yz' - y'z)R = (u + u')S' + (uu'^3 + u'u^3)S + (u^3 + u'^3)I'.$$

With similar formulas for the other two terms, the sum of the three provides the announced formula.

If we use the formulas for the scalar product and for the cross product, then

$$[u, u', u''] = u(u'^2u''^4 + u'^4u''^2) + u'(u''^2u^4 + u''^4u^2) + u''(u^2u'^4 + u^4u'^2),$$

and this is also given by the proposed formula for the mixed product. □

Proposition 4.4 *We have*

$$[u, u', u''] = \begin{vmatrix} x & x' & x'' \\ y & y' & y'' \\ z & z' & z'' \end{vmatrix} = \begin{vmatrix} u & u' & u'' \\ u^2 & u'^2 & u''^2 \\ u^4 & u'^4 & u''^4 \end{vmatrix} \in \{0, 1\}.$$

The value is 1 *if and only if* (u, u', u'') *is a basis.*

Proof The first determinant is equivalent to the definition of $[u, u', u'']$—and it is in $\{0, 1\}$—and the second is equivalent to the last formula in the previous proof. Let us remark that another determinant expression for $[u, u', u'']$—the so called *conjunctive determinant* of Definition 4.13—will be obtained in Proposition 4.14. □

4.1.3 The Equation $a \times u = b$

Proposition 4.5 *In* $\mathbb{F}_8$, *let* $a \neq 0, b \neq 0$; *the equation*

$$a \times u = b.$$

has a solution if and only if $\langle a, b\rangle = 0$, *and then there are two solutions*:

$$u_1 = a^4b^3, \qquad u_2 = a^2b.$$

Proof If $a \times u = b$, then $0 = [a, a, u] = \langle a, b\rangle$, and the condition is necessary.

For the converse, using the double cross product formula (in Proposition 4.1), with $u_0 = a \times b$ we have $a \times (a \times b) = \langle a, a\rangle b$, and if $\langle a, a\rangle \neq 0$ we get a solution $u'_1 = \langle a, a\rangle^{-1} a \times b$, and another $u'_2 = u'_1 + a$. But if $\langle a, a\rangle = 0$ (i.e. if $a = 1, R', S'$ or I') this method fails. So if we start with the algebraic formula for $\times$ given in Proposition 4.3, and $(au(a + u))^2 = b$, we obtain $au(a + u) = b^4$, $au^2 + a^2u = b^4$, and with Proposition 3.13 we obtain solutions if and only if $(aa^2b^4)^3 + (ab^4) + (a^2)^2 = 0$, or equivalently if $(ab)^4 + (ab)^2 + ab = 0$. Then the solutions are $u_1 = a^5(a^2)^3(b^4)^6 = a^4b^3$ and $u_2 = a(a^2)^4(b^4)^2 = a^2b$. Of course, $u_2 = u_1 + a$. □

Proposition 4.6 *In* $\mathbb{F}_8$, *a convenient change of variable* $x = \lambda u$ *transforms any second degree equation* $ax^2 + bx + c = 0$ *into a vectorial division problem* $A \times u = B$, *and so the equation obtains a geometrical meaning. Conversely, the multiplication and the algebraic structure of field of* $\mathbb{F}_8$, *not only help solving algebraic equations, but also geometrical linear problems.*

Proof Propositions 3.13 and 4.5 provide a passage between $ax^2 + bx + c = 0$ and $A \times u = B$: with $x = \frac{b}{a}y$ and $u = av$ these equations are equivalent to $y^2 + y = \frac{ac}{b^2}$ and $u^2 + u = (AB)^4$, and if we take A and B such that $AB = a^2b^3c^2$, the second degree equation is equivalent to the vectorial division, the correspondence between solutions being given by $x_i = \frac{b}{aA}u_i$, for $i = 1, 2$. □

4.2 The Canonical Logic of $\mathbb{F}_8$

4.2.1 Definition of ∧ and ¬

Proposition 4.7 *If we define the* canonical conjunction ∧, *the canonical disjunction* ∨, *and the* canonical negation ¬ *on* $\mathbb{F}_8$ *by*

$$u \wedge u' = xx'R + yy'S + zz'I, \qquad u \vee u' = (x \vee x')R + (y \vee y')S + (z \vee z')I,$$
$$\neg u = (x+1)R + (y+1)S + (z+1)I,$$

then we obtain on $\mathbb{F}_8$ *a structure of a boolean algebra, with atoms* R, S *and* I, *with 'false'* $= 0$ *and 'truth'* $= 1$, *with also* $+$ *as the 'symmetric difference', i.e.*

$$u + u' = (u \wedge \neg u') \vee (\neg u \wedge u').$$

Proof Obviously, the structure is boolean because it is componentwise in $\mathbb{F}_2$; it could be named 'canonical' because it is associated to the very special basis (R, S, I) characterized among all the bases by an arithmetical property (see Proposition 6.2). The $\neg u = u + 1$ results from $1 = R + S + I$. □

4.2.2 Relations Between the Logic, the Field Structure, and the Geometry

Proposition 4.8 *If we employ the squaring* $(-)^2$, *the canonical conjunction* ∧ *and the cross product* ×, *then the product of* $\mathbb{F}_8$ *is*:

$$uu' = (u \wedge u')^2 + u \times u' + (u \times u')^2.$$

The boolean operations are expressible with the field operations:

$$u \wedge u' = u^4u'^4 + u^4u'^2 + u^2u'^4 + u^2u' + uu'^2, \qquad \neg u = u + 1,$$
$$u \vee u' = u^4u'^4 + u^4u'^2 + u^2u'^4 + u^2u' + uu'^2 + u + u',$$
$$u \Rightarrow u' = u^4u'^4 + u^4u'^2 + u^2u'^4 + u^2u' + uu'^2 + u' + 1.$$

Proof For the relation between $u \wedge u'$, uu', and $u \times u'$, we compute

$$u^2u' + uu'^2 = uu'(u + u') = (u \times u')^4 = (zx' - z'x)R + (xy' - x'y)S + (yz' - y'z)I,$$

and also $(u^2u' + uu'^2)^2$ and $(u^2u' + uu'^2)^4$, and then we verify the formula for uu', or, equivalently, the formula for $u \wedge u'$ (using $u^8 = u$, $v^8 = v$). Then $u \vee u' = u \wedge u' + u + u'$, $u \Rightarrow u' = u \wedge u' + u + 1$. □

Remark 4.9 In principle, our formula $uu' = (u \wedge u')^2 + u \times u' + (u \times u')^2$ in Proposition 4.7 provides a construction respectful of the circular symmetry or the symmetry of the situation: on the one hand, the × is circular and symmetric—$R \times S = I$, $S \times I = R$, $I \times R = S$, and $R \times S = S \times R, \ldots$, the $(-)^2$ is circular—$R^2 = S$, $S^2 = I$, $I^2 = R$; on the other hand, the ∧ is symmetric $R \wedge R = R$, $R \wedge S = S \wedge R = 0, \ldots$. It is a kind of decomposition of the product into circular and non-circular components.

Proposition 4.10 *We have*

$$\langle u, u' \rangle = \mathrm{tr}\big(u \wedge u'\big).$$

Proof From the formula for uu' in Proposition 4.8 we get by addition of uu', $(uu')^2$ and $(uu')^4$, $\langle u, u' \rangle = u \wedge u' + (u \wedge u')^2 + (u \wedge u')^4 = \mathrm{tr}(u \wedge u')$. □

Remark 4.11 The value of $\langle u, u' \rangle$ depends only on $u \wedge u'$, and the canonical conjunction could be seen as an enriched scalar product; so it could be considered as a kind of 'geometrical operation'. The operation $(-)^2$ also has a geometrical meaning : it is a rotation $R \mapsto S \mapsto I \mapsto R$. Hence our formula $uu' = (u \wedge u')^2 + u \times u' + (u \times u')^2$ is a geometrical reconstruction of the field law in the 3-dim space $\mathbb{F}_8$, which is possible because of the characteristic 2.

Of course, in characteristic 0 the situation would be completely different, and it is well known that it is impossible to construct a field structure on $\mathbb{R}^3$; rather the cross product in $\mathbb{R}^3$ could be understood as a part of a field structure on $\mathbb{R}^4$ (the quaternion field). The same can be do with $\mathbb{F}_2^3$ in $\mathbb{F}_2^4$, but also in addition to that, in characteristic 2 the situation could be tightened in 3 dimensions (as we have seen in the field $\mathbb{F}_8$).

Proposition 4.12 *In* $\mathbb{F}_8$, *the operation* $(-)^2$ *commutes with* $\times$, $+$, $\wedge$, *and* $\neg$:

$$\big(u \times u'\big)^2 = u^2 \times u'^2, \qquad \big(u + u'\big)^2 = u^2 + u'^2,$$
$$\big(u \wedge u'\big)^2 = u^2 \wedge u'^2, \qquad (\neg u)^2 = \neg\big(u^2\big).$$

So $(-)^2$ *is linear and boolean.*

Proof We have $u \times u' = (yz - zy')R + (zx' - xz')S + (xy' - yx')I$, $(u \times u')^2 = (xy' - yx')R + (yz' - zy')S + (zx' - xz')I$, which is equal to $u^2 \times u'^2$ with $u^2 = zR + xS + yI$ and $u'^2 = z'R + x'S + y'I$. For the second formula the two members are equal to $zz'R + xx'S + yy'I$. □

Definition 4.13 In $\mathbb{F}_8$ equipped with $\wedge$ and $+$, we define the *conjunctive determinant* of 3 elements $u, u', u'' \in \mathbb{F}_8$ as $\det_\wedge(u, u', u'') = u \wedge u'^4 \wedge u''^2 + u^4 \wedge u' \wedge u''^2 + u^2 \wedge u' \wedge u''^4 + u \wedge u'^2 \wedge u''^4 + u^4 \wedge u'^2 \wedge u'' + u^2 \wedge u'^4 \wedge u''$,

$$= \begin{vmatrix} u & u' & u'' \\ u^2 & u'^2 & u''^2 \\ u^4 & u'^4 & u''^4 \end{vmatrix}_\wedge,$$

this notation means that in order to expand this 'determinant', we have to use $\wedge$ instead of product.

Proposition 4.14 *In* $\mathbb{F}_8 = \mathbb{F}_{2^3}$ *equipped with squaring* $(-)^2$ *and conjunction* $\wedge$, *the field product is*

$$uu' = u^2 \wedge u'^2 + u \wedge u'^4 + u^4 \wedge u' + u^4 \wedge u'^2 + u^2 \wedge u'^4,$$

the cross product is

$$u \times u' = u^4 \wedge u'^2 + u^2 \wedge u'^4,$$

the scalar product is

$$\langle u, u' \rangle = u \wedge u' + u^2 \wedge u'^2 + u^4 \wedge u'^4,$$

and the mixed product is the conjunctive determinant from Definition 4.13, namely

$$[u, u', u''] = \underset{\wedge}{\det}(u, u', u'').$$

Proof The formula $u \wedge u' = u^4u'^4 + u^4u'^2 + u^2u'^4 + u^2u' + uu'^2$ from Proposition 4.7 will be completely 'reversed', with an expression of uu' as a composition of $\wedge$ and $(-)^2$; we have to add just the following:

$$u^2 \wedge u'^2 = uu' + uu'^4 + u^4u' + u^4u'^2 + u^2u'^4,$$
$$u \wedge u'^4 + u^4 \wedge u' = u^4u' + uu'^4,$$
$$u^4 \wedge u'^2 + u^2 \wedge u'^4 = u^4u'^2 + u^2u'^4.$$

The last formula also implies the announced formula for $u \times u'$.

For the scalar product we expand $uu' + u^2u'^2 + u^4u'^4$. For the mixed product we expand $\langle u, u' \times u'' \rangle = \langle u, u'^4 \wedge u''^2 + u'^2 \wedge u''^2 \rangle$, assuming the commutations from Proposition 4.12: $[u, u', u''] = u \wedge u'^4 \wedge u''^2 + u^4 \wedge u' \wedge u''^2 + u^2 \wedge u' \wedge u''^4 + u \wedge u'^2 \wedge u''^4 + u^4 \wedge u'^2 \wedge u'' + u^2 \wedge u'^4 \wedge u''$, i.e. the announced conjunctive determinant (cf. Definition 4.13). □

Proposition 4.15 *We have*

$$Ru = R' \wedge u^4 + S' \wedge u^2 + I \wedge u,$$

and the squared boolean expression of geometrical operations:

$$R \times u = S \wedge u^4 + I \wedge u^2,$$
$$\langle R, u \rangle = I \wedge u^4 + S \wedge u^2 + R \wedge u.$$

Proof A consequence of formulas in Proposition 4.14. □

5 Presentation of $\mathbb{P}_8$ by Boolean Combination of Logical Avatars

5.1 From Powers of u to Sums of Conjunctions of u, u^2, u^4

Proposition 5.1 *Given $u \in \mathbb{F}_8$ we can express powers of u as sums of conjunctions of the powers u, u^2, u^4*:

$$u = u, \qquad u^2 = u^2, \qquad u^3 = u + u^4 + u \wedge u^2, \qquad u^4 = u^4, \qquad u^5 = u^2 + u^4 + u \wedge u^4,$$

$u^6 = u + u^2 + u^2 \wedge u^4, \qquad u^7 = (u + u^2 + u^4) + (u \wedge u^2 + u^2 \wedge u^4 + u^4 \wedge u) + u \wedge u^2 \wedge u^4.$

And conversely, the previous conjunctions are polynomials:

$$u \wedge u^2 = u + u^3 + u^4, \qquad u^2 \wedge u^4 = u + u^2 + u^6, \qquad u^4 \wedge u = u^2 + u^4 + u^5,$$
$$u \wedge u^2 \wedge u^4 = u + u^2 + u^3 + u^4 + u^5 + u^6 + u^7.$$

Proof With the formula for uu' (Proposition 4.14), and the commutation of $(-)^2$ with $\wedge$ (Proposition 4.12), we expand $u^3 = uu^2$, $u^5 = uu^4$, $u^6 = u^2u^4$, and $u^7 = u^6u$. □

5.2 Variant with Cross Product

Proposition 5.2 *We have*:

$$u \wedge u^2 = u^4 + u \times u^2, \qquad u^2 \wedge u^4 = u + u^2 \times u^4, \qquad u^4 \wedge u = u^2 + u^4 \times u.$$

Proof From Proposition 4.14 with have $u \times u' = u^4 \wedge u'^2 + u^2 \wedge u'^4$, and we obtain $u \times u^2$, etc. □

5.3 Conjunctive Avatars

Definition 5.3 For u a variable on $\mathbb{F}_8$, the set of *conjunctive avatars* or *avatars* of u is the set of functions

$$\mathbb{A}(u) = \{u, u^2, u^4, u \wedge u^2, u^2 \wedge u^4, u^4 \wedge u, u \wedge u^2 \wedge u^4\},$$

and they can be expressed with product or with cross products (Propositions 5.1 and 5.2).

In fact, we have:

$$\mathbb{A}(0) = \{0\}, \qquad \mathbb{A}(1) = \{1\},$$
$$\mathbb{A}(R) = \mathbb{A}(S) = \mathbb{A}(I) = \{0, R, S, I\},$$
$$\mathbb{A}(R') = \mathbb{A}(S') = \mathbb{A}(I') = \{0, R, S, I, R', S', I'\}.$$

Our notations for avatars of u, or elements of $\mathbb{A}(u)$, are: $u^{(1)} = u$, $u^{(2)} = u^2$, $u^{(3)} = u \wedge u^2$, $u^{(4)} = u^4$, $u^{(5)} = u \wedge u^4$, $u^{(6)} = u^2 \wedge u^4$, $u^{(7)} = u \wedge u^2 \wedge u^4$. And also we introduce the notation $u^{(0)} = u^0 = 1$.

Proposition 5.4 *The 'avatarian' functions* $(-)^{(i)}$ *in Definition 5.3 are organized in a commutative monoïd* $\mathbb{A}$ *given by the table*:

∘	1 2 4	3 5 6	7
1	1 2 4	3 5 6	7
2	2 4 1	6 3 5	7
4	4 1 2	5 6 3	7
3	3 6 5	7 7 7	7
5	5 3 6	7 7 7	7
6	6 5 3	7 7 7	7
7	7 7 7	7 7 7	7

Proof We have just to check the compositions. For the associativity, in the expression $a \circ (b \circ c) = (a \circ b) \circ c$, both sides are equal to 7 if one of the a, b, or c is 7, or if two of them are in $\{3, 5, 6\}$; in the other cases, the compositions are multiplications of numbers modulo 7, so are associative. □

5.4 Construction of $\mathbb{P}_8$

5.4.1 $\mathbb{P}_8$ by Polynomial Expressions

Proposition 5.5 *Any function $Z : \mathbb{F}_8^k \to \mathbb{F}_8$ is a polynomial with variables $u_1, u_2, \ldots, u_k$ with coefficients in $\mathbb{F}_8$.*

Proof We know that if $u - w \neq 0$, then $(u - w)^7 = 1$, and then for any $w \in \mathbb{F}_8$ the *indicator function* of w is the polynomial function

$$[w](u) = 1 - (u - w)^7 = \begin{cases} 1 & \text{if } u = w, \\ 0 & \text{if } u \neq w. \end{cases}$$

Then if $E \subseteq \mathbb{F}_8$, the *indicator function* or the *characteristic function* of E is the sum

$$[E](u) = \sum_{w \in E} [w](u) = \begin{cases} 1 & \text{if } u \in E, \\ 0 & \text{if } u \notin E. \end{cases}$$

and more generally, when $k = 1$, for an arbitrary function Z we have

$$Z(u_1) = \sum_{z \in \mathbb{F}_8} \sum_{\{w \in \mathbb{F}_8; Z(w) = z\}} z[w](u_1);$$

if $k = 2$, then Z is given by

$$Z(u_1, u_2) = \sum_{z \in \mathbb{F}_8} \sum_{\{(w_1, w_2) \in \mathbb{F}_8{}^2; Z(w_1, w_2) = z\}} z[w_1](u_1)[w_2](u_2),$$

expression in which, in fact,

$$z[w_1](u_1)z[w_2](u_2) = z \wedge [w_1](u_1) \wedge [w_2](u_2),$$

as the functions $[u_1]$ and $[u_2]$ are with values in $\{0, 1\}$. Let us remark that every function $Z : A^k \to A$ on a commutative unitary ring A is a polynomial if and only if A is a finite field [9]; it is the case of $\mathbb{F}_8$. We remark that the point in the Heisler's theorem is that if A is a finite unitary commutative ring and if the indicator of 1, i.e. $[1](u)$ is a polynomial, then as $[1](0) = 0$, we have $[1](u) = ug(u)$ with $g(u)$ a polynomial, and then if $u \neq 0$ we have $ug(u) = 1$, i.e. $g(u)$ is an inverse of u. Here we just need that $[1](u) = 1 - (1 - u)^7$. □

5.4.2 $\mathbb{P}_8$ by Sums of Conjunctions of Avatars and Constants

Proposition 5.6 *Any function $Z : \mathbb{F}_8^k \to \mathbb{F}_8$ with variables $u_1, u_2, \ldots, u_k$ is a sum of conjunctions of constants in $\mathbb{F}_8$ and the various u_i, ${u_i}^2$ and ${u_i}^4$, $1 \leq i \leq k$.*

For example, in the case $k = 2$, any function Z has a presentation where $d_{i,j}$ is a sum of elements of $\{0, 1, R, S, I\}$, $0 \leq i, j \leq 7$:

$$Z(u_1, u_2) = \sum_{i,j} d_{i,j} \wedge {u_1}^{(i)} \wedge {u_2}^{(j)}.$$

So the full logic of $\mathbb{F}_8$, i.e. the Post–Malcev full iterative algebra $\mathbb{P}_8 = \mathbb{P}(\mathbb{F}_8) = \bigcup_{n \geq 1} \mathbb{F}_8^{{\mathbb{F}_8}^n}$ of all functions of all arities on $\mathbb{F}_8$ (as defined in [13] and [10]), is generated by its canonical boolean operations $\wedge$ and $\neg$, the 3 constant functions R, S, I and *the 3-circular automorphism $(-)^2$.*

Proof It is a consequence of Proposition 5.5 and details in its proof, and Proposition 4.14. Any monomial $c{u_1}^{n_1}{u_2}^{n_2} \cdots$ in Z in Proposition 5.5 is reducible to the case where $n_j \leq 7$, and we apply the formula for the product in Proposition 4.14, using commutation of $(-)^2$ with $\wedge$ (Proposition 4.12). For example, any monomial $c{u_1}^m{u_2}^n$ is a sum of terms of the form $d \wedge ({u_1}^{p_1} \wedge {u_1}^{p_2} \wedge {u_1}^{p_3} \wedge {u_2}^{q_1} \wedge {u_2}^{q_2} \wedge {u_2}^{q_3})$, with $p_i, q_j \in \{0, 1, 2, 4\}$. This is explicitly done with Proposition 5.1. See also Proposition 4.15.

Another way is to use $Z(u_1, u_2) = \sum_{Z(w_1,w_2)=z} z \wedge [w_1](u_1) \wedge z[w_2](u_2)$, and to represent directly each indicator $[w]$ with $\wedge$, using Proposition 5.1 to expand $[w](u) = 1 + (u - w)^7$. But also we have

$$u \wedge u^2 \wedge u^4 = [1](u) \quad \text{and} \quad (u + w + 1) \wedge (u + w + 1)^2 \wedge (u + w + 1)^4 = [w](u).$$

The fact that the coefficients $d_{i,j}$ could be limited to values $0, 1, R, S, I$ results from $R' = R + S$, $S' = S + I$, $I' = I + R$. □

5.4.3 $\mathbb{P}_8$ by Sums of Products by Constants of Conjunctions of Avatars

Proposition 5.7 *With the hypothesis and notations of Proposition 5.6, every function could be written as*

$$Z(u_1, u_2) = \sum_{i,j} c_{i,j} \big({u_1}^{(i)} \wedge {u_2}^{(j)}\big),$$

where $c_{i,j}$ *is a sum of elements of* $\{0, 1, R, S, I\}$, $0 \le i, j \le 7$, *i.e. with canonical boolean operations* $\wedge$ *and* $\neg$, *the* 3 *bijective linear functions* $R^{\cdot} : w \mapsto Rw$, $S^{\cdot} : w \mapsto Sw$, $I^{\cdot} : w \mapsto Iw$ and *the* 3-*circular automorphism* $(-)^2$.

6 Auto-Dual Bases, Change of Bases in $\mathbb{F}_8$, R, S, I Borromean Generations of $GL_3(\mathbb{F}_2)$ and of $\mathbb{P}_8$

Dual bases and change of coordinates are exposed, and auto-dual bases in $\mathbb{F}_8$ are recognized. This allows generating the simple group $GL_3(\mathbb{F}_2)$ by 3 linear transformations $R^{\cdot}, S^{\cdot}, I^{\cdot}$. So this group is 'borromean'. We have also another borromean presentation by 3 linear involutions A, B, C. Then the Post–Malcev algebra $\mathbb{P}_8$ could be generated by the canonical boolean functions and the three linear involutions A, B, C.

6.1 Dual Bases

Two bases $\beta = (e_1, e_2, e_3)$ and $\beta^* = (e_1^*, e_2^*, e_3^*)$ are *dual* if $\mathrm{tr}(e_i e_j^*) = \delta_{i,j}$, where $\delta_{i,j}$ is Kronecker's symbol (with value 1 if $i = j$, and 0 if $i \neq j$). A basis $\beta = (e_1, e_2, e_3)$ is said to be *strictly auto-dual* if $\mathrm{tr}(e_i e_j) = \delta_{i,j}$, and *auto-dual* if, for a permutation σ on $\{1, 2, 3\}$, β and $\beta_\sigma = (e_{\sigma 1}, e_{\sigma 2}, e_{\sigma 3})$ are dual, i.e. such that $\mathrm{tr}(e_i e_j^*) = \delta_{i,j}$, with $e_j^* = e_{\sigma(j)}$.

Proposition 6.1 *If* $\beta = (e_1, e_2, e_3)$ *is a basis of* $\mathbb{F}_8$ *over* $\mathbb{F}_2$, *then we obtain a dual basis* $\beta^* = (e_1^*, e_2^*, e_3^*)$ *by*:

$$e_1^* = e_2 \times e_3, \qquad e_2^* = e_3 \times e_1, \qquad e_3^* = e_1 \times e_2,$$

and then $(\beta^*)^* = \beta$, *and the coordinates of* $u = u_1 e_1 + u_2 e_2 + u_3 e_3$ *are*:

$$u_1 = \mathrm{tr}\big(u e_1^*\big) = [u, e_2, e_3], \qquad u_2 = \mathrm{tr}\big(u e_2^*\big) = [e_1, u, e_3], \qquad u_3 = \mathrm{tr}\big(u e_3^*\big) = [e_1, e_2, u],$$

that is to say,

$$u_1 = u^4 + \big(e_2 e_3 + (e_2 + e_3)^2\big) u^2 + e_2 e_3 (e_2 + e_3), \quad \textit{etc.}$$

Also we have

$$u_1 = \big[\{e_1, e_1 + e_2, e_1 + e_3, e_1 + e_2 + e_3\}\big](u).$$

In particular, we obtain

$$e_i^* = \sum_{j=1,2,3} \langle e_i^*, e_j^* \rangle e_j, \qquad e_i = \sum_{j=1,2,3} \langle e_i, e_j \rangle e_j^*.$$

Proof By construction, $\langle e_1, e_1^* \rangle = [e_1, e_2, e_3] = 1$, $\langle e_1, e_2^* \rangle = [e_1, e_3, e_1] = 0$, etc. For $(\beta^*)^* = \beta$, for example, we have $e_2^* \times e_3^* = (e_3 \times e_1) \times (e_1 \times e_2) = ((e_3 \times e_1) \cdot e_2) e_1 + ((e_3 \times e_1) \cdot e_1) e_2 = [e_3, e_1, e_2] e_1 = e_1$, etc. Then $[u, e_2, e_3] = u_1 [e_1, e_2, e_3] + u_2 [e_2, e_2, e_3] + u_3 [e_3, e_2, e_3] = u_1 + 0 + 0 = u_1$, etc. The next formula comes from the formula for $[u, e_2, e_3]$ in Proposition 4.3; and the last two formulas are an application of the previous.

For the relation with $[\{e_1, e_1 + e_2, e_1 + e_3, e_1 + e_2 + e_3\}]$ we can argue directly that the four vectors $e_1, e_1 + e_2, e_1 + e_3, e_1 + e_2 + e_3$ are different, and the four others are $0, e_2, e_3, e_2 + e_3$, i.e. exactly those u with component $u_1 = 0$ on e_1. So it is the mixed product $[u, e_2, e_3]$, because this one is 0 if and only u, e_2, e_3 are linearly dependent, that is to say, $u \in \{0, e_2, e_3, e_2 + e_3\}$. □

An element u is said to be *normal* over $\mathbb{F}_2$ if (u, u^2, u^4) is a basis, which is called a *normal basis*. If furthermore u is *primitive*, i.e. if the powers of u generate $\mathbb{F}_8\backslash\{0\}$, then the basis is said to be *normal primitive*.

Proposition 6.2 *There are* 28 *bases of* $\mathbb{F}_8$, *or* 168 *when the order of terms is specified. Up to a circular permutation, there is only one* normal basis:

$$\kappa = (R, S, I) = \kappa^*,$$

which is even a normal primitive basis. *Up to a circular permutation, this* κ *is also the only* strictly auto-dual basis, *and there are* 3 *other auto-dual bases* (*not strict*), *which are*:

$$r = \big(R', I', 1\big), \qquad s = \big(1, S', R'\big), \qquad i = \big(S', 1, I'\big),$$

each one being its own dual, but with another order of terms:

$$r^* = \big(I', R', 1\big), \qquad s^* = \big(1, R', S'\big), \qquad i^* = \big(I', 1, S'\big).$$

Up to the order of terms, each basis $\phi = (f_1, f_2, f_3)$ *is of the form*

$$\phi = t_\phi \beta = (t_\phi e_1, t_\phi e_2, t_\phi e_3),$$

with $\beta = (e_1, e_2, e_3)$ *one of the four auto-dual bases* κ, r, s *or* i, *and with* $t_\phi = f_1 + f_2 + f_3$. *The dual of such a basis* $\lambda\beta$ *is given by*

$$\phi^* = (t_\phi \beta)^* = t_\phi^{-1} \beta^*.$$

Proof There are 35 sets of 3 distinct elements $\neq 0$ of $\mathbb{F}_8$, and those which are not among the 7 which are lines in the Fano plane are exactly those which are bases. In particular, (R', S', I') is not a basis. A trio (u, v, w) of three distinct elements $\neq 0$ is a basis if and only if in the Fano plane the 3 points do not form a line, i.e. if and only if $u + v + w \neq 0$.

The existence of normal basis in a finite field comes back to K. Hensel (1888), in a non-constructive manner. Even more is true: there is always a primitive normal basis [11]. There is not always an auto-dual basis in a finite field; for example, $\mathbb{F}_{16}$ has no such basis. But there exists such a basis in $\mathbb{F}_{2^n}$ if n is odd [12, p. 73, p. 129, ex. 3.77]; it is the case of $\mathbb{F}_8$. For $\mathbb{F}_8$ the basis (R, S, I) is precisely primitive normal and strictly auto-dual, and the only one in this case (with also, of course, (S, I, R) and (I, R, S)).

We can give an explicit proof of this last point, taking as in Proposition 6.1, a basis $\beta = (e_1, e_2, e_3)$, its dual $\beta^* = (e_1^* = e_2 \times e_3, e_2^* = e_3 \times e_1, e_3^* = e_1 \times e_2)$; then $\beta = \beta^*$ if and only if $\langle e_i, e_j\rangle = \delta_{i,j}$, if and only if

$$e_1 = e_2 \times e_3, \qquad e_2 = e_3 \times e_1, \qquad e_3 = e_1 \times e_2.$$

Then $\langle e_1, e_1\rangle = [e_1, e_2, e_3] = 1$, and so $e_1 \in \{R, S, I, 1\}$, and similarly $e_2, e_3 \in \{R, S, I, 1\}$. In fact, $e_1 \neq 1$ because $1 \times X \in \{R', S', I', 0\}$ (see the table in Proposition 4.2). So $\{e_1, e_2, e_3\} = \{R, S, I\}$.

With the same table in Proposition 4.2 we verify that $R' \times I' = 1$, $I' \times 1 = I'$, $R' \times 1 = R'$, and so r and r^* are dual bases (not strictly). The same is available for s and for i.

Up to a permutation of e_1, e_2 and e_3, the other case of duality is

$$e_1 = e_2 \times e_3, \qquad e_3 = e_3 \times e_1, \qquad e_2 = e_1 \times e_2.$$

In this case, $\langle e_1, e_1\rangle = 1$, $\langle e_2, e_2\rangle = 0$, $\langle e_3, e_3\rangle = 0$, i.e. $e_1 \in \{R, S, I, 1\}$, $e_2, e_3 \in \{R', S', I'\}$. If, for example, $e_2 = R'$ and $e_3 = S'$, then $e_1 = R' \times S' = 1$.

To conclude our proof, we have just to check that the different values of $t\beta$, with $t \in \mathbb{F}_8\backslash\{0\}$ and $\beta \in \{\kappa, r, s, i\}$ provide exactly the 28 possibilities of bases.

Finally, for the computation of the dual bases, we put $\underline{\phi} = \{f_1, f_2, f_3\}$, $\lambda\underline{\phi} = \{\lambda f_1, \lambda f_2, \lambda f_3\}$, and $\underline{\phi}^* = \{f_2 \times f_3, f_3 \times f_1, f_1 \times f_2\}$, in such a way that, with $\underline{\kappa} = \{R, S, I\}$, $\underline{r} = \{R', I', 1\}$, etc. and $\lambda \in \mathbb{F}_8\backslash\{0\}$, we have to verify that $(\lambda\underline{\kappa})^* = \lambda^{-1}\underline{\kappa}$, $(\lambda\underline{r})^* = \lambda^{-1}\underline{r}^*$, etc. Because of symmetry, only these two cases are to be checked. We do it with the tables for product and cross product given in the proof of Proposition 3.7 and in Proposition 4.2. □

Proposition 6.3 *If $\epsilon = (e_1, e_2, e_3)$ and $\phi = (f_1, f_2, f_3)$ are two bases, and if $u = u_1 f_1 + u_2 f_2 + u_3 f_3$, let $T(u)$ be the element with the same coordinates on ϵ:*

$$u = u_1 f_1 + u_2 f_2 + u_3 f_3, \quad T(u) = u_1 e_1 + u_2 e_2 + u_3 e_3$$

in such a way that

$$T(f_1) = e_1, \qquad T(f_2) = e_2, \qquad T(f_3) = e_3,$$
$$T^{-1}(e_1) = f_1, \qquad T^{-1}(e_2) = f_2, \qquad T^{-1}(e_3) = f_3,$$

then, with $v = v_1 e_1 + v_2 e_2 + v_3 e_3$, we have

$$T(u) = \mathrm{tr}\big(u f_1^*\big) e_1 + \mathrm{tr}\big(u f_2^*\big) e_2 + \mathrm{tr}\big(u f_3^*\big) e_3,$$
$$T^{-1}(v) = \mathrm{tr}\big(v e_1^*\big) f_1 + \mathrm{tr}\big(v e_2^*\big) f_2 + \mathrm{tr}\big(v e_3^*\big) f_3.$$

This transformation T is linear and if necessary more explicitly denoted by $T = T^{\epsilon\leftarrow\phi}$. Its matrix relative to ϕ is $\Theta = (\mathrm{tr}(f_i^\star e_j))$.

Furthermore, if coordinates of u on ϕ and ϵ are given, $u = \sum_j u_j e_j$ and $u = \sum_i u_i' f_i$, then the exchange of coordinates is given by composition with Θ:

$$u_i' = \sum_j \mathrm{tr}\big(f_i^\star e_j\big) u_j.$$

Proof It is an immediate application of Proposition 6.1. The matrix of T with source basis ϕ and target ϵ is I_3, and the matrix of T relative to ϕ is the matrix of Id with source basis ϵ and target ϕ. The last formula comes from the description of e_j on ϕ: $e_j = \sum_i \mathrm{tr}(f_i^* e_j) f_i$, etc. □

Proposition 6.4 *With the notations of Proposition 6.2, we consider the 3 linear transformations r°, s°, i°, sending κ to r, κ to s, and κ to i. Their matrices relatively to κ are—abusively—denoted only by r,s and i:*

$$r = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad s = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \qquad i = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$

and their polynomial forms are (cf. convention 1, and Proposition 3.12.):

$$\underline{r}(u) = R'u^4 + u^2 + I'u; \qquad \underline{s}(u) = S'u^4 + u^2 + R'u; \qquad \underline{i}(u) = I'u^4 + u^2 + S'u.$$

Proposition 6.5 *The multiplications by R, S or I, used in Proposition 5.7, and denoted by $R^\cdot$, $S^\cdot$, $I^\cdot$, are given by matrices denoted only by:*

$$R = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad S = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \qquad I = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

and the polynomial forms (cf. Proposition 3.12.):

$$\underline{R}(u) = Ru; \qquad \underline{S}(u) = Su; \qquad \underline{I}(u) = Iu.$$

Proposition 6.6 *The three cross products with R, S, I, i.e. $u \mapsto R \times u$, $u \mapsto S \times u$, $u \mapsto I \times u$, are given by the matrices:*

$$R^\times = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad S^\times = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad I^\times = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

with the polynomial forms (cf. Proposition 3.12.):

$$\underline{R^\times} u = Su^4 + Iu^2, \qquad \underline{S^\times} u = Iu^4 + Ru^2, \qquad \underline{I^\times} u = Ru^4 + Su^2.$$

Furthermore—as in the Lie algebra $\mathfrak{so}(3)$ of the Lie group SO(3)*—we have the commutator relations:*

$$[R^\times, S^\times] = I^\times, \qquad [S^\times, I^\times] = R^\times, \qquad [I^\times, R^\times] = S^\times.$$

Proposition 6.7 *The inverses of matrices r, s, i are*

$$r^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \qquad s^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \qquad i^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

with the polynomial forms

$$\underline{r^{-1}}(u) = Ru^4 + Iu^2 + Ru; \qquad \underline{s^{-1}}(u) = Su^4 + Ru^2 + Su; \qquad \underline{i^{-1}}(u) = Iu^4 + Su^2 + Iu,$$

and we have

$$\underline{r^{-1}} + \underline{s^{-1}} + \underline{i^{-1}} = \mathrm{tr}\,.$$

Proof We can use Propositions 3.10 and 3.12, or directly compute $r^{-1} = r^6$, $s^{-1} = s^6$, $i^{-1} = i^6$. □

Proposition 6.8 *We have*

$$\underline{r} + \underline{s} + \underline{i} = (-)^2,$$
$$R = ir^2, \qquad S = rs^2, \qquad I = si^2.$$

Proof The formula $R = ir^2$ was given in [7, Proposition 15, p. 152]. It is easy to check, as well as the formula for $(-)^2$. □

Proposition 6.9 *Given $\phi = (f_1, f_2, f_3) = t_\phi \beta$ one of the 28 bases of $\mathbb{F}_8$, as in Proposition 6.2, let ϕ° be the linear map sending κ to ϕ, and let F its matrix relative to κ. Then F is of the form HB, with $H \in \{I_3, R^\cdot, S^\cdot, I^\cdot, R'^\cdot, S'^\cdot, I'^\cdot\}$ and $B \in \{I_3, R^\circ, S^\circ, I^\circ\}$, and every element M of $\mathrm{GL}_3(\mathbb{F}_2)$ could be written in a unique way as a composition of such an HB and a permutation P:*

$$M = HBP.$$

Proof It is [7, Proposition 17, p. 153]. □

Proposition 6.10 *The group $\mathrm{GL}_3(\mathbb{F}_2)$ is generated by $r^\circ, s^\circ, i^\circ$, as well as by their inverses.*

Proof It is as in [7, Proposition 18, p. 154], a consequence of Proposition 6.9. □

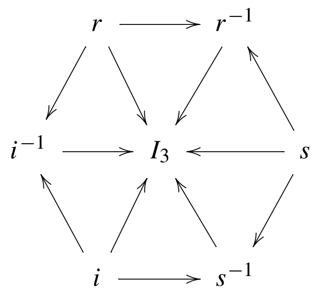
Proposition 6.11 *With the hypothesis and notations of Proposition 5.6, every function in $\mathbb{P}_8$ can be written with canonical boolean operations $\wedge$ and $\neg$, and the three linear maps $r^\circ, s^\circ, i^\circ$ with matrices r, s, i.*

Proof With Proposition 6.8, in Proposition 5.7 we can replace the operations $(-)^2$, $R^\cdot, S^\cdot, I^\cdot$ by $r^\circ, s^\circ, i^\circ$. □

Proposition 6.12 *With the hypothesis and notations of Proposition 5.6, every function in $\mathbb{P}_8$ can be written with canonical boolean operations $\wedge$ and $\neg$, and the three linear maps $r^{\circ-1}, s^{\circ-1}, i^{\circ-1}$ with matrices r^{-1}, s^{-1}, i^{-1}.*

Proof It results from Proposition 6.11 and the fact that the r, s, i are formulable with their inverses, in $\mathrm{GL}_3(\mathbb{F}_2)$ (Proposition 6.10). □

Remark 6.13 If in the picture of the hexagon in Definition 2.1 we emphasize that $R' = R^{-1}$, then now we get an analogous decoration of the hexagon by elements of $\mathbb{P}_8$:



7 A, B, C and $R^\times, S^\times, I^\times$ Borromean Generations of $\mathrm{GL}_3(\mathbb{F}_2)$ and of $\mathbb{P}_8$

7.1 Generation by A, B, C

Proposition 7.1 *In Proposition 6.10, we introduce, with $r^6 = r^{-1}$, etc.*

$$r^t = rir^6, \qquad s^t = srs^6, \qquad i^t = isi^6;$$

they are the transposed matrices of r, s, i, and we define

$$A = r^t i^t, \qquad B = s^t r^t, \qquad C = i^t s^t.$$

These A, B, C are the matrices of 3 transvections given by

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A^2 = B^2 = C^2 = I_3,$$

$$r = ACB, \qquad s = BAC, \qquad i = CBA,$$

and the transposed matrices of A, B, C are

$$A^t = (CB)^2, \qquad B^t = (AC)^2, \qquad C^t = (BA)^2,$$

and $\mathrm{GL}_3(\mathbb{F}_2)$ is generated by A, B, C.

Furthermore, we have the polynomial forms:

$$\underline{A}(u) = R'u^4 + Iu^2 + Su; \qquad \underline{B}(u) = S'u^4 + Ru^2 + Iu; \qquad \underline{C}(u) = I'u^4 + Su^2 + Ru.$$

Proof So the borromean structure of $\mathrm{GL}_3(\mathbb{F}_2)$ can act on the set $\mathbb{F}_8$, which is also a boolean algebra, and so we obtain another borromean presentation of $\mathbb{P}_8$, as in the next proposition. For polynomial forms we use Proposition 3.10. □

Proposition 7.2 *We have*

$$\underline{A}+\underline{B}+\underline{C}+\mathrm{Id}_{\mathbb{F}_8}=(-)^2,$$
$$R=(CB)^2A(CB), \qquad S=(AC)^2B(AC), \qquad I=(BA)^2C(BA).$$

Proof For $(-)^2$ it is immediate by addition of the 3 matrices A, B, C, and for $R^{\cdot}$ from Proposition 6.8 we have $R = ir^2$, and with Proposition 7.1, $R=(CB)^2A(CB)$. □

Proposition 7.3 *The commutators of the* A, B, C *are*:

$$[A,B]=AB+BA=\begin{bmatrix}0&0&0\\0&0&0\\1&0&0\end{bmatrix},$$
$$[B,C]=BC+CB=\begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix},$$
$$[C,A]=CA+AC=\begin{bmatrix}0&0&0\\0&0&1\\0&0&0\end{bmatrix},$$

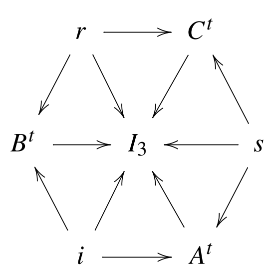
and

$$\underline{[A,B]}+\underline{[B,C]}+\underline{[C,A]}=(-)^4.$$

Proposition 7.4 *With the hypothesis and notations of Proposition* 5.6, *every function in* $\mathbb{P}_8$ *can be written with canonical boolean operations* $\wedge$ *and* $\neg$, *and the three linear involutive transvections given by* A, B, C.

Proof From Propositions 6.11 , 7.1, and 7.2. □

Remark 7.5 If in the picture of the hexagon in Definition 2.1 we emphasize that $R' = SI$, then we get an analogous decoration of the hexagon by elements of $\mathbb{P}_8$:



7.2 Generation by $R^\times$, $S^\times$, $I^\times$

Proposition 7.6 *We have*

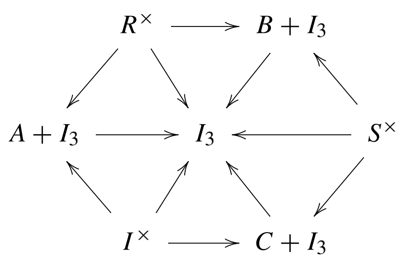
$$R^\times S^\times = A + I_3, \qquad S^\times R^\times = [B, C],$$
$$S^\times I^\times = B + I_3, \qquad I^\times S^\times = [C, A],$$
$$I^\times R^\times = C + I_3, \qquad R^\times I^\times = [A, B].$$

Proof An immediate verification with matrices. □

Proposition 7.7 *With the hypothesis and notations of Proposition 5.6, every function in $\mathbb{P}_8$ can be written with canonical boolean operations $\wedge$ and $\neg$, and the three cross products given by $R^\times$, $S^\times$, $I^\times$.*

Proof A consequence of Propositions 7.4 and 7.6. □

Remark 7.8 If in the picture of the hexagon in Definition 2.1 we emphasize that $R' = SI$, then now we get an analogous decoration of the hexagon by elements of $\mathbb{P}_8$:



8 Conclusion: Hexagonal Presentations of an 8-Valued Logic, Going from a Decoration by $\mathbb{F}_8$ to Decorations by $\mathbb{P}_8$

We consider that the logic of an object W 'is' the organization of $\mathbb{P}(W)$ the Post–Malcev algebra of functions of all arities on this object, $f : W^k \to W$, and especially the logic of $\underline{8} = \{\underline{0}, \underline{1}, \underline{2}, \underline{3}, \underline{4}, \underline{5}, \underline{6}, \underline{7}\}$ or of the cube $\{0, 1\}^3$ is the organization of functions $f : (\{0, 1\}^3)^k \to \{0, 1\}^3$, i.e. the algebra $\mathbb{P}(\{0, 1\}^3) = \mathbb{P}_8$.

At first we have shown that the set $\{0, 1\}^3$ as a field $\mathbb{F}_8$ can be presented as a Fano plane plus a zero, as a hexagon (cf. Definition 2.1 and Sect. 3), and we began with a decoration of an hexagon by elements of $\mathbb{F}_8$.

Then studying the arithmetic and the geometry on $\mathbb{F}_8$, we proved that $\mathbb{P}(\mathbb{F}_8) = \mathbb{P}_8$ can be generated by a boolean calculus with conjunctive avatars.

After that, we proved that $\mathbb{P}_8$ itself admits hexagonal generations, by canonical boolean operations plus (r, s, i) or plus (r^{-1}, s^{-1}, i^{-1}), or plus A, B, C, or $R^\times$, $S^\times$, $I^\times$, and we have drawn corresponding decorations of the hexagon.

Now to conclude, forgetting our arithmetical and geometrical tools and intermediary arguments, as the different ways of thinking with hexagons and the avatars, we express our main result from Proposition 7.7 in layman's terms:

Theorem 8.1 *Given a set with* 8 *elements, represented as* $\{0,1\}^3$, *the set of all the functions* $f : (\{0,1\}^3)^k \to \{0,1\}^3$, *for all* $k \in \mathbb{N}$, *is generated by composition of the* 6 *following functions* (*modulo* 2) *of arities* 2 *and* 1:

$$\big((x,y,z),(x',y',z')\big) \mapsto \big(x+x', y+y', z+z'\big);$$

$$\big((x,y,z),(x',y',z')\big) \mapsto \big(x.x', y.y', z.z'\big);$$

$$(x,y,z) \mapsto (x+1, y+1, z+1);$$

$$(x,y,z) \mapsto (0,z,y); \qquad (x,y,z) \mapsto (z,0,x); \qquad (x,y,z) \mapsto (y,x,0).$$

References

1. Béziau, J.-Y.: The power of the hexagon. Log. Univers. **6**(1–2), 1–43 (2012)
2. Blanché, R.: Sur l'opposition des concepts. Theoria **19** (1953)
3. Blanché, R.: Structures Intellectuelles. Vrin, Paris (1966)
4. Dickson, L.E.: Linear Groups. Dover, New York (1958) (1st ed. 1900)
5. Guitart, R.: L'idée de logique spéculaire. In: Journées Catégories, Algèbres, Esquisses, Néo-esquisses, Caen, 27–30 September 1994
6. Guitart, R.: Moving logic, from Boole to Galois. In: Colloque International "Charles Ehresmann: 100 ans", 7–9 October 2005, Amiens. Cahiers Top Géo Diff Cat, vol. XLVI-3, pp. 196–198 (2005)
7. Guitart, R.: Klein's group as a Borromean object. In: Cahiers Top. Géo. Diff. Cat, vol. L-2, pp. 144–155 (2009)
8. Guitart, R.: A hexagonal framework of the field $\mathbb{F}_4$ and the associated Borromean logic. Log. Univers. **6**(1–2), 119–147 (2012)
9. Heisler, J.: A characterization of finite fields. Am. Math. Mon. **74**, 537–538, 1211 (1967)
10. Lau, D.: Function Algebras on Finite Sets. Springer, Berlin (2006)
11. Lenstra, H.W. Jr., Schoof, R.J.: Primitive normal basis for finite fields. Math. Comput. **48**, 217–231 (1987)
12. Lidl, R., Niederreiter, H.: Introduction to Finite Fields and Their Applications. C.U.P., Cambridge (1994)
13. Malcev, A.I.: Iterative algebra and Post's varieties. Algebra Log. **5**, 5–24 (1966) (Russian)
14. Moore, E.H.: Mathematical Papers, Chicago Congress of 1893, pp. 208–242; Bull. Am. Math. Soc., December 1893
15. Ribenboim, P.: L'arithmétique des Corps. Hermann, Paris (1972)
16. Sesmat, A.: Logique: 1. les Définitions, les Jugements. 2. Les Raisonnements, la Logistique. Hermann, Paris (1951)

R. Guitart (✉)
IMJ-PRG Université Paris Diderot, Bâtiment Sophie Germain, 75013 Paris, France
e-mail: rene.guitart@orange.fr

The Move from One to Two Quantifiers

Wilfrid Hodges

Abstract Ibn Sīnā (Persian, 980–1037) made a dramatic extension to Aristotle's syllogistic by adding quantifiers over times or situations, thus introducing multiple and mixed quantification. The extension is unlike anything in the Latin Scholastic logic, but related extensions to syllogistic appear later in work of Leibniz and of Peirce's student Mitchell. Ibn Sīnā's version was the most integrated and systematic of these three, but at the same time it was the furthest from modern perspectives. We examine from a modern point of view the limitations, proof-theoretic and otherwise, which Ibn Sīnā's highly original introduction of multiple quantification failed to overcome.

Keywords Ibn Sina · Peirce · Syllogism · Multiple quantifiers · Mixed quantification · Scope

Mathematics Subject Classification (2000) 01A30 · 03B65

1 Introduction

By adopting the motif of 'Universal Logic', Jean-Yves Béziau [6] has performed a valuable service. He has encouraged us all to look at the larger features of logic, those that are not limited to a single subdiscipline of logic or to a single tradition. In honouring him I will call attention to a topic that lies on the frontier between modern logic and the old Aristotelian logic, namely sentences with two quantifiers.

One of the most conspicuous differences between Aristotle's logic and modern logic is that Aristotle recognised only one quantifier per sentence. In the turbulent second half of the 19th century, when the foundations were being laid for today's logic, it doesn't appear that anybody was much exercised about the number or type of quantifiers in Aristotle's syllogistic. But today you may well have heard it said that quantifier complexity was the big difference between the old logic and the new.

Discoveries in the history of logic give us a chance to test this viewpoint. We now know that Ibn Sīnā, the early 11th century Arabic–Persian logician and polymath (980–1037), known to the Latins as Avicenna, had a sophisticated extension of Aristotle's syllogistic that included two or more quantifiers per sentence. Was it the big breakthrough? Apparently not. But it gives us an excellent testbed to check what else was needed to get where we are today. We can identify a number of things that stand between Ibn Sīnā's logic and ours. Mostly they are things not mentioned in histories of logic, which shows that we are still some way from understanding the levers that drive change and progress in logic.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_9

We owe an apology to the memory of Ibn Sīnā for treating his logic in this way, as if it was a step in some inevitable progress towards the European logic of the 20th and 21st centuries. His logic did generate a distinctive style of formal logic that is still widely taught in the Arabic and Persian worlds. It compares very favourably with what was available in post-medieval Europe before the 1880s. But my concern here is with what Ibn Sīnā himself said, and not with its effects in the later history of logic.

Ibn Sīnā will be our main topic, because his work is both the most significant and the least well known advance into multiple quantification before Frege. But for a fuller picture we will also visit Galen and Alexander of Aphrodisias in the second century AD, and Oscar Mitchell and C.S. Peirce in the late 19th century.

I warmly thank Johan van Benthem, Saloua Chatti, Ben Morison, Zia Movahed, Stephen Read and Dag Westerståhl for their helpful comments, but blame me for any errors of fact or judgment.

2 Aristotle

Oversimplifying a little, Aristotle's non-modal logic was a study of the inference relations between four kinds of sentence, known to the later Latins by the first four vowels:

$$\begin{array}{ll} \text{(a)} & \forall x(Cx \to Bx) \text{ ('Every } C \text{ is a } B\text{')} \\ \text{(e)} & \forall x(Cx \to \neg Bx) \text{ ('No } C \text{ is a } B\text{')} \\ \text{(i)} & \exists x(Cx \wedge Bx) \text{ ('Some } C \text{ is a } B\text{')} \\ \text{(o)} & \exists x(Cx \wedge \neg Bx) \text{ ('Not every } C \text{ is a } B\text{')} \end{array} \tag{1}$$

I will refer to sentences of these forms as *(Aristotle's) assertoric sentences*. The theory of inference relations between them is *assertoric syllogistic*. Aristotle also allows sentences where the quantifier is removed and a single element named, as in 'c is a B'; these are known as *singular assertoric sentences*. Later Aristotelian logicians sometimes treated 'c is a B' as if it said 'Everything that is c is a B', to bring it within the ambit of the quantified sentences.

A typical example of a valid inference form using assertoric sentences is the following, known to the Latins as *Barbara*; the vowels of the name indicate that we have three 'a' sentences.

Every C is a B.
Every B is an A. (2)
Therefore every C is an A.

Inference forms like this are called *moods*. A mood is a pattern for a single inference step.

The classical terminology is not very systematic; in particular, we can say that a syllogism is *valid*, meaning that the conclusion does follow from the premises, but also we can say that a sequence of sentences *is a syllogism* meaning that it's a valid one. Likewise one can speak of a *valid mood*, but also speak of 'the moods' meaning only the valid ones. So the terminology leaves the status of invalid syllogisms or moods in limbo. A proper systematisation would introduce the notion of a 'putative syllogism' or a 'pre-syllogism', and pick out syllogisms as those putative syllogisms that are valid. But I will try to get by without introducing extra terminology along these lines.

The main practical application of assertoric syllogistic takes the following form: we show that an inference expressed in natural language is valid, by demonstrating that it can be paraphrased into one of the valid moods. For this the argument needs to be split into two premises and a conclusion, and parts of the premises and conclusion have to be identified as *terms*, i.e. the parts represented by the letters C, B, A. Although most of the examples offered by Aristotle and later Aristotelian logicians are fairly simple, in principle a term could be a piece of text of any length and complexity. In particular, a term can contain as many quantifiers as you like.

One of the best sets of examples of complex terms is in three sections of Ockham's *Summa Logicae* dealing with 'syllogismi ex obliquis'. For example—and excuse my stumbling attempt to represent his stumbling Latin (see [29, p. 386, iii.1.9 27f]):

> Nobody is seen by a certain donkey;
> everything that laughs sees somebody; (3)
> therefore nothing that laughs is a certain donkey.

Note the two quantifiers in the second premise, one universal and the other existential. It takes some thinking to work out which of these quantifiers goes to the beginning of the formal sentence and which is hidden inside a term. Try starting from 'Not every donkey sees somebody' and aim for 'Not every donkey is a laugher'.

Examples like Ockham's are neat brainteasers, but they do nothing to increase the power of assertoric syllogistic. We will see below that what seem at first to be significant extensions of assertoric syllogistic sometimes turn out to reduce to assertoric syllogisms by suitable paraphrases.

3 Adding Just Universal Quantifiers

Aristotle created his assertoric syllogistic during the fourth century BC. Its development during the next 400 years is obscure, but we know that Aristotle's *Prior Analytics* was available as a teaching text in the Roman Empire by the first century AD. By the mid second century various accessory books were appearing, among them elementary handbooks of logic by Apuleius and Galen, and a detailed commentary on the *Prior Analytics* by Alexander of Aphrodisias. These books acknowledged the existence of other traditions in logic besides that of Aristotle. Galen is known to have written several works on logic, but the only one to survive is his eclectic introductory text *Institutio Logica*, which has reached us via a single very poor medieval manuscript.

Aristotle's assertoric sentences have just one quantifier each. They can talk about just one thing at a time, not about relationships between more than one thing (unless the relationships are buried inside terms, where they are unreachable by the syllogistic rules). In Chap. xvi of his *Institutio Logica*, Galen observes that 'arithmeticians and calculators' use arguments which depend essentially on comparing more than one thing. He takes an example from Euclid's *Elements*:

> Since things equal to the same thing are also equal to one another, and the first and second have been shown equal to the third, the first would thus be equal to each of them. ([18, p. 39f, xvi.6]; trans. Kieffer [19, p. 50]) (4)

The text is not entirely secure, but Galen appears to be challenging the Aristotelian logicians to show how to carry out such arguments using assertoric syllogisms ([18, p. 38, xvi.1]; Kieffer [19, p. 49]).

Modern commentators note a passage in Alexander of Aphrodisias' commentary on the *Prior Analytics* that can be read as a response to this challenge, though Alexander himself doesn't phrase the matter in those terms. Here from a modern perspective is what he does. He notes that just as we can talk about single elements, we can talk about ordered pairs. (He uses a plural rather than a dual, but the context and his use of *allēlois* show that he is thinking of pairs.) All of Galen's examples are either unquantified or universally quantified; we can pack two individuals into a single pair, and two universal quantifiers into a single quantifier over pairs. Or indeed we can do the same thing with triples or quadruples; some of Galen's examples need quadruples.

The most complicated example that Alexander himself gives involves three elements. It can be paraphrased as follows:

Every triple where the first is greater than the second and the second is greater than the third is a triple where the first is greater than the third.
ABC is a triple where the first is greater than the second and the second is greater than the third. (5)
Therefore ABC is a triple where the first is greater than the third.

(For the original see [2, 344.20–27], trans. Mueller [4, p. 28].) This is a proof of a particular instance of the transitivity of 'greater than', from an axiom stated as first premise. It forms a syllogism in mood *Barbara* with two singular sentences. In view of examples like this one, I will refer to this use of ordered sets of individuals as *Alexander's trick*.

You might object that the conclusion we want is that A is greater than C, not that ABC is a triple where the first is greater than the third. This would be a shallow objection. All Aristotelian logicians up to the late 19th century reckoned only to validate single inference steps, one at a time. Rephrasing before an inference step, or after it, or between two inference steps, was entirely normal procedure. (The rephrasing might be a paraphrase, or it might—as here—discard some of the content.) Not many logicians commented on this feature, but at least Ibn Sīnā and Leibniz noted it as a fact about logic. Frege commented too, vehemently; for Frege the steps of rephrasing between inference steps were pieces of reasoning not under the control of logic, and logic had to be rejigged to remove these gaps. I discussed this feature of pre-19th-century logic in more detail in [12] §5, under the name of *local formalising*, and there is a concrete example in (31) below. As Frege correctly observed, local formalising was a gigantic obstacle between Aristotelian logic and modern logic. The introduction of multiple quantifiers by itself does nothing to remove the obstacle.

As (5) above illustrates, Alexander's trick allows you to have multiple homogeneous quantifiers (i.e. multiple quantifiers that are all universal or all existential) for free in syllogistic. The set of valid argument forms need not be affected at all; the entire work of extending the logic can be shifted into the procedures for applying the argument forms.

Nevertheless, Alexander's trick did allow more flexibility in applications of syllogisms. I haven't followed it through the literature thoroughly, but it seems to have been a stock in trade of Aristotelian logic from Alexander onwards, and quite possibly mathematical writers were another source besides Alexander. Ibn Sīnā used the trick freely. John Stuart Mill also used it repeatedly in his sample geometrical 'train of reasoning' to show that

the angles at the base of an isosceles triangle [are] equal.
(see [20, pp. 141–143]) (6)

Note that even his statement of the theorem quantifies universally over pairs of angles.

See Morison [22, pp. 107f] for a different take on Alexander's trick and the argument (5). Note also Ian Mueller's puzzling remark:

The development of the logic of relations in the nineteenth century has made clear that Alexander is barking up the wrong tree here. (Mueller [3, p. 17]) (7)

Alexander's trick seems to me pretty anodyne, not much of a step up any tree. But of all later logicians, the person who made the deepest use of Alexander's idea was C. S. Peirce; he used it precisely to advance the logic of relations. We will see how in Sect. 6 below.

4 Ibn Sīnā's Double Quantification

Ibn Sīnā starts from Aristotle's assertoric a-, e-, i- and o- sentences. He notes that when we put a sentence into one of these forms, we may find ourselves packing quite a lot into the predicate, and the structure of what we put there may be relevant to inferences. He notes also that some of what goes there may not even be stated explicitly, but we all understand that it is meant. In the examples that he first gives us, the extra information packed into the predicate is about the time at which the predicate holds. So there is a second variable in the predicate, a variable ranging over times. (In modern notation we can use Latin letters x, etc. for the *object variables* which are quantified over in the assertoric forms, and Greek letters τ, etc. for the *time variables* needed for expressing the extra information identified by Ibn Sīnā.)

For example, Ibn Sīnā claims:

Everything that moves is a body. (8)

But we all understand, first, that a thing which is once a body is always a body, at least while it still exists. And second, we all understand that (8) wouldn't be meant just as a statement about the present moment. It expresses a truth about anything that moves at any time.

So we write $E(x, \tau)$ for 'x exists at time τ', and we formalise the statement (8) as

$$\forall x\big(\exists\tau C(x, \tau) \to \forall\tau\big(E(x, \tau) \to B(x, \tau)\big)\big). \tag{9}$$

(Here $C(x, \tau)$ is 'x moves at time τ', $B(x, \tau)$ is 'x is a body at time τ'.) Ibn Sīnā claims that no property can truthfully be ascribed to a thing that doesn't exist. So there are some background truths about the relation E, which we can write as axioms of the form $\forall x\forall\tau(B(x, \tau) \to E(x, \tau))$.

This illustrates Ibn Sīnā's general recipe for turning an assertoric sentence into a doubly quantified one. We take the assertoric form; if it is negative we count the negation as part of the predicate. We add a time variable τ to both subject and predicate. The subject takes the quantifier $\exists\tau$. The predicate, say $B(x, \tau)$, is buried inside a more complicated formula $\theta(B, x)$ in which B occurs only positively. The original assertoric sentence and the formula $\theta(B, x)$ together determine the doubly quantified sentence. I will refer to the extra information contained in θ as the *Avicennan adjunction*. Ibn Sīnā himself describes it as being 'attached to' the predicate, or as 'added to' the predicate, or as a 'condition on' the predicate; sometimes he refers to it briefly as the 'modality'.

Just as we classify the assertoric sentence as a-, e-, i- or o-, we can classify the Avicennan adjunctions. Ibn Sīnā himself didn't use the letters a-, e-, i- and o-; instead he classified the assertoric sentences as ‘universal affirmative’, etc. But he did propose some names for the Avicennan adjunctions, and happily they provide us with some consonants to match the vowels of the classification of assertorics. For example, the Avicennan adjunction in the case (8) above can be written

$$\forall\tau\big(E(x,\tau) \rightarrow B(x,\tau)\big), \tag{10}$$

and Ibn Sīnā calls sentences with this adjunction on their predicate *ḍarūrī*, so we will call them d. The sentence (8) is an example of an a–d sentence, with Aristotelian classification a and Avicennan classification d.

Likewise we have e–d, i–d and o–d sentences. An example of an o–d sentence might be

There is a person who never writes. (11)

This comes from the o-sentence ‘There is a person who doesn’t write’, counting ‘doesn’t write’ as the predicate that goes for the B in $\theta(B,x)$.

To cut a long story short, here are the four main forms in the Avicennan classification. Note that in two cases the form of θ depends on the subject term C.

(d) x is a B for all the time while x exists.

$$\forall\tau\big(E(x,\tau) \rightarrow B(x,\tau)\big). \tag{12}$$

(ℓ) x is a B for all the time while x satisfies the subject term.

$$\forall\tau\big(C(x,\tau) \rightarrow B(x,\tau)\big). \tag{13}$$

(m) x is a B at some time while x satisfies the subject term.

$$\exists\tau\big(C(x,\tau) \wedge B(x,\tau)\big). \tag{14}$$

(t) x is a B at some time while x exists.

$$\exists\tau\big(E(x,\tau) \wedge B(x,\tau)\big). \tag{15}$$

This gives 16 sentence forms, which we can label e–m, i–ℓ, etc. Note that a–d sentence entails the corresponding ℓ sentence, an ℓ sentence entails the corresponding m sentence, and an m sentence entails the corresponding t sentence. (For this last we need the background fact that everything exists at some time.) Ibn Sīnā’s best account of all this is in his *Easterners* [16], but it can also be found in the first book of his *Qiyās* [17].

I said ‘main forms’ because the whole picture that Ibn Sīnā paints is much less tidy than this four-way Avicennan classification suggests. The chief complication is an asymmetry that Ibn Sīnā notices between the existentially quantified and the universally quantified sentences, which leads him to define several other kinds of existentially quantified forms whose logic he never develops. (Later Arabic logicians worked hard to develop some of these forms.) The reasons for this asymmetry are very important for understanding Ibn Sīnā’s position and how it could have developed; we come back to them in Sect. 9 below. But meanwhile I will refer to the sentences of the forms a–d, etc., with an Aristotelian classification and an Avicennan one, as *two-dimensional sentences*. (The name is stolen from Mitchell, see Sect. 6 below.)

There are a further two possible directions that Ibn Sīnā indicates. One is to take the quantifier in θ to the front of the sentence, so that the main quantification is over times. In his examples it becomes rapidly clear that in fact he is thinking in terms of quantification over possible situations—or does he really think that (for example) there was or will be a time when nothing is coloured black (*Qiyās* [17, 133.7])? There are obvious links here to distinctions one might make between de dicto and de re modalities; see Zia Movahed [23] for a study of these notions in Ibn Sīnā and some later Arabic logicians.

The other direction is a suggestion that Ibn Sīnā makes briefly at *Qiyās* [17, 41.12f]; some passages in *Burhān* (his commentary on the *Posterior Analytics*) may also have it in mind. The suggestion is that the $E(x, \tau)$ or $C(x, \tau)$ might be replaced by some other 'affirmative' condition. The suggestion is certainly interesting and shows a very welcome flexibility of thinking. We will come back to it in Sect. 7 below.

5 Avicennan Syllogisms

An *Avicennan syllogism* is the same as an assertoric syllogism except that the assertoric sentences are replaced by two-dimensional ones. Under what circumstances is a two-dimensional syllogism valid? The following fact can be proved.

Fact 5.1 *Let T be a set of two-dimensional sentences and ϕ a two-dimensional sentence. For T to entail ϕ the following are necessary and sufficient*:

(a) *If we remove the Avicennan adjunctions so that what remains is a set T' of assertoric sentences and an assertoric sentence ϕ', then there is a valid assertoric syllogism with premises T' and conclusion ϕ'.*
(b) *T and ϕ meet a further condition which refers only to the classification of the Avicennan adjunctions in ϕ and the sentences of T, not to the Aristotelian classification of the sentences involved* (*and in fact not to the 'figure' either*).

With his lack of metatheory there is no way that Ibn Sīnā himself could have proved this fact, even for two-premise syllogisms. But of course he could have formulated it on the basis of experience, and verified it for individual arguments. He always proceeds as if he knew it.

Since the condition (b) is independent of the Aristotelian classifications and the figure, we can state it for one single mood, and this determines its form for all other moods. The simplest mood to use is the one Ibn Sīnā himself chooses, namely *Barbara*. To state the condition it suffices to say what are the optimally valid two-dimensional versions of *Barbara*, where 'optimally valid' means that the premises can't be weakened or the conclusion strengthened without losing validity. Figure 1 is not a quotation from Ibn Sīnā. It was calculated by modern methods [15] on the basis of Ibn Sīnā's own definitions. He was certainly aware of its contents, apart from some probable haziness about the difference between (iii) and (iv).

The forms (i)–(v) of *Barbara* are only the optimally valid ones, and Ibn Sīnā does consider some others that are not optimally valid.

It will help to have examples of the five forms (i)–(v).

For (i) Ibn Sīnā offers the example (*Qiyās* 203.10):

Every human moves sometimes;
every moving thing is a body throughout its existence (16)
therefore every human is a body throughout its existence.

Fig. 1 The optimal Avicennan moods for *Barbara*

	minor	major	conclusion
(i)	t	d	d
(ii)	d	ℓ	d
(iii)	ℓ	ℓ	ℓ
(iv)	m	ℓ	m
(v)	t	t	t

One example that Ibn Sīnā gives for (ii) is

> Snow is coloured white throughout its existence;
> everything coloured white emits wide-spectrum light so long as it is coloured white; (17)
> therefore all snow emits wide-spectrum light throughout its existence.
> (*Qiyās* 129.1f)

Because of a connection with some earlier Peripatetic discussions, I will call (ii) the *Theophrastan* mood.

For (iii), here is an example made up from sentences that Ibn Sīnā uses in other contexts:

> Every writer moves all the time while writing;
> every moving thing changes all the time while moving; (18)
> therefore every writer changes all the time while writing.

Likewise an example for (v) made up from sentences that he uses is:

> Every human breathes in sometimes;
> everything that breathes in breathes out sometimes; (19)
> therefore every human breathes out sometimes.

There remains (iv), which for modern logicians is probably the most interesting mood. Sadly Ibn Sīnā lets us down badly here; it's not even clear that he identifies this as an interesting form. So the following example is mine, not his (except for the minor premise, which is Ibn Sīnā's from *Qiyās* [17, 22.12]):

> Everyone who travels from Tehran to Baghdad visits Kermanshah during the travel;
> everybody who visits Kermanshah is in the birthplace of Doris Lessing throughout the time while they are visiting Kermanshah; (20)
> therefore everybody who travels from Tehran to Baghdad is in the birthplace of Doris Lessing at some point during their travel.

Let us call this mood the *Kermanshah* mood. Section 7 below will explain the reason for paying special attention to the Theophrastan and Kermanshah moods.

6 Mitchell and Peirce

We leave Ibn Sīnā for a moment and fast forward to Johns Hopkins in Baltimore in 1883. Frege had published *Begriffsschrift* in 1879, and by 1883 there were people in Johns

Hopkins who knew of its existence, and of Schröder's review of it, but nobody in Johns Hopkins seems to have been aware of its contents. In 1880, the lead logician in Johns Hopkins, Charles S. Peirce, had published a paper [25] in which, among other things, he developed the rudiments of propositional logic. Near the end of the paper Peirce added some remarks about symbolism for a calculus of relations. For example, if ℓ means 'lover of' and s means 'servant of', then he offered $\ell \circ_s$ for 'whatever is not a non-lover only of a servant of'. So clearly there are some subtleties here, but he gives us nothing along the lines of first-order relational logic.

In 1883, the logic group at Johns Hopkins published a volume 'Studies in Logic'. The volume contained a chapter by Oscar H. Mitchell entitled 'On a new algebra of logic' ([21]; see Dipert [9] on Mitchell himself). In this chapter Mitchell proposes to extend syllogisms in a way similar to Ibn Sīnā's extension above. He notes that in an assertoric sentence the predicate can be taken to contain a reference to time, and he suggests ways of adding a second quantification to express something about the time. So his predicate term is one that today we would write with two variables, $B(x, \tau)$, just like Ibn Sīnā's. But from this point on he follows a different path. His subject term refers only to the objects, but he introduces a second subject term to deal with the times. In his preferred notation, the first subject term is U, the second is V and the predicate term is F. There are separate quantifiers over the two subject terms. These quantifiers are written as subscripts to F. The first quantifier is written 1 if it is universal and u if it is existential; the second again 1 for universal and v for existential. Mitchell is aware that there is a difference between 'For every object in U there is a time in V such that' and 'There is a time in V such that for every object in U'. He introduces an ad hoc notation for this, writing v' instead of v to show that only a single time is involved.

Unless I miss something, Mitchell's notation becomes incoherent at this point. He explains that F can stand for 'a polynomial function of class terms'. As illustration he uses b for 'the Browns' and i for 'ill', so that the polynomial $(\bar{b} + i)$ means 'either not a Brown or ill', i.e. 'if a Brown then ill'. The subscripts then tell us where to add quantifiers, so that, for example, $(\bar{b} + i)_{11}$ means 'Every one of the Browns was ill during every part of the year'. But how we find 'year' in this notation is obscure to me; maybe it's supplied separately by explaining what V is, but in that case what is U?

In spite of the incoherence of the notation, Mitchell's general drift is clear. He is constructing a logic in which we can have sentences of forms like

$$\begin{aligned} &\forall u \forall v (u \in U \wedge v \in V \to Fuv), \\ &\exists u \forall v \big(u \in U \wedge (v \in V \to Fuv)\big), \\ &\exists v \forall u \big(v \in V \wedge (u \in U \to Fuv)\big). \end{aligned} \tag{21}$$

But he has no proper variables, and he has no sensible notation for indicating the relative scopes of the quantifiers. In fact, in his English explanations of the scopes, he insists on putting U before V regardless of the scopes, for example 'every part of U, during the same part of V, is F'.

Clearly, Mitchell is not operating on the same level as Ibn Sīnā. But the curious fact is that just two years after the publication of Mitchell's paper, Peirce [26] presents a logic equivalent to full first-order logic, and he gives major credit to Mitchell's paper for this breakthrough. The editors of the Indiana edition of Peirce's works include a very revealing preparatory note of Peirce from summer 1884 [27], where he sets out what he has learned from Mitchell. The opening words of the note are:

> I begin by briefly restating the principles of the logical algebra, in the light of conceptions derived especially from the study of the important memoir of Professor O.H. Mitchell entitled "On a New Algebra of Logic". (22)

After a few remarks on propositional logic, Peirce turns to quantification.

The first thing he attributes to Mitchell is to separate the quantification from the rest of the sentence. (As we saw, Mitchell separates it by putting it into a subscript.)

> We now pass from the consideration of a single individual to that of the whole universe; and first we have that rude distinction of "some" and "all" which may be said to discriminate logic from mathematics. After the whole Boolian school had for thirty years been puzzling over the problem of how to take account of this distinction in their notation, without any satisfactory result, Mr. Mitchell, by a wonderfully clear intuition, points out that what is needed is to enclose the whole Boolian expression in brackets, and to indicate to what proportion of the universe it refers by exterior signs. (23)

The key point seems to be that the quantification is applied to the *whole of the rest of the sentence*. Frege achieved the same thing in *Begriffsschrift* by writing a quantifier at the top of its formula, so that its scope is everything below it.

The second thing that Peirce attributes to Mitchell is an approach to 'relative logic' by considering 'propositions of two dimensions'. Here Peirce offers us a picture of 'a multidimensional universe' as follows:

$$\begin{array}{ccccccccccc}
. & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & .
\end{array} \qquad (24)$$

We remark at once that this is easily recognisable as a picture of Alexander of Aphrodisias' collection of ordered pairs of elements of the universe. Alexander quantified over the whole lot. It seems that Peirce extracted from Mitchell's paper the idea that we can have two kinds of quantifier, one which operates on columns and the other on rows. More precisely, if the array consists of the pairs

$$\begin{array}{cccc}
A:A & A:B & A:C & \dots \\
B:A & B:B & B:C & \dots \\
C:A & C:B & C:C & \dots \\
\vdots & \vdots & \vdots &
\end{array} \qquad (25)$$

(Peirce's notation for the pairs) then we can think of 'For every x there is a y' as taking a product indexed by the rows, followed by a sum indexed by the columns.

There is clearly a mathematical picture behind this: writing $\phi(A, B)$ for a formula about the pair $A : B$, we read 'For every x there is a y such that $\phi(x, y)$' as

$$\prod_x \sum_y \phi(x, y) \equiv \begin{cases} (\phi(A, A) \vee \phi(A, B) \vee \phi(A, C) \vee \cdots) \\ \quad \wedge (\phi(B, A) \vee \phi(B, B) \vee \phi(B, C) \vee \cdots) \\ \quad \wedge (\phi(C, A) \vee \phi(C, B) \vee \phi(C, C) \vee \cdots) \\ \quad \wedge \cdots \end{cases} \tag{26}$$

What Peirce has done here is to move Mitchell's quantifier subscript to the front of the formula, reassign the letters u and v as indices of rows or columns, and introduce the product and sum symbols $\prod$, $\sum$ to indicate whether the quantification on each index is 'All' or 'Some'. This allows Peirce to do something that evaded Mitchell, namely to express the relative scopes of the quantifiers by the order of the indexed sum and product symbols $\prod_x$, $\sum_y$.

Thus Peirce introduced prenex quantifiers, and used the order of quantifiers within the quantifier prefix to indicate their relative scopes, exactly in analogy with indexed products and sums in mathematics. Peirce's use of prenex quantifiers was standard in mathematical logic for decades after Peirce; for example, it is the default in *Principia Mathematica*. Early mathematical logicians generally preferred it to Frege's style of quantification, not least because Frege had no existential quantifier.

Certainly the major credit here should go to Peirce rather than to Mitchell. But we ought to take seriously Peirce's repeated statements that he got the ideas by reading Mitchell's paper.

7 The Expressive Power of Ibn Sīnā's Predicative Syllogistic

We return to Ibn Sīnā. The next three sections will examine ways in which Ibn Sīnā's introduction of two-quantifier sentences fell short of its full potential. The comparison between his syllogistic and the contributions of Mitchell and Peirce will emerge gradually.

What valid inferences can be validated using Ibn Sīnā's syllogistic that couldn't already be validated using Aristotle's assertoric syllogistic? This question depends on what kinds of thing can be expressed by a single two-dimensional sentence. But even if sentences taken separately can express new things, the way they interact in the syllogistic as a whole has the effect that not all the new expressive power shows up as new valid inference forms.

Fact 7.1 *Every valid Avicennan syllogism using only d and t sentences can already be validated within Aristotle's assertoric syllogistic, taking suitable terms that incorporate the time variables. If the syllogism has an optimally valid form, it can be validated by a single two-premise assertoric mood; if not optimal, it can still be validated by a compound three-premise assertoric mood (where the extra premise expresses for example that an a–d sentence entails the corresponding a–t sentence).*

This fact is fairly devastating. In Fig. 1, the Avicennan moods of types (i) and (v) involve only d and t sentences, so they yield no new validations. If we look at a list of the sections of *Qiyās* in which Ibn Sīnā discusses various kinds of syllogism:

ii.4	assertoric and d	
iii.1	d, ℓ and m	
iv.1	t	
iv.2	ℓ, m and t	
iv.3	d and t	(27)
iv.4	t	
iv.5	d and t	
iv.6	all	

we see that out of eight listed sections, at least five have no chance of contributing any new valid inference patterns not already covered by Aristotle's assertoric syllogistic.

It gets worse. Fact 7.1 covered only what can be validated using paraphrases like those exploited by Ockham as in Sect. 2. If we add what can be covered by Alexander's trick, we have:

Fact 7.2 *Every valid Avicennan syllogism using only a–ℓ, e–ℓ, i–m or o–m sentences can already be validated within Aristotle's assertoric syllogistic by Alexander's trick. As in Fact 7.1, if the syllogism is optimal it can be validated by a single two-premise assertoric syllogism.*

In Fig. 1, it follows that moods of type (iii) give no new validations. So within the optimal Avicennan forms of *Barbara*, the whole task of extending Aristotle's valid inference patterns rests on the Theophrastan forms (ii) and the Kermanshah forms (iv). Certainly the forms (ii) are new; Ibn Sīnā knew it and boasted of it. But Ibn Sīnā's favourite examples of (ii) involve only universal quantification. About the Kermanshah forms (iv), which are subtler and involve a closer interaction between the two quantifiers, Ibn Sīnā is strangely silent. The later Arabic logicians, the better of whom certainly understood the Kermanshah forms, lamented that Ibn Sīnā was lax about distinguishing ℓ from m.

These facts are massively disappointing. But one should put the blame in the right place; the fault is in Ibn Sīnā's use of his system, not in the system itself. In the case of the Kermanshah forms, they were there in Ibn Sīnā's syllogistic but he failed to exploit them. There were other things that he could have done, even within his own framework of ideas, that would have given other essentially new valid inference forms. Before I illustrate this, let me point out two significant constraints on the expressive power of Ibn Sīnā's sentences taken one at a time.

First, Ibn Sīnā's logic is *two-sorted*. By itself this is not a limitation on the expressive power. Even the restriction of the second sort to 'times' may not be a real restriction, since Ibn Sīnā often shows a tendency to interpret 'time' rather generously, for example to include situations. But Ibn Sīnā uses the sorts in a highly restricted way. For example, he never has two-dimensional sentences in which two or more quantifiers on the same sort have to be nested. The effect is that, just as Aristotle's sentences could only talk about one element at a time, Ibn Sīnā can only talk about two elements at the same time, and the two elements must be of different sorts. So although he has binary relations, there is no chance for any of them to be transitive, linearly ordered and so forth.

Second, Ibn Sīnā's two-dimensional sentences lie entirely within the Andréka–Németi–Van Benthem guarded fragment of first-order logic [5]. (More precisely, every quantifier except initial quantifiers is relativised to an atomic formula containing all the variables that are free in the subformula following the quantifier.) This is a huge limitation. The compensation is that guardedness itself allows the use of some logical techniques that might be helpful for extracting the properties of Ibn Sīnā's system.

There are two formal extensions that Ibn Sīnā could have made to his logic that would have significantly improved its expressive power without breaking either of these two limitations. One is to apply his Avicennan adjunctions with suitable adjustments to the subject term as well as to the predicate term. The other is to adopt his own suggestion mentioned at the end of Sect. 4, of replacing the $E(x, \tau)$ or $C(x, \tau)$ in the Avicennan adjunction by some other suitable 'affirmative' expression (which we can cash in as meaning an atomic formula).

These two formal extensions, added to the more general interpretation of the time sort mentioned above, would have allowed Ibn Sīnā to incorporate within his system an argument discussed by Leibniz:

> [All] writing is an art.
> Therefore everybody who learns writing learns an art. (28)
> (Mughnai [24, p. 153], my translation)

The first sentence is an assertoric a-sentence. The second can be put in the following form:

$$\forall x \big(\exists \tau \big(learns(x, \tau) \wedge writing(\tau)\big) \rightarrow \exists \tau \big(learns(x, \tau) \wedge art(\tau)\big)\big) \tag{29}$$

with the object sort ranging over people and the time sort ranging over some class containing all arts. The formal extensions just described would allow this form; the *learns* relation assures guardedness, and there is no nesting of time quantifiers.

In short, Ibn Sīnā's predicate syllogistic is, as it stands, rather unimpressive in its expressive power. But this seems to be mainly the result of a failure to follow through with the original ideas. The system itself is much more impressive than the clothes Ibn Sīnā puts on it. Mohammad Maarefi in Tehran is investigating ways in which Ibn Sīnā could easily have let his system slip over into undecidability.

Ibn Sīnā was a busy man, working in less than ideal circumstances and apparently with no students or colleagues willing to keep him up to the mark in logic. Buried in his collected correspondence we find a note, probably made by Ibn Sīnā to himself in around 1030, about a set of questions sent to him by one of his most loyal followers, Bahmanyār bin al-Marzubān:

> As for the requests after that, I have not opened my eye to any part of them for consideration; his disdain for logic, aided by that Šayẖ in that, grieved me far too much. (Trans. Reisman [28, p. 220]) (30)

It's hard to overestimate the value of researching in a group, where ideas are shared and challenged competitively. Ibn Sīnā lacked this resource.

8 The Lack of a Proof Theory

After stating the inference (28), Leibniz gives a proof of it. The following translation picks out the main logical points; there are some further subtleties that are best studied in the original Latin.

If not, then someone who learns writing doesn't learn an art.
If possible, let Titius be learning writing but not learning an art.
Then some writing is learned by Titius.
But all writing is an art. (31)
Therefore some art is learned by Titius.
Therefore Titius is learning an art.
(Mugnai [24, p. 153], my translation)

This proof is extraordinarily interesting as a transition between Aristotelian and modern methods. The general drift of the argument is what a mathematician today would call a proof from first principles: the inference is proved without using any background assumptions or lemmas, simply from the structure and meanings of the sentences involved. On the other hand, at least one move is phrased as it is in order to fit within an Aristotelian paradigm. This is the step from 'Titius is learning writing' to 'Some writing is learned by Titius'. It's a typical example of local formalising, where in mid argument we change the choice of terms. The aim in this particular case is to make 'writing' into a term for an application of the syllogistic mood *Darii* in lines 3–5 of (31).

As far as I know, there are no examples of non-trivial proofs from first principles anywhere in Ibn Sīnā's logical writings. He simply lacks the notion. So what does he do instead when he wants to show that a mood is valid?

Mainly he follows the paradigm set by Aristotle in the *Prior Analytics*. Some moods are declared 'perfect', which means that no detailed proof of their validity is needed; they are self-evidently valid. For other valid moods, Aristotle has several techniques of reduction to moods already known to be valid. For Ibn Sīnā the purpose of these techniques is to allow us to see the validity of the given mood intuitively, by a short series of steps each of which is self-evident. One technique is conversion, for example replacing a sentence 'Some C is a B' by the self-evidently equivalent 'Some B is a C'. The second technique, slightly less self-evident, is contraposition: we keep one premise fixed, and we deduce the negation of the other premise from the fixed premise and the negation of the conclusion. The third is ecthesis, where we introduce an auxiliary compound syllogism S with three premises, in such a way that the validity of S gives that of the mood under discussion, and S can be shown valid by using syllogisms already proved valid.

These techniques are not adequate for Avicennan syllogisms. The main problem arises already at the first step: no Avicennan syllogism is convincingly described as self-evident. The reason is the double level: in each sentence we need to think about both the Aristotelian component and the Avicennan one. Ibn Sīnā responds to this problem in several ways. The commonest is a kind of 'talking through', familiarising the reader with what does and what doesn't follow from each of the premises. This kind of discussion might have generated convincing proofs from first principles, but that never seems to happen in Ibn Sīnā's writings. All that Ibn Sīnā achieves is to help the reader to develop an intuitive sense of what each premise says. A second response that he uses is to suggest slogans, like 'Possibly possible is possible', or 'The conclusion takes its form from the major premise'. These slogans give some plausibility to the inference, but hardly provide logical proofs.

To my eye—and this is admittedly speculation—the most damaging effect of Ibn Sīnā's choice of an unsuitable proof theory was that he never developed proof techniques that reach below the level of a syllogism. If we look at Leibniz's argument above, we see he has a method for handling existentially quantified sentences: he drops the quantifier and introduces the name of an arbitrary individual, in the case above 'Titius'. Deducing a contradiction about Titius counts as deducing a contradiction from the original premises. Today we understand this move in terms of existential instantiation and existential generalisation; but forget the names, the key point is that these moves are moves specifically for an existential quantifier, not for a whole syllogism. Peirce achieved something equivalent when he separated the two parts of Mitchell's quantifier notation into expressions for two separate quantifiers, so that the quantifiers could be manipulated one at a time.

Ibn Sīnā did realise that the Aristotelian conditions for validity of a syllogism are separate from the Avicennan ones. The Avicennan ones are covered by his advice to 'take care of the conditions' (one of his most characteristic phrases). But this separation of conditions was still a separation of conditions *on a whole syllogism*. The more complex the syllogism, the further it was from first principles.

It's interesting to contrast Ibn Sīnā with some leading Latin Scholastics who also studied sentences with mixed quantifiers. It seems that they also missed the idea of separate inference rules involving separate quantifiers, but for the opposite reason to Ibn Sīnā. They did develop techniques for handling separate quantifiers, but the techniques were not proof rules. They were techniques of analysis that were used primarily for identifying and rejecting invalid inferences.

In more detail: From at least the mid 13th century, Scholastic logicians used expansions of initial universal quantifiers along the lines

Every man runs ≡
Man *a* runs and man *b* runs and man *c* runs and . . . (32)

They gave similar expansions of initial existential quantifiers, using 'or' instead of 'and'. This device appears, for example, in William of Sherwood [30, pp. 35–38] in the mid 13th century. By the early 14th century we find the same device being used as a test of whether a quantifier can count as initial, or in other words (to use our post-Fregean terminology) whether it has wide scope. For example, in the sentence

For every magnitude there is a smaller magnitude. (33)

we can't replace the 'there is a' by a long disjunction:

For every magnitude, *a* is smaller than it; or for every magnitude, *b* is smaller than it; or for every magnitude, *c* . . . (34)

This test shows that 'there is a' doesn't have wide scope. Logicians pointed out that this distinction allows us to formulate the rule that we can't derive from a sentence with a 'there is' quantifier not in initial position another sentence got by moving that quantifier to the front. This identifies and blocks the false inference rule that takes (33) to 'There is a magnitude smaller than every magnitude'. A number of examples along these lines appear in Burley [8, pp. 117–124] and Buridan [7, p. 874]. All this is very much a treatment of mixed quantification. But proof rules for mixed quantification are not developed, and a fortiori no rules are given for multiple quantifiers in combination with other logical items. In fact, it's my understanding that the Scholastics, unlike the Arabic logicians, didn't even have a word for 'quantifier'.

This brief account may be unfair to some Scholastic writers that I don't know about. If so I can only apologise for my ignorance.

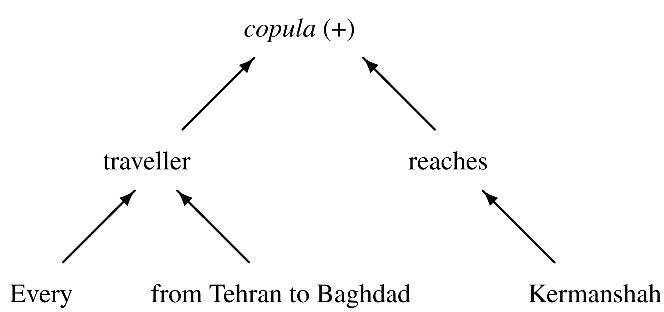
So far we have not mentioned what often seems the most disastrous part of Ibn Sīnā's proof theory. This is his use of the modal section of Aristotle's *Prior Analytics* as a template. It didn't work, mainly for the reasons mentioned above—Aristotle's proof paradigm was simply not suitable for the job. It also led some modern commentators to conclude that Ibn Sīnā had no original logic at all and was simply trying to find 'a way to understand Aristotle'. But whether this modal ingredient really militated against the development of Ibn Sīnā's logic is not so clear. We won't explore the question further here.

9 The Existential/Universal Mismatch

Why is it that Mitchell and Peirce found it easy to separate out the logical features of two quantifiers in a sentence, and Ibn Sīnā didn't? It seems that we can point to one major reason. Ibn Sīnā failed to do something sensible, not because he was ignorant but because he saw something else that distracted him. What he saw was a set of linguistic facts about existential quantification.

Ibn Sīnā's view of logic was deeply syntactic. He believed that it's impossible for us to manipulate compound ideas directly; we can only do it through language. He also believed that the structures of language reflect underlying structures of meaning, though not always faithfully—the reflection can't be exact because different languages use different structures, a point that he often mentions. And above all, for Ibn Sīnā the notion of validity of a natural language argument rests ultimately on intuitions that we have as users of the language in question. As a result of all this, Ibn Sīnā's logical writings are dotted with discussions of the structures of propositions, which allows us to picture in some detail how he thinks of these structures. (For some further information see [13], though this is an incomplete account of work in progress.)

For example, consider how Ibn Sīnā handles quantifiers. Here is a sketch of how he sees one sentence that we have already considered:



(35)

The structure shown is incomplete because it leaves out what needs to be added in order to specify that the reaching of Kermanshah happens while travelling from Tehran to Baghdad. There doesn't seem to be a right answer about where to add the required adjunction without some grasp of the workings of the time quantifier. But already from the diagram above we see a problem with the object quantifier 'Every'. By attaching it to 'traveller' Ibn Sīnā loses control over its scope. This is precisely why Frege moved the 'Every' up to the top when he introduced the notion of scope in his *Begriffsschrift*. Peirce achieved the same effect by moving the quantifier to one end of the sentence, as we saw.

Even with only a single quantifier there are still scope issues, for example about whether a negation lies within the quantifier scope. Ibn Sīnā had problems with these issues, but I leave these on one side here. [14] has further material on Ibn Sīnā's problems with the notion of scope.

We return to the issue of separating out individual quantifiers. Ibn Sīnā, like all Aristotelian logicians, sees every proposition as one of a pair, namely the proposition itself and its contradictory negation. So an early and essential part of explaining the meaning of a form of proposition is to explain how we find its negation. The contradictory negation of a universally quantified proposition is an existentially quantified proposition, and vice versa. So taking negations should set up a bijection, up to logical equivalence, between universally quantified propositions and existentially quantified ones.

Here Ibn Sīnā stubs his toe on some facts of language. Existential quantification is often expressed by indefinite pronouns. Most languages that one ever comes across, including Arabic, have a number of different indefinite pronouns for use in different contexts. But there is no corresponding complexity on the universally quantified side. So in the facts of language we don't find the required bijection. As an introductory example in English, consider

(i) A fox got into the chicken coop last night.
(ii) Some foxes got into the chicken coop last night. (36)

Then consider

No foxes got into the chicken coop last night. (37)

Isn't (37) the contradictory negation of both (i) and (ii)?

Ibn Sīnā gives several examples to illustrate different kinds of existential quantification. He is not as explicit as one would like about the distinctions he sees between them. Manuela Giolfo and I have a project which will compare Ibn Sīnā's distinctions in this area with distinctions made by the linguists of his time; in the light of soundings we expect to find a good deal of overlap. Until that work is complete, the comments below are provisional.

Ibn Sīnā's examples illustrate two dimensions of variation. The first dimension is what Martin Haspelmath in his study of indefinite pronouns [11, pp. 45–48] describes as follows:

the speaker may or may not be able to identify the referent of the indefinite pronoun. (38)

Haspelmath gives examples in Russian, Kannada, Lithuanian, German and English.

The second dimension is quantity, which in English we can illustrate with (i) and (ii) in (36). One of Ibn Sīnā's classifications is *muntašir*, which means 'widespread' or 'scattered around'; maybe 'Quite a lot of foxes got into the chicken coop last night' would be an example. This dimension of quantity can arise even if no indefinite pronouns are used.

For existentially quantified sentences the result is that Ibn Sīnā makes distinctions that vanish when we take contradictory negations. So inevitably the distinctions vanish in logic too; they never reappear in Ibn Sīnā's discussions of the validity of inference forms.

An extra twist comes in when Ibn Sīnā considers $\forall\exists$ sentences, i.e. sentences of the form 'For every x there is a y ...'. In these sentences Ibn Sīnā distinguishes between different things that can be expressed or intended by the 'there is a y'. His general picture is that the sentence states the existence of a certain set of ys for each x. As in the simple existential case above, the set may be known or unknown to the speaker, and it may be a singleton or a large set. Examples that Ibn Sīnā gives regularly are 'The (i.e. every) moon is eclipsed' (where we know exactly when this occurs, at least in terms of the physical cause) and 'Every human breathes' (where we know that it occurs often but we have no idea exactly when). And there is now the extra feature that the set may be fixed or it may vary with x.

If our chief concern is the second quantifier, then the breakdown in the correlation between sentences and their contradictory negations shows up when we try to take the negation of a $\forall\exists$ sentence. The problem is that the distinctions that language makes for the second quantifier when it is existential get lost when the second quantifier is universal. Here is one of Ibn Sīnā's own comments on this problem. This is from *Easterners*, whose text is in a bad state; so several details of the translation below may be wrong, but I think the general sense comes through.

> Someone might say: Let the negative proposition that is the opposite of this affirmative proposition be that 'Some C is not a B' in a time or situation which is the one determined by the sentence 'Every C is a B'; ... so that the first sentence is contrived to carry the determination [used in the second]. So [for example] the original sentence is contrived to carry a determination of a time or situation that is not known, and then the second sentence is contrived to have a truth value which is false [or true] in every situation where the affirmative proposition was true or false. There are two things wrong with the content of this remark. One is that if the affirmative sentence 'For every C there is a time when ...' [is true], then the negative sentence 'There is a C such that for all times it is false that ...' doesn't have to be false. The other is that when the affirmative proposition that yields this negative proposition is false, the negative proposition can also be false— which is impossible. (*Easterners*[16, 78.11–18]) (39)

A modern counterpart is that if you explain a sentence $\forall x \exists y \phi(x, y)$ as meaning $\forall x \phi(x, f(x))$, then you need to be aware that $\exists x \neg\phi(x, f(x))$ doesn't express the negation of the original sentence. The modern Skolem function f corresponds to Ibn Sīnā's 'determination' (*farḍ*).

In practice Ibn Sīnā abandons these linguistic distinctions between different forms of existential quantification when he turns to syllogistic inferences. We can read him as putting forward the view that logic has a formal structure which has itself no linguistic content, and that any questions of language come at the point where logic is applied to concrete inferences; and similarly for questions of metaphysics. I certainly do think this is his view, but there is a danger of reading him as being clearer than he is about the issues.

The main reason for setting out the facts discussed above is that these facts provide one explanation of why Ibn Sīnā never came near Mitchell's analysis, let alone Peirce's. Basically, Ibn Sīnā doesn't think of a sentence of the form 'For every x there is a y such that $\phi(x, y)$' as having two separate quantifiers. His instinct is to read the sentence as having a universal quantifier 'For every x' plus some kind of functional dependence of y on x, where the functional dependence could take any of several forms. From this position there is no easy route to the insights of Mitchell and Peirce.

10 Conclusions

First, the fact that Aristotle in his logic considers only single quantifiers in propositions is a conspicuous difference between his logic and modern logic. But it may not be a fundamental one. There are ways of adding further quantifiers to Aristotle's logic without breaking some of the most significant limitations of that logic.

Second, we can identify some of those significant limitations, which couldn't be dislodged merely by adding multiple quantifiers. Above all they include (a) local formalising (see Sects. 3, 8 above), and (b) a proof theory that treated syllogisms as wholes, rather than analysing them down to their separate logical components. When Ibn Sīnā added second quantifiers he made the problem (b) worse, not better. (Limitation (b) is closely related to what I called 'top-level processing' in [12].)

Identifying these limitations is not in any way to impugn Aristotle's genius. Rather the opposite: the more clearly we see the number of factors that had to be taken into account in building up a workable logic, the more extraordinary his achievement becomes.

Third, Ibn Sīnā should take his place as one of the most original thinkers of the Aristotelian logic tradition. Probably many other logicians in that tradition were better formal calculators than he was, and he could be a dreadful expositor. But he had a very distinctive insight into the nature and requirements of logic, which led him in some quite different directions from the European medievals. He needs to be read on his own terms, and—so far as the facts of life allow—in his own language.

This raises an obvious question: where to go for further information? Ibn Sīnā's fullest and most helpful explanations of his predicative syllogistic are in his *Qiyās* [17] and the surviving fragment of his slightly later work *Easterners* [16]. Unfortunately, most of this material is available only in Arabic, and for *Easterners* we lack even a reliable edition of the Arabic text. In present circumstances, by far the best thing that an English reader can do to understand what Ibn Sīnā is up to is to read Dimitri Gutas' book [10]; it touches only superficially on the logic, but as background on Ibn Sīnā himself and his relevant writings it is revelatory.

There is a recent professional translation of the logic section of Ibn Sīnā's *Najāt* by Asad Q. Ahmed [1]. But this work of Ibn Sīnā was apparently written some ten years before *Qiyās*, and in the interim it seems that Ibn Sīnā did a lot to crystallise the differences between his viewpoint and that of Aristotle (on which see Gutas [10] *passim*). It would be difficult to extract the facts about Ibn Sīnā's logic reported above just from the Arabic of *Najāt*, let alone from an English translation.

References

1. Ahmed, A.Q.: Avicenna's Deliverance: Logic. Oxford University Press, Oxford (2011). With introduction by T. Street
2. Alexander of Aphrodisias: Alexandri in Aristotelis Analyticorum Priorum Librum I Commentarium. Reimer, Berlin (1883). Ed. M. Wallies
3. Alexander of Aphrodisias: On Aristotle Prior Analytics 1.23–31. Duckworth, London (2006). Trans. I. Mueller
4. Alexander of Aphrodisias: On Aristotle Prior Analytics 1.32–46. Duckworth, London (2006). Trans. I. Mueller
5. Andréka, H., Németi, I., van Benthem, J.: Modal languages and bounded fragments of predicate logic. J. Philos. Log. **27**, 217–274 (1998)

6. Beziau, J.-Y.: The relativity and universality of logic. Synthese (2014). doi:10.1007/s11229-014-0419-0. Online March 2014
7. Buridan, J.: Summulae de Dialectica. Yale University Press, New Haven (2001). Trans. G. Klima
8. Burley, W.: On the Purity of the Art of Logic: The Shorter and the Longer Treatises. Yale University Press, New Haven (2000). Trans. P.V. Spade
9. Dipert, R.R.: The life and logical contributions of O.H. Mitchell, Peirce's gifted student. Trans. Charles S. Peirce Soc. **30**(3), 515–542 (1994)
10. Gutas, D.: Avicenna and the Aristotelian Tradition: Introduction to Reading Avicenna's Philosophical Works. Brill, Leiden (1988)
11. Haspelmath, M.: Indefinite Pronouns. Oxford University Press, Oxford (1997)
12. Hodges, W.: Traditional logic, modern logic and natural language. J. Philos. Log. **38**, 589–606 (2009)
13. Hodges, W.: Ibn Sina, Frege and the grammar of meanings. Al-Mukhatabat **5**, 29–60 (2013)
14. Hodges, W.: Notes on the history of scope. In: Kossak, R., Villaveces, A. (ed.) Logic Without Borders, Essays in Honour of Jouko Väänänen. De Gruyter (to appear)
15. Hodges, W.: Mathematical background to the logic of Ibn Sīnā. Notes at http://wilfridhodges.co.uk/arabic44.pdf
16. Ibn Sīnā: Manṭiq al-mašriqiyyīn (Easterners). Al-Maktaba al-Salafiyya, Cairo (1910)
17. Ibn Sīnā: Al-qiyās (Syllogism). Ed. S. Zayed, Cairo (1964)
18. Kalbfleisch, C.: Galeni Institutio Logica. Teubner, Leipzig (1896)
19. Kieffer, J.S.: Galen's Institutio Logica: English Translation, Introduction, and Commentary. Johns Hopkins Press, Baltimore (1964)
20. Mill, J.S.: A System of Logic Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation, 8th edn. Longmans Green, London (1872)
21. Mitchell, O.H.: On a new algebra of logic. In: Peirce, C.S., et al. (eds.) Studies in Logic, pp. 72–106. Little Brown and Company, Boston (1983)
22. Morison, B.: Logic. In: Hankinson, R.J. (ed.) The Cambridge Companion to Galen, pp. 66–115. Cambridge University Press, Cambridge (2008)
23. Movahed, Z.: De re and de dicto modality in Islamic traditional logic. Sophia Perennis **2**, 5–14 (2010)
24. Mugnai, M.: Leibniz' Theory of Relations. Franz Steiner, Stuttgart (1992)
25. Peirce, C.S.: On the algebra of logic. Am. J. Math. **3**, 15–157 (1880)
26. Peirce, C.S.: On the algebra of logic: a contribution to the philosophy of notation. Am. J. Math. **7**, 180–202 (1885)
27. Peirce, C.S.: On the algebra of logic (Second paper), Summer 1884. In: Kloesel, C.J.W., et al. (eds.) Writings of Charles S. Peirce, a Chronological Edition, Volume 5, 1884–1886, pp. 111–115. Indiana University Press, Bloomington (1993)
28. Reisman, D.C.: The Making of the Avicennan Tradition. Brill, Leiden (2002)
29. William of Ockham: Summa Logicae. University of St. Bonaventure, St. Bonaventure (1974). Ed. Philotheus Boehner et al.
30. William of Sherwood: Treatise on Syncategorematic Words. University of Minnesota Press, Minneapolis (1968). Trans. N. Kretzmann

W. Hodges (✉)
Herons Brook, Sticklepath, Okehampton, Devon, England EX20 2PY, UK
e-mail: wilfrid.hodges@btinternet.com

On the Contrary: Disjunctive Syllogism and Pragmatic Strengthening

Laurence Horn

Abstract Bosanquet's dictum that "The essence of formal negation is to invest the contrary with the character of the contradictory" (Bosanquet in Logic, vol. 1. Clarendon, Oxford, 1888) describes the tendency for contradictory (apparent wide-scope) negation to be semantically or pragmatically strengthened to contrary readings whenever possible. Strengthening to a contrary instantiates the inference schema of disjunctive syllogism or modus tollendo ponens: from $\phi \vee \varphi$ and $\neg\phi$, infer φ. The role of disjunctive syllogism is instantiated in a variety of strengthening shifts in natural language where a disjunctive excluded-middle premise is pragmatically presupposed in relevant contexts. In a range of apparently quite diverse phenomena—negative strengthening in lexical and clausal contexts (e.g. "neg-raising"), apparent scope adjustments with negated plural definites and bare plurals, epistemic strengthening of weak implicature in both main and embedded contexts, and children's word learning strategies, among others—can be collected under the umbrella the general preference for strengthening to contrariety via disjunctive syllogism. This can be modelled using the Square of Opposition Aristotle describes in Chap. 46 of the *Prior Analytics I*, which I dub the SINGULAR SQUARE, to formalise his analysis of the interrelations among singular expressions (*it's good/it isn't good/it's not-good/it isn't not-good*).

Keywords Contrariety · Disjunctive syllogism · Excluded middle · Pragmatic strengthening · Singular square

Mathematics Subject Classification Primary 03A99 · Secondary 03B65 · 03B99

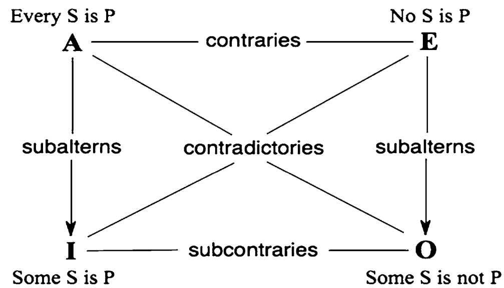
> 'I know what you're thinking about,' said Tweedledum; 'but it isn't so, nohow.'
> 'Contrariwise', continued Tweedledee, 'if it was so, it might be; and if it were so, it would be; but as it isn't, it ain't. That's logic.'
>
> —Lewis Carroll, *Through the Looking-Glass*

1 Introduction: The Asymmetric Square

It is an honour to participate in this Festschrift for Jean-Yves Béziau, a prime mover in the realms of "universal logic" and the application of Aristotelian logic (and the geometric representations thereof). This contribution is an exercise in applied neo-Aristotelian logic, exploring the genesis of the asymmetry of the Square of Opposition as reflected in the operation of contradictory and contrary opposition in natural language.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_10

Fig. 1 Traditional square of opposition



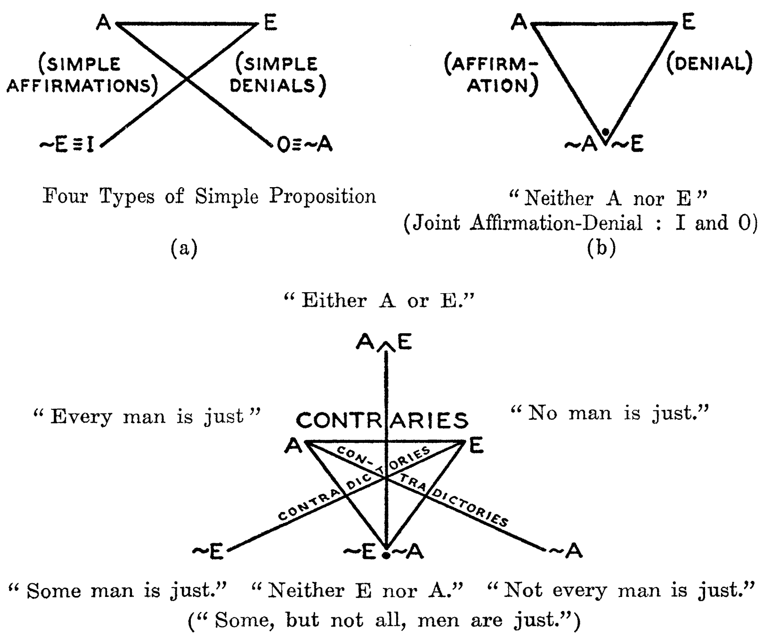


Fig. 2 Jacoby's geometries of opposition. From Paul Jacoby ([45], pp. 38 and 44) with permission from © 1950 Philosophy Documentation Center

What is generally termed the "Aristotelian Square of Opposition" in fact postdates the Stagyrite by approximately eight centuries. Both the traditional Apuleius–Boethius square (typically drawn as an oblong rectangle as in Fig. 1) and its triangular and hexagonal reshapings due to Jacoby, Sesmat, Blanché, among others (see [7, 45, 74]), as sampled in Figs. 2 and 3, are symmetrical around the vertical axis and in the traditional case around the horizontal axis as well (beyond the asymmetry of the subaltern relations).

(For related symmetric geometries, see [36] and papers in [6].) But nature abounds in both horizontal and vertical asymmetry, culminating in the three-sided square.

The two central species of opposition (*Categories* 11b17ff., *Metaphysics* 1022b23ff.) are contradiction and contrariety; the subcontrary "opposition" is defined and named not by Aristotle but (by virtue of its placement below the contraries) by those cartographers mapping the square. In fact, for Aristotle these "contradictories of contraries" are "only verbally opposed." Here are the semantic differentiae of the three relations:

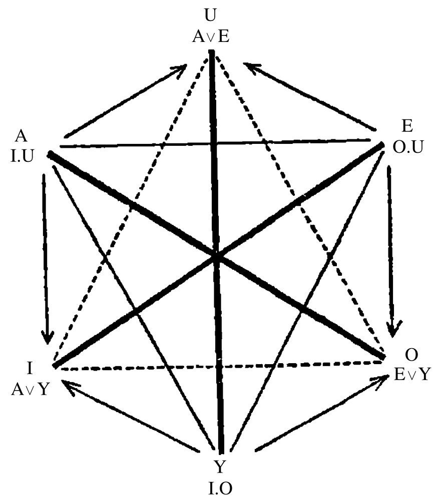


Fig. 3 Blanché's hexagon of opposition. From Robert Blanché ([7], p. 56) with permission from © 1969 Librairie Philosophique J. Vrin

(1) **CONTRADICTORIES** (**A/O**, **I/E**) "split the true and false between them"
(*All dogs are fish*/*Not all dogs are fish; Some men are bald*/*No men are bald*)
CONTRARIES (**A/E**) can be simultaneously false but not simultaneously true
(*All men are bald*/*No men are bald*)
SUBCONTRARIES (**I/O**) can be simultaneously true but not simultaneously false
(*Some men are bald*/*Some men aren't bald*)

The relation of subalternation holding between a universal and its corresponding particular is motivated by Aristotle's observation that whatever holds of all also holds of some, and if a property belongs to none it does not belong to all (*Topics* 109a3).[1]

Contradictories and contraries are distinguished by their interaction with Aristotle's two indemonstrable axioms: the Law of (Non-)Contradiction, which governs both varieties of opposition, and the Law of Excluded Middle, which applies to contradictory but not contrary oppositions (*Prior Analytics* 51b20–22, 32–33; also in *Categories*, *Metaphysics*, et al.); see [39]. Here are the two laws in their propositional form, where ¬ and © represent contradictory and contrary negation, respectively.

(2) **Law of Non-Contradiction** (**LNC**): For any proposition p, not (p and not-p)

$$\vdash \neg(p \wedge \neg p);\ \vdash \neg(p \wedge \copyright p)$$

Law of Excluded Middle (**LEM**): For any proposition p, (p or not-p)

$$\vdash (p \vee \neg p);\ \nvdash (p \vee \copyright p)$$

For natural language, the guiding principle is:

[1] What of existential import? Which of the four statement forms entail or presuppose that the set over which the quantifier ranges is non-null and how does this affect the subaltern and other relations? In particular, if *all* is import-free while *some* is not, doesn't this vitiate the Square? The fact that other operators (binary connectives, adverbs, modals, deontics) for which existential import is irrelevant can be mapped onto the Square makes such a step as unappealing as it is unnecessary. This leaves a number of options for dealing with questions of import and quantification and their relation to the Square; see Horn [38] for discussion.

(3) **MaxContrary**:

Contrariety tends to be maximised in natural language. **Subcontrariety tends to be minimised in natural language.**

1.1 On the Contrariety "Operator"

I use ©p to represent a contrary of p, so that by definition p and ©p cannot both be true but can both be false. (Others have used κ or R for a one-place non-truth-functional contrariety operator corresponding to my ©.) It should be noted that © is not a proposition-forming operator in the way ¬ is, since a given proposition may have logically distinct contraries, while this is not the case for contradictories, a point Geach [22, pp. 71–73] demonstrates with the example in (4).

(4) (a) Every cat detests every dog.
 (b) No cat detests every dog.
 (c) There is no dog that every cat detests.

While (4a) may have two syntactically distinct contradictories, *Not every cat detests every dog* and *It's not every dog that every cat detests*, co-contradictories of a given proposition will always have the same truth conditions. But (4a) allows two contraries with distinct truth conditions, (4b) and (4c). Similarly, (5a) allows the three non-identical contraries in (5b–d).

(5) (a) I believe that you're telling the truth.
 (b) I believe that you're not telling the truth.
 (c) I don't believe that you're telling the truth or that you're not telling the truth; I haven't made up my mind yet.
 (d) I don't believe that you're telling the truth or that you're not telling the truth; I haven't given the matter any thought.

Thus while we can speak of the contradictory of a proposition, Geach observes, we cannot (pace [62]) speak of **the** contrary, but only of **a** contrary, of a proposition.

But as Humberstone [43, fn. 6] points out in response to Geach's critique of McCall, the lack of uniqueness "does not prevent one from exploring the logical properties of an arbitrarily selected contrary for a given statement." For our purposes, the crucial logical property of contrariety is that it unilaterally entails contradiction:

(6) (a) $©p \vdash \neg p$
 (b) $\neg p \nvdash ©p$

In this respect, contrariety is a quasi-modal notion approximating logical impossibility, given the parallel between (6) and the fact that $\Box\neg p$ entails $\neg p$ but not vice versa. (See [10, 43, 62] for additional considerations.)

1.2 The Three-Cornered Square and the Missing O

We have noted that natural language makes a case for geometrical asymmetry. The reason is that in a wide variety of languages the values that map onto the southeast corner of the Square are systematically restricted in their potential for lexicalization. This crucial asymmetry was perhaps first recognised by St. Thomas Aquinas, who observed that whereas in the case of the universal negative (**A**) "the word 'no' [*nullus*] has been devised [!] to signify that the predicate is removed from the universal subject according to the whole of what is contained under it", when it comes to the particular negative (**O**), we find that

> there is no designated word, but 'not all' [*non omnis*] can be used. Just as 'no' removes universally, for it signifies the same thing as if we were to say 'not any' [i.e. 'not some'], so also 'not all' removes particularly inasmuch as it excludes universal affirmation.
>
> (Aquinas, *in Arist. de Int.* [64, pp. 82–83])

Thus alongside the quantificational determiners corresponding to *all*, *some*, *no*, we find no **O** determiner **nall*; we have the quantificational adverbs *always, sometimes, never*, but never a **nalways* (= 'not always', 'sometimes not'). We may have univerbations for *both* (*of them*), *one* (*of them*), and *neither* (*of them*), but never for **noth* (*of them*) (= 'not both', 'at least one... not'); alongside the connectives *and*, *or*, and sometimes *nor* (= 'and not'), we never see **nand* (= 'or not', 'not... and'). Schematically:

(7) The lexicalization asymmetry of the Square of Opposition

	DETERMINERS/ QUANTIFIERS	QUANT. ADVERBS	BINARY QUANTIFIERS	CORRELATIVE CONJUNCTIONS	BINARY CONNECTIVES
(A)	all α, everyone	always	both (of them)	both... and	and
(I)	some α, someone	sometimes	one (of them)	either... or	or
(E)	no α, no one (= all¬/¬some)	never (= always¬)	neither (of them) (=both¬/¬either)	neither... nor (=[both... and] ¬)	nor (=and¬)
(O)	*nall α, *neveryone (= some¬/¬all)	*nalways (= ¬always)	*noth (of them) (= either¬/¬both)	*noth... nand (=[either... or]¬)	*nand (=and¬/¬or)

As argued elsewhere ([32, Chap. 4], [35, §4.5], [40]; cf. also [46, 50, 56]), the motivation for this asymmetry is pragmatic: the relation of mutual implicature between the positive and negative subcontraries results in the superfluity of one of the subcontraries for lexical realisation and—given the functional markedness of negation (see [35] for a comprehensive review)—the superfluous, unlexicalised subcontrary will always be **O**. Thus instead of the symmetrical arrangement of **A**, **I**, **E**, and **O** in Fig. 1, we obtain a truncated square with no lower right corner.

The "Histoire d'*O" phenomenon extends to the modals and deontics and is also reflected in the general tendency toward **O** > **E** drift, as manifested by the northward movement of **O**-corner lexical items or collocations toward **E**. Further illustrations of the asymmetry are provided by epistemic and deontic modals, causatives, and prohibitives (see [80, 81]), supporting the general finding that the existence of a lexicalised **O** form implies the existence of a lexicalised **E** counterpart, but not vice versa, and that a lexicalised **E** form tends to be more opaque and to be semantically and distributionally less constrained than the corresponding lexicalised **O** form (if any); see [40, 41] for examples and discussion. But our focus will be on a different set of instantiations of MaxContrary and its interaction with disjunctive syllogism.

2 Polarising Disjunctions and Excluded Middles

As noted, contradictory terms (*black/non-black*, *odd/even*) exclude any middle term, an entity satisfying the range of the two opposed terms but falling under neither of them, a hat that is neither black nor not-black, an integer which is neither odd nor even. Contraries, by contrast, allow a middle: my shirt may be neither black nor white, my sister neither happy nor sad. Yet in certain remarkable circumstances the gap between the contraries narrows and even disappears, the middle effectively excluded or swallowed up.

The centrifugal politics and/or theology of polarisation tends to force every entity within the range of the polar contraries to choose one of the two terms to fall under. In this setting, evoked by revolutionists of every stripe, everything is black or white, there are no shades of gray. Hitler, for one, was described by a contemporary as operating under "a two-valued classification scheme where everything was Black (dark, evil, Jewish) or White (pure, good, Aryan)" [69, 194]. But others, from Jesus on, have endorsed their own binary taxonomies:

(8) (a) He who is not with me is against me. (Jesus, in *Matthew* 12:30 and *Luke* 11:23)
  (b) Keiner oder alle. Alles oder nichts. (Bertolt Brecht)
  (c) O con noi o contro di noi. (Benito Mussolini)
  (d) Either you're part of the solution or you're part of the problem. (Eldridge Cleaver)
  (e) Either you are with us or you are with the terrorists. (George W. Bush)

In his seminal investigation of gradable terms, Sapir [72, p. 133] points to the existence of a psychological excluded middle: "Three-term sets [*superior/average/inferior*, *good/moderate/bad*, *big/medium/small*, *warm/lukewarm/cool*] do not easily maintain themselves because psychology, with its tendency to simple contrast, contradicts exact knowledge, with its insistence on the norm, the 'neither nor'." Given this preference for simple, either-or contrast, the middle term, occupying a ZONE OF INDIFFERENCE, tends to be "quasi-scientific rather than popular in character."

This yields a productive, if context-dependent, process. Polar contraries p and q are treated as mutually exhaustive as well as mutually inconsistent; when we can eliminate all values but p and q, we obtain the disjunction in (9a), which—despite its formal contrariety—functions as an instance of (9b), i.e. the Law of Excluded Middle (LEM).

(9) (a) $p \vee q$
  (b) $p \vee \neg p$

LEM applies where it "shouldn't", based on the possibility of establishing such pragmatic disjunctions between semantic contraries in a given context. As the neo-Hegelian Sigwart observed, the efficacy of LEM derives from the establishment of such pragmatic disjunctions between semantic contraries:

> We are able, on the ground of our knowledge and of the particular contents of our subjects and predicates, to frame two positive statements, of which we know [as with] contradictory judgments that while both cannot be true together, neither can both be false; and in this case we gain, by denial of either member of the disjunction, a definite, unambiguous affirmation. [76, p. 155]

This pattern of inference has been long been recognised as the DISJUNCTIVE SYLLOGISM (DjS), the law of MODUS TOLLENDO PONENS, or—if you're a Stoic (cf. [60])—the "fifth indemonstrable syllogism":

(10) $p \vee q$

$\underline{\neg p}$

$\therefore q$

As Sigwart recognises, the crucial step is the first one, the initial establishment in context of the disjunction: "If we could solve all difficult questions by starting right off with 'it is either so or so'—'he is either mentally healthy or diseased in mind', 'the number is either odd or even'—then indeed the principle of excluded middle would be an invincible weapon. No disjunction of contraries, no LEM; no LEM, no DjS.

But when does this strategy of "divide and assert" extend to semantic contraries? We have touched on those cases that arise from the polarising tendency of revolutionary credos. When the centre does not hold, there are only two possibilities: if everything is either good or evil, and something isn't good, what else can it be? If *evil* expands to cover the territory of 'not good', *not good* is essentially reduced to 'evil'. Thus a formal contradictory (*not good*, vis-à-vis *good*) is strengthened in terms of the relevant scale to yield the assertion of a contrary (*bad*, *evil*). In such cases, warns Sigwart [76, p. 195], "The opposition of predicates [e.g. *good* vs. *evil*, *white* vs. *black*] has substituted itself unnoticed for the mere negation, and the negative statement [*x is not good*, *y is not white*] seems to tell us more than it really does." Note, however, that while *x is not good* is, in Sigwart's words, "understood as if it applied to the truth of the proposition with the opposite predicate", i.e. *x is evil*, the parallel strengthening is less plausible when the negative statement itself involves implicit negation: *x is not evil* $\neq$ *x is good*.

For Sigwart's contemporary and rival, the fellow-neo-Hegelian philosopher Bosanquet [11, p. 306], this strengthening of "mere contradictories" to contraries is fundamental: "The essence of formal negation is to invest the contrary with the character of the contradictory" In support of his claim that "negation always involves contradiction between contraries" rather than simple contradictory opposition, on the one hand, or pure contrariety, on the other, Bosanquet cites the apparent "mere contradiction" between *He is good* and *He is not good*, where the latter is semantically a relatively weak "non-informative form" which in practice is "filled in", "so that from 'he is not good' we may be able to infer something more than that 'it is not true that he is good' " (p. 310). As a related illustration of the same tendency to fill in a literal contradictory, Bosanquet (p. 337) recognises "...the habitual use of phrases such as [*I do not believe it*], which refer grammatically to a fact of my intellectual state but actually serve as negations of something ascribed to reality... Compare our common phrase 'I don't think that'—which is really equivalent to 'I think that ___ not'."

Schematically, at least three scenarios for contrary negation in contradictory clothing can be recognised English, one (not observed by Bosanquet) for the conventionalised strengthening of negative affixes (cf. [47, p. 144], [35, §5.1], [37]) and the two noted just above, the simple litotes or understatement of *it's not good* and the "neg-raising" effect of *I don't think that p*:

(11) (a) *contrary readings for affixal negation* (conventionalised strengthening)

He is unfriendly	(stronger than, i.e. unilaterally entails ¬[He is friendly])
She was unhappy	(stronger than, i.e. unilaterally entails ¬[She was happy])
I disliked the play	(stronger than, i.e. unilaterally entails ¬[I liked the play])

(b) *litotes/understatement in simple denials* (online pragmatic strengthening)
He's not happy with it (pragmatically stronger than ¬[He's happy with it])
She doesn't like fondue (pragmatically stronger than ¬[She likes fondue])
I'm not optimistic that p (pragmatically stronger than ¬[I'm optimistic that p])

(c) *"neg-raising" effects* (short-circuited implicature; cf. [42])
I don't believe it'll snow (≈ I believe it won't snow)
I don't want you to go (≈ I want you not to go)
It's not likely they'll win (≈ It's likely they won't)

In each case, the negation of an unmarked positive scalar value implicates a stronger (contrary) negation, based on a pragmatically motivated assumed disjunction: In a context licensing the pragmatic assumption p ∨ q, to assert ¬p is to R-implicate q, where R-based implicature (on the dualistic Q vs. R model of [34, 35]) is a strengthening rather than upper-bounding implicature that may undergo subsequent conventionalisation.

3 Mapping Negation on the Singular Square

Well before Apuleius and Boethius designed the traditional "Aristotelian" Square of Opposition relating the **quantified** statement types in (1), the Stagyrite himself—in Chap. 46 of *Prior Analytics I*—provided architectural instructions for building a square for **singular** predicate-subject statement types. This SINGULAR SQUARE, as we shall dub it, distinguishes positions for two varieties of negation, predicate denial (with the semantics of wide-scope sentential negation) and predicate term negation (essentially narrow scope constituent negation), corresponding to contradictory (**O**-vertex) and contrary (**E**-vertex) opposition respectively (see [35] for elaboration). Crucially, the two operators are not truth-conditionally interchangeable but can be characterised by unilateral entailment:

> The expressions 'it is a not-white log' and 'it is not a white log' do not imply one another's truth. For if it is a not-white log, it is a log: but that which is not a white log need not be a log at all. (*Prior Analytics I*, 51b28–30)

If we rotate the square Aristotle defines (51b36ff.) for representing the relations in question and replace his Greek-letter vertex labels $A/B/\Gamma/\Delta$ with the more familiar Latin mnemonics for the vertices, we obtain the diagram in Fig. 4.

The distinction between the two negations extends to predicate adjective statements:

> Hence it is evident that 'it is not-good' is not the negation of 'it is good'. If of any statement is either an affirmation or a negation, and this ['it's not-good'] is not a negation, it must be a sort of affirmation, and must have a negation of its own, which is 'is not not-good' [*ouk estin ouk agathon*]...
> If it is true to say 'it is not-white' [*estin ou leukon*], it is true also to say 'it is not white' [*ouk esti leukon*] but not vice versa. (*Prior Analytics* 51b31–42)

Note that in Greek (as in Latin), the key distinction between predicate term negation and predicate denial is marked by word order, as seen in the English analogues of the nonfinite versions of the two constructions, e.g. *einai mê leukon* 'to be not-white' (**E**) vs. *mê einai leukon* 'not to be white' (**O**). Figure 5 gives the singular square on which *it's not good* (lit., 'good not is') is distinguished from *it's not-good* (lit., 'not good is'), and the double negation *is not not-good* is distinguished from the simple affirmative *is good*, given

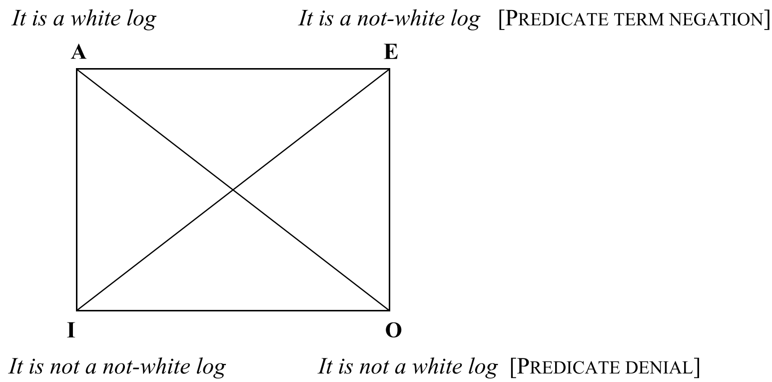


Fig. 4 Singular square: white log edition

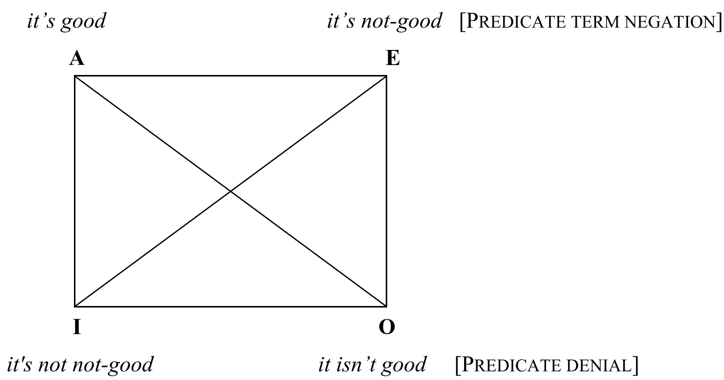


Fig. 5 Singular square for predicate adjectives

that "a thing cannot be both good and not-good", but it can be neither (*Prior Analytics* 51a36, *De Int*. 19b18–30; cf. [35, p. 16] and references therein).

3.1 Neg-Raising and MaxContrary: The Insufficiency of Pragmatic Strengthening

As St. Anselm (1033–1109) observes, *"non... omnis qui facit quod non debet peccat, si proprie consideretur"*: not everyone who does what he *non debet* ('not-should') commits a sin, if the matter is considered strictly (i.e. with the contradictory reading of negation as suggested by the surface structure). The problem, as he sees it, is the practice of using *non debere peccare* to convey the contrary *debere non peccare*, rather than the literal wide-scope contradictory (= 'it is not a duty to sin'). A man who does what is not his duty does not necessarily sin thereby, but (because of the interference of the NR reading) it is hard to stipulate, e.g. (12a)—the proposition that a man need not marry—without seeming to commit oneself to the stronger (12b), an injunction to celibacy ([28, 193ff.]; cf. [31, 86], [33, p. 200]).

(12) (a) non debet ducere uxorem lit., 'NEG [he should take a wife]'
(b) debet non ducere uxorem lit., 'he should NEG [take a wife]'

For Henry [28, p. 193], Anselm's take on the scope interaction of the negative and modal is "complicated by the quirks of Latin usage," spurring his recognition that " '*non debet*', the logical sense of which is 'It isn't that he ought', is normally used not to mean exactly what it says, but rather in the sense more correctly expressed by '*debet non*' ('he ought *not*')." Henry's assessment parallels that of logicians and epistemologists seeking to debug English of a similar flaw, from Quine's dismissal [68, pp. 145–6] of the "familiar quirk of English whereby 'x does not believe that p' is equated to 'x believes that not p' rather than to 'it is not the case that x believes that p' " as an "idiosyncratic complication" to Hintikka's complaint [29, p. 15] that "the phrase 'a does not believe that p' [$\sim B_a$p in his notation] has a peculiarity…in that it is often used as if it were equivalent to 'a believes that ∼p' [$B_a \sim$ p]" and Deutscher's acknowledgement [17, p. 55] that *I do not believe that* p can be "unfortunately ambiguous" between disbelief and simple non-belief. Grammarians, too, find NR readings more to be censured than explained; for Vizitelly [84], *I don't believe I'll go* and *I don't think it will rain* are "solecisms now in almost universal use. Say, rather, 'I believe I will not go'; 'I think it will not rain.' " In fact, Quine's "quirk"—the lower-clause understanding of higher-clause negation over a semantically coherent but somewhat variable range of predicates and operators—is far more than an unfortunate foible or error in English (or Latin) usage; cf. Prince [67]. For some, it reflects a rule of grammar.

Fifty years ago, Fillmore [19, p. 220] proposed "the Transposition of NOT (EVER)" as a cyclical rule within a fragment of generative grammar, basing his analysis on the (putative) paraphrase relation between (13b) and (one reading of) (13a).

(13) (a) I don't believe that he wants me to think that he did it.
(b) I believe that he wants me to think that he didn't do it.

The golden anniversary of this rule has now been marked by a major monograph [16], which marshals old and new evidence in support of a syntactic approach to neg-raising (NR). In so doing, Collins & Postal are bucking a trend (as seen in the chronology reviewed in [33] and [35, §5.2]). Over the years, empirical and theoretical considerations have gradually led linguists down the trail blazed by Jackendoff [44, p. 291]: "The synonymy between *John thinks that Bill didn't go* and one reading of *John doesn't think that Bill went* is inferential in character and has nothing to do with the syntactic component—it may even have nothing to do with the semantic component." Further, it is unclear whether (13a,b) are equipollent; as recognised a half-century earlier by the grammarian Poutsma [66, p. 105], "the shifting of *not* often has the effect of toning down the negativing of a sentence."

The result has been largely a return to the traditional view purveyed in grammars of English that tend to subsume NR (if it's noticed at all) under the more general pattern of formal contradictory negation (paraphrasable as 'It is not that case that p') strengthened to a contrary, as when *She's unhappy* or even *She's not happy* are understood as making a stronger negative claim than a mere denial that she is happy. These can be viewed as expressing lexicalised and virtual (or on-line) contrariety respectively; cf. (11b,c) above. Jespersen [47, p. 53], for example, presents the use of *I don't think* to "really mean" *I think he has not come* as both an instance of specialisation of negative meaning and an illustration of "the strong tendency in many languages to attract to the main verb a negative which should logically belong to the dependent nexus."

But just how and why is the hearer led to strengthen the weaker contradictory, *I don't believe that p* understood as ¬(I believe that p), to the force of the contrary *I believe that* ¬p? The classic recipe is given by Bartsch [5] in a paper that wears its conclusion on its sleeve: "'Negative transportation' gibt es nicht." Like Jackendoff, Bartsch [5, p. 3] rejects any ambiguity for (14), but proposes that while (14a) entails (14b), there arises in certain contexts a mirroring *pragmatische Implikation* from (14b) to (14a).

(14) Peter doesn't believe that Hans is coming.
  (a) Peter believes that Hans is not coming.
  (b) It is not the case that Peter believes that Hans is coming.

This implication derives from the assumption that the subject *a* can be assumed to have given some thought to the truth of the complement p and come to some conclusion about it, rather than that *a* hasn't thought about p or is neutral as whether p or ¬p. Propositional attitudes ('think', 'believe', 'want') express the subject's cognitive or psychological stance toward the complement, inducing a disjunctive pragmatic presupposition of the form "[*a* believes that p] or [*a* believes that ¬p]." Thus so-called neg-raising is not a rule of grammar or semantic interpretation but a (mere) pragmatic implication; (14a,b), while semantically distinct, can express the same information relative to a given *Sprechsituation*. Bartsch's inference schema in (15) is an instance of the disjunctive syllogism expressed in (10) above.

(15)	(i) $F(a, \text{p}) \vee F(a, \neg\text{p})$	(the pragmatically presupposed disjunction)
	(ii) "$\neg F(a, \text{p})$"	(the proposition asserted)
	(iii) $F(a, \neg\text{p})$	(the proposition conveyed)

The key step is the presupposed disjunction in (i): if you can assume I either want to go or want to stay (= not-go) and I say (out of diffidence, politeness, cowardice, etc.) that I don't want to go, you can infer that I want to stay.

This is a congenial story, and others have arrived independently at something similar. Klooster [51] undertakes his own quasi-Bartschean analysis of NR by invoking a 'BLACK AND WHITE' EFFECT for verbs like *think* and *want*, given that "in a discourse where judgements and intentions are relevant, but reserving or deferring them are not, verbs of the considered type are easily interpreted as dichotomous" [51, pp. 3–4]. This suggests a neutralisation of contradiction and contrariety for such predicates—

> In a sentence containing a matrix verb of the type in question, its contrary can thus (indirectly) be expressed simply by introducing negation. That is, where P is an NR verb, x the subject, and p the complement clause, the following seems to hold: $\neg P(x, \text{p})$ iff $P(x, \neg\text{p})$.

—but there's many a slip between what "seems to hold" and what does hold. Klooster ultimately rejects this equivalence, given (inter alia) the non-synonymy of the higher-clause and lower-clause versions of pairs like that in (15) and the difficulty of pinning down exactly what qualifies as "an NR verb."

In fact, Bartsch's proposed solution to the NR puzzle cannot handle variation within and across languages as to just which NR candidates can instantiate F in (15). When is the middle-excluding disjunction in (15(i)) actually assumed? Membership in the class of propositional attitudes is not necessary; as recognised by Anselm, the patron saint of neg-raising, strengthening is available for a range of a variety of deontic and epistemic predicates like *debere* that are not obviously propositional attitudes, e.g. *be supposed to, falloir* 'must', *be advisable, be likely*, and so on (see Horn [35, pp. 308ff.]] and references therein). Nor is membership in the class of epistemic propositional attitudes a sufficient condition, given that factives like *know* and *realise* and related strong epistemics like *be certain* and *be sure* fail to allow lower-neg or contrary readings or the apparent ambiguity in contexts like (14).

And then there is the problem of variation, both within and across languages. While German *hoffen* and Dutch *hopen* neg-raise, their English sister *hope* (usually) doesn't; Latin *sperare* 'hope' neg-raised but its French derivative *espérer* doesn't (while *souhaiter* 'wish, hope' does). Parenthetical *guess* is a neg-raising propositional attitude in Southern U.S. English but not in other U.S. or U.K. varieties. And so on; see [42] for further discussion and a proposed fix. Gajewski [21] takes the existence of lexical exceptions (non-NR attitude predicates) as indicating that excluded middle for NR must be treated as a "soft" presupposition, while Collins & Postal [16] see this muddle as further evidence for a grammatical rather than pragmatic account of NR readings.

But Bartsch's model of neg-raising as pragmatic strengthening via disjunctive syllogism, while not a good fit for the purpose for which it was designed, turns out to function as an excellent template for several other linguistic phenomena where assumed disjunctions are contextually invoked to massage contradictories into virtual contraries with functional excluded middles, e.g. bare plurals, plural definites, mass terms, conjunctions, and conditionals.

Before we move on to those cases, it would be useful to see how we might represent neg-raising on the adapt the Singular Square to represent neg-raising as pragmatic strengthening. In Kingsley Amis's 1979 novel *Jake's Thing*. Jake's eponymous thing is the problem he's been having in getting aroused by his wife Brenda, a situation for which the two of them are receiving some well-meant but ultimately ineffective therapy. (As usual, boldface is added.)

> "You're not enjoying this are you, me stroking you? Your face went all resigned when I started. Are you?"
> "I'm **not disenjoying** it."
> "Thanks a lot", said Brenda, stopping stroking. [3, p. 57]

As our hero later confides to a friend (p. 217), "I'm supposed to be working out what I feel about her. I **don't dislike** her, which is a start of a kind." Thus we have Fig. 6, in which the coerced innovative predicate appears in brackets:

Jake goes on to fill his friend in on the "libido therapy" he's been undergoing:

> [M]y "therapist" works on the principle that the way of getting to want to do something you don't want to do is to keep doing it. Which seems to me a handy route **from not...pause...wanting** to do it **to not-wanting, wanting not,** to do it. But I am paying him to know best. Brenda wants affection, physical affection... My chap is always on at me to go through the motions of it on the principle I've described. I'm a bit scared of being shifted **from not-pause-wanting** to do that **to not-wanting** to do it. [3, pp. 217–8]

Jake's unwanted shift from not-pause-wanting to not-wanting (i.e. wanting not) represents the usual **O** > **E** drift from contradictory to contrary as indicated by the arrow (see Fig. 7).

What prompts Jake's innovative verb forms (*not...pause...wanting*, *not-pause-wanting*) is his recognition that the simple negative (*not wanting*) will inevitably be understood as a neg-raised **E** contrary rather than the literal **O** contradictory he intends to express.

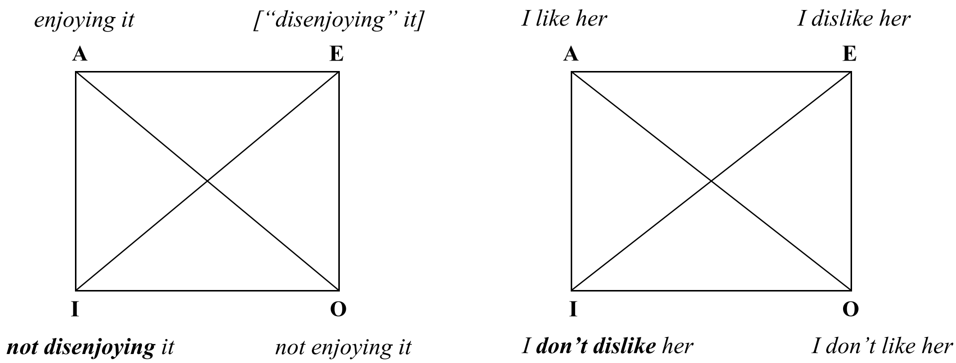
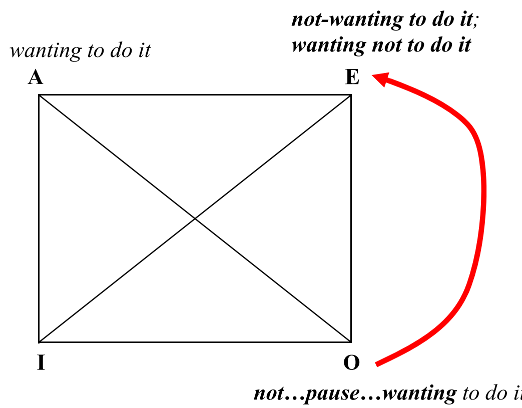


Fig. 6 Jake’s non-canceling double negations

Fig. 7 Jake on not-wanting



3.2 Beyond Bartsch: More Grist for the MaxContrary Mill

In the Bartschian picture of NR, it is the nature of the predicate that evokes the Law of Excluded Middle as a disjunction between contraries $p \vee q$, so that the formal negation of p results not merely in a simple contradictory $\neg p$ but in the affirmation of q. In other cases, LEM arises from the nature of the term phrases, in particular the subject, in the negated sentence. One instance of this is “ALL-OR-NONE” effect identified by Janet Fodor [20, pp. 158–168] in which an apparent sentential negation with a definite plural or generic bare plural scopes under the (explicit or implicit) quantification within subject and object terms. Thus if I saw the boys, this generally amounts to my having seeing them all, but the possibility for negation to outscope a universal as in (16a) is unavailable with the plural definite in (16b).

(16) (a) I didn’t see all the boys, but I did see some of them.
(b) #I didn’t see the boys, but I did see some of them.

Similarly, the bare plurals in (17b), unlike their explicitly quantified counterparts in (17a), “leave no room for disagreements about different women.”

(17) (a) All women enjoy washing dishes.
All women do not enjoy washing dishes. [on ‘not all’ reading]

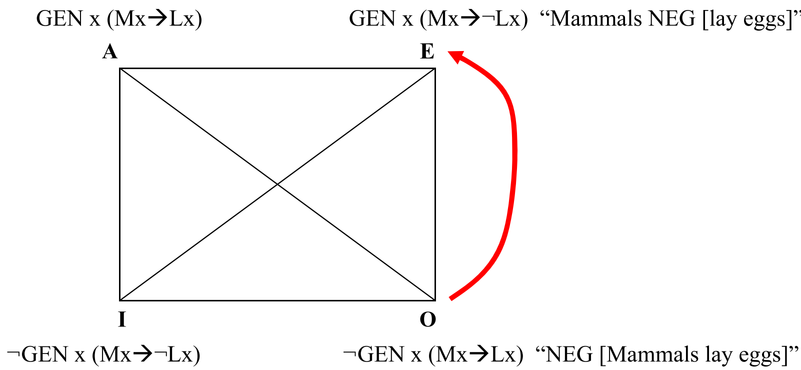


Fig. 8 Generic strengthening to the contrary

(b) Women enjoy washing dishes.
Women do not enjoy washing dishes.

In each case, the set in question, whether designated as a definite or bare plural, behaves monolithically, rendering any literal wide-scope contradictory reading of the negative versions difficult or impossible to get. Similarly, Leslie [53, p. 39] points out that in cases like (18) "we do not take negations as taking wide scope over the bare plural", with the result that both sentences come up false.

(18) (a) Fair coins come up heads.
(b) Fair coins do not come up heads.

Fodor's all-or-none for bare and definite plurals resurfaces as the GENERIC EXCLUDED MIDDLE of von Fintel [85, p. 31]—"When a kind is denied to have a generic property P_k, then any of its individuals cannot have the corresponding individual-level property P_i." As von Fintel notes, his principle is a direct descendant of Löbner's HOMOGENEITY or UNIFORMITY PRESUPPOSITION [55], which is explicated through the interpretation of negative responses to questions like *Do mammals lay eggs* or *Are the children asleep?* Löbner's homogeneity principle—"If the predicate P is false for the NP, its negation not-P is true for the NP"—is later reformulated as the PRESUPPOSITION OF INDIVISIBILITY: "Whenever a predicate is applied to one of its arguments, it is true or false of the argument as a whole" [57, p. 239]. Either mammals lay eggs or mammals don't lay eggs.

These different formulations are clearly variations on a theme. In each case, the members of a set A either homogeneously exhibit a property (e.g. egg-laying) or homogeneously exhibit the opposed property (e.g. non-egg-laying); the possibility that there might be an $a \in A$ in one camp and a $b \in A$ in the opposite camp is excluded from consideration. The Law of Excluded Middle, in the form of the all-or-none, homogeneity, or indivisibility, strengthens apparent wide-scope sentential negation ('No, it's not the case that mammals [in general] lay eggs') into a contrary of the positive ('Mammals are such they in general [don't lay eggs]'), by virtue of the nature of the implicitly quantified terms with which negation interacts, or rather fails to interact. Schematically, we have Fig. 8, where GEN is the usual implicit quantifier for bare plurals:

Negation also tends to scope under mass nouns: it's hard for me to deny your claim in (19a) by uttering (19b) without seeming to commit myself to a characterisation of (all) meat, precisely as with the bare plural generics in (17) and (18). The "middle" possibility that some meat might be healthful and some not healthful is tacitly excluded.

(19) (a) Meat is good for you.
(b) (No,) meat isn't good for you. [= ©(19a), not ¬(19a)]

Some time before Fodor, Löbner, and von Fintel, Aristotle formulated his own all-or-none (more technically, both-or-neither) for the case of "dialectical" or conjoined questions like that of A in the exchange in (20).

(20) (A) Are Coriscus and Callias at home?
(B) No.

According to Aristotle, B's negative response to A is ill-formed if just one of them is home, for "this is exactly as though he had asked 'Are Coriscus and Callias at home or not at home?', supposing them to be both in or both out" (*Sophistical Refutations* 175b40–176a17; cf. also *De Int*. 20b12). In his commentary, Ackrill [2, p. 145] points out that B's negative answer above technically just commits him to an inclusive disjunction of negations (= 'Coriscus isn't at home or Callias isn't at home'), contra Aristotle's view that A's question "normally presupposes that they are both in or both out, and that the answer 'no' inevitably accepts this presupposition". But note the difference between (21) and (22), where B is assumed to know that Austria is in the E.U. but Switzerland isn't:

(21) (A) Are both Austria and Switzerland in the E.U.?
(B) No.

(22) (A) Are Austria and Switzerland in the E.U.?
(B) #No.

While Ackrill's objection would hold for the exchange in (21), Aristotle's misgivings about (20B) as a possible answer if just Coriscus is at home generalise to the case of (22), where excluded middle and MaxContrary are operative.

In characterising Aristotle's view of (20) as invoking a presupposition, Ackrill anticipates the Fodor–Löbner view on the all-or-none for bare and definite plurals,[2] as well as the Bartsch–Gajewski treatment of NR via excluded middle. We can translate the Aristotelian inference into the DS-normal form of the schema in (20′), along the lines of (15) above:

(20′) (i) Cory and Cally are at home or Corey and Cally are not at home.
(ii) "No, ¬(Cory and Cally are at home)"
(iii) Cory and Cally are not at home [i.e. they're both out]

But is excluded middle really a presupposition at all (even a soft one), or is it rather an implicature? Note that all-or-none "presuppositions" can be cancelled or overridden:

(23) Almost all the new media of that day [17th c. France] were working, in essence, for kinglouis.gov. Even later, full-fledged totalitarian societies **didn't burn books. They burned some books**, while keeping the printing presses running off such

[2]Another natural extension of MaxContrary is to Conditional Excluded Middle (cf. [77, 85, 87] for discussion and references), based on the plausibility of the assumed disjunction: [if A then C] ∨ [if A then ¬C]). Given this disjunction, the denial of A > C amounts to the assertion of A > ¬C.

quantities that by the mid-fifties Stalin was said to have more books in print than Agatha Christie. (Adam Gopnik, New Yorker, 14 & 21 Feb. 2011, p. 125)

(24) DIANE [laughs]: No, David. You would hate it. You hate people.
DAVID: **I don't hate people. I hate…some people.**
("The Good Wife", CBS TV, 4 Mar. 2012)

The boldface rebuttals in such cases can perhaps be considered to be instances of metalinguistic or echoic negation, but this is not a requirement, as (25) (*gratia* Elena Herburger) and (26) indicate, although for me the full rebuttals in the B_2 versions strike me as a bit less natural qua contradictory sentence negations than the reduced versions.

(25) (A) The children are asleep.
(B_1) Not true, I can hear someone traipsing around upstairs.
(B_2) ?No, the children aren't asleep, I can hear someone traipsing around upstairs.

(26) (A) Mammals give live birth.
(B_1) No, you're wrong: platypuses lay eggs.
(B_2) ?No, mammals don't give live birth—platypuses lay eggs.

Krifka [52] has argued that the homogeneity effect is not a presupposition but derives from pragmatic strengthening. He observes that while the universal reading is preferred in positive plural predications like (27), the negated existential is preferred in (28).

(27) (a) The windows are made of security glass.
(b) $\forall x[x \subseteq \text{THE WINDOWS} \rightarrow \text{MADE OF SECURITY GLASS}(x)]$ (←preferred)
(c) $\exists x[x \subseteq \text{THE WINDOWS} \wedge \text{MADE OF SECURITY GLASS}(x)]$

(28) (a) The windows are not made of security glass.
(b) $\neg\exists x[x \subseteq \text{THE WINDOWS} \wedge \text{MADE OF SECURITY GLASS}(x)]$ (←preferred)
(c) $\neg\forall x[x \subseteq \text{THE WINDOWS} \rightarrow \text{MADE OF SECURITY GLASS}(x)]$

The relevant generalisation is that "In predications on sum individuals, the logically stronger interpretation is preferred" [52, p. 12]; that is to say, "If grammar allows for a stronger or weaker interpretation of a structure, choose the one that results in the stronger interpretation of the sentence, if consistent with general background assumptions" [52, (39)]. Krifka suggests that this principle might be assimilated to the R-based implicatures of [34, 35] that license a speaker to underspecify the force of her utterance while counting on the hearer to recover, in Krifka's formulation, "the strongest possible interpretation that is consistent with the background knowledge."

4 Disjunctive Syllogism and Implicature Strengthening

In saying *Some Fs are G*, a speaker is typically understood to be implicating that not all Fs are G. Under what conditions is such an inference warranted? While perhaps not immediately recognisable as an analogous case of MaxContrary or pragmatic strengthening, a closer examination reveals the operation of the same inferential schema we have outlined in the previous two sections in the strengthening of a weak or primary quantity (scalar) implicature to the strong or secondary implicature typically (but not invariably) recovered by the hearer from a speaker's assertion containing a relatively weak scalar value.

By Grice's first submaxim of quantity—"Make your contribution as informative as is required (for the current purposes of the exchange)" [26, p. 26]—or what I have formulated as the Q Principle (see [34, 54]), a speaker uttering (29a) will suggest (29b).

(29) (a) Some Spaniards speak Basque.
(b) For all the speaker knows, not all Spaniards speak Basque.
(c) (The speaker believes/knows that) Not all Spaniards speak Basque.

In the terminology of [32, 35], the move from (29a) to (29b) is a scalar implicature, since it operates off the implicitly invoked scale <*some, many, most, all*>. The use of a weaker value on the scale implicates (ceteris paribus) that the speaker was not in a position to have used a stronger scalar competitor salva veritate (see [30] for a fuller picture and [24, 49] for complications); if the speaker knew that all Spaniards speak Basque, she should have said so. But in most circumstances, the recipient of (29a) will tend to infer not just (29b) but the stronger (29c).

More generally, the hearer recovers the implicature in (30a), where W and S are weak(er) and strong(er) scalar alternatives and $\mathbf{B}_a$ is the belief operator ('a believes that') of Hintikka 1962. But a hearer who can invoke what Geurts [24, pp. 29ff.], following Sauerland [73] and van Rooij & Schulz [82], calls the COMPETENCE ASSUMPTION—"the speaker knows the relevant facts" and hence "is not undecided about the truth of the stronger alternative"—will tend to strengthen (30a) to (30b):

(30) (a) Given $<W, S>$, a speaker a uttering "$\ldots W \ldots$" Q-implicates $\neg\mathbf{B}_a(\ldots S \ldots)$.
(b) Given $<W, S>$, a speaker a uttering "$\ldots W \ldots$" Q-implicates $\mathbf{B}_a\neg(\ldots S \ldots)$.

It will be noticed that the competence assumption is essentially Bartsch's (and Sigwart's) assumed disjunction between contraries. We can thus represent the inferential strategy into the now familiar DjS-normal form:

(31) (i) $\mathbf{B}_a(\ldots S \ldots) \vee \mathbf{B}_a\neg(\ldots S \ldots)$ (competence assumption)
(ii) "$(\ldots W \ldots)$" (speaker a's assertion)
(iii) $\underline{\neg\mathbf{B}_a(\ldots S \ldots)}$ (weak Q-implicature from (ii), via Quantity Maxim)
(iv) $\mathbf{B}_a\neg(\ldots S \ldots)$ (the strong Q-implicature, via DjS from (i), (iii))

This reasoning schema is actually foreshadowed by Mill [63] in his proto-Gricean rebuttal to Hamilton's analysis [27] of *some* as 'some only, some but not all':

> No shadow of justification is shown... for adopting into logic a mere sous-entendu of common conversation in its most unprecise form. If I say to any one, "I saw some of your children today", he might be justified in inferring that I did not see them all, not because the words mean it, but because, if I had seen them all, it is most likely that I should have said so: **though even this cannot be presumed unless it is presupposed that I must have known whether the children I saw were all or not.** [63, p. 442, emphasis added]

To provide a singular square for the move from (30a) to (30b) (or from (31iii) to (31iv)), we need to add a piece of notation: p$[a, b]$ represents the proposition p with any instances of b replaced by instances of a. Thus, in Fig. 9, p$[S, W]$ takes a proposition containing a weaker scalar value W (e.g. *some, possible, or*) and replaces it with one containing the stronger competitor S (e.g. *all*, *certain/necessary*, *and*).

The same (neo-Bartschean) means for deriving strong implicatures via assumed disjunction can be invoked in a different context. In [13, 14], and subsequent work, Gennaro Chierchia and his colleagues have argued that the global implicatures predicted by the

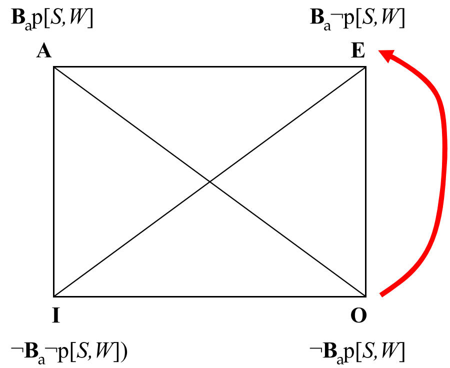


Fig. 9 Scalar implicature strengthening

Gricean model are empirically unsupported and that the facts require a localist account in which implicature calculation is part of the computational system rather than determined by Gricean assumptions about rationality and cooperation. Chierchia notes, for example, that in uttering a sentence like (32a), I seem to implicate not just (32b) but the stronger (32c). (Examples are from [23, 70].)

(32) (a) George believes that some of his advisors are crooks.
(b) [According to me] it is not the case that George believes that all of his advisors are crooks: $\neg\mathbf{B}_g$ (all of g's advisors are crooks)
(c) [According to me] George believes that not all of his advisors are crooks.

On Chierchia's localist analysis, the move from *some* in (32a) to *not all* in (32c), and hence the resultant communication of 'some but not all', is delivered not through pragmatic inference but by stipulating the inference to *not all* or *only some* as a default within the computational system.

Geurts's response (see [23, p. 68], [24, p. 169]), following Ben Russell [70], is that "We may assume that the competence assumption holds not only for the speaker but for the subject of the belief report", allowing us to derive the key disjunction:

(33) B_g (all of George's advisors are crooks) $\vee$ B_g $\neg$(all of George's advisors are crooks)

Then we obtain (36c) from (32b) + (33) via DjS, essentially as in (30). As Geurts observes, this approach has the virtue of allowing the generation of both the weak implicature in (32b) and the strong implicature in (32c) by the same general mechanism, modulo the contextual plausibility of the second-order excluded middle "competence assumption" in (33).

Not all propositional attitudes allow easy access to the relevant disjunction, and in those cases the "local implicature" effect will be attenuated. For example, according to Sharvit & Gajewski [75], *certain* licenses local implicatures, so (34a) implicates (34b).

(34) (a) John is certain that the boss or her assistant have disappeared.
(b) John is certain that the boss or her assistant but not both have disappeared.

But is this correct? Does the hearer really infer that the speaker of (34a) intends to convey (34b)? What about (35)?

(35) Our chair is certain that some of the students we admitted will accept their offers.

This doesn't generally implicate that the chair is certain that not all the admitted students will accept. (Cf. [71] for a general account of such cases within a Bayesian implementation of Gricean global implicature.)

While the jury is still out on globalist/pragmatic vs. localist/grammatical approaches to the range of phenomena associated with scalar implicature, the relation between the strengthening of weak to strong implicatures surveyed in this section and the analogous strengthening in the instances of DjS discussed in the earlier sections is something that should not be ignored.

5 Word Learning, Beastly Inferences, and Disjunctive Syllogism

> When presented a trial with "doll" and "megaphone" as the two objects and told to take the megaphone, the [20–31 month old] children were able to do so—not because they knew the label "megaphone", but because they knew the name of the other object was "doll." [83, p. 192]

Our last set of illustrations of the interaction of MaxContrary with disjunctive syllogism involves word learning (by humans and others) and its relation to what psychologists call "learning by exclusion." The key step in the Vincent-Smith et al. study excerpted above is the operation of the principle that would come to be called mutual exclusivity. The young children in the study recognise, through their nascent mastery of Theory of Mind, that the "megaphone" in question must be one of two objects presented to them, the other being a familiar object—the doll. Later work by developmental psycholinguists, in particular the Contrast principle of [15] and the Mutual Exclusivity principle of [59], has demonstrated that children tacitly assume bidirectional uniqueness between objects and words, ruling out both synonymy and homonymy. If α is known to be a doll it cannot (also) be a megaphone, and if *doll* labels a particular kind of object, it cannot also label a different object.[3] If the familiar word denotes the familiar object, the unfamiliar word must pick out the unfamiliar item. The child's correct identification of the megaphone is rendered in DS-normal form in (36) and mapped onto the Singular Square in Fig. 10.

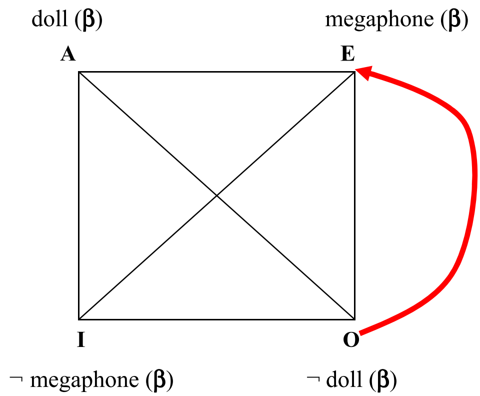
(36)	(i) $\alpha \in M \vee \beta \in M$	[inference via Theory of Mind]
	(ii) $\alpha \in D$	[Prior knowledge]
	$\underline{\text{(iii) } \alpha \notin M}$	[from (ii) by Mutual exclusivity]
	(iv) $\beta \in M$	[from (i), (iii) by Disjunctive syllogism]

Nor is it only (human) children who reason in this way. In their 2004 study, Kaminski et al. [48] describe how Rico, a border collie, was able to acquire over 200 different object labels, inferring the names of new objects by exclusion (i.e. disjunctive syllogism) and conclude that the manifestation of such inferential behaviour cannot be restricted to species-specific language acquisition. (See [9, 58] for other interpretations of Rico's achievements.)

In follow-up work, Call [12] attests successful (non-word) inferential learning by exclusion in great apes (chimpanzees, gorillas, bonobos, orangutans), as demonstrated by

[3] If a child knows the label for a given object, an unfamiliar label must pick out an unfamiliar object or a subpart (*trachea*) or property (*pewter*) of a familiar one. Bloom [8, pp. 65–87] provides an excellent survey of late 20th century research on word learning, mutual exclusivity, and the role of Gricean pragmatics and "theory of mind."

Fig. 10 Mutual exclusivity square



their ability to select the correct (non-empty) food bin after they had witnessed food being discarded from the sole alternative bin. Erdöhegyi et al. [18] further extend the results from Rico to 35 pet dogs volunteered for their project involving a forced choice task to pick out the non-hidden toy. However, the dogs in their study employed disjunctive reasoning only in absence of social-communicative cues which override perception-based information in which their humans deliberately misled the dogs by gazing at the empty container. This finding supports the familiar slogan "social-dog, causal-ape". Aust et al. [4] investigate the ability to reason by exclusion among humans (children and undergraduates), dogs, and pigeons. Such inferential reasoning is confirmed in humans and (most) dogs tested, but not every creature in their study displayed the requisite inferential leap; Aust et al. concede that "the ability to make inferences by exclusion—which requires logical reasoning independent of perceptual features—may be out of a pigeon's reach" (p. 595).

More recently, Grassmann et al. [25] take the experimental behaviour of 2-year-olds toward novel items in their study to support the social/pragmatic nature of the general inference strategies employed (based on common ground and disjunctive reasoning) of which the lexical Contrast and Mutual Exclusivity principles may be seen as special cases. Pilley & Reid [65] explore the word-learning capabilities of Chaser, a border collie who managed to out-Rico Rico by acquiring the proper names of 1022 objects learned over a three-year period through her use of inferential reasoning by exclusion to learn the names of unfamiliar objects, replicating the findings in the Kaminski et al. [48] study.[4]

Nor should these 21st century revelations of the canine application of disjunctive syllogism be surprising. As we have noted, the DjS inferential pattern is sometimes known as Modus Tollendo Ponens. But the Stoics invoked two patters in which disjunction participated, their fourth and fifth indemonstrable syllogisms [60]:

(37) 4TH INDEMONSTRABLE SYLLOGISM — 5TH INDEMONSTRABLE SYLLOGISM

4th Indemonstrable Syllogism	5th Indemonstrable Syllogism
p or q	**p or q**
p	**not-p**
not-q	**q**

[4]Rico's "fast mapping" can be seen in action at http://www.youtube.com/watch?v=kW27XF21ORs#t=12, while Chaser performs with co-authors Pilley and Reid at http://www.youtube.com/watch?v=Hi8HFdPMsiM.

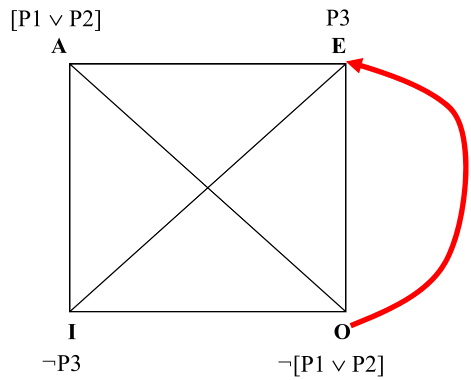


Fig. 11 Canine square

The Stoics' endorsement of the fourth syllogism reflects their view that true disjunctions are always exclusive (inclusive *or* was recognised but treated as "paradisjunction"); the exclusive reading of *or*, with the bitstring <0110> rather than <1110>, validates both of the above inferential schemata. But the fifth indemonstrable syllogism is familiar to us in the form of disjunctive syllogism (or Modus Tollendo Ponens). And familiar not only to us, but to our canine companions. Here is the 2nd c. CE philosopher Sextus Empiricus reporting on an earlier observation by Chrysippus of Soli (279–206 BCE):

> According to Chrysippus, who was certainly no friend of non-rational animals, the dog ... uses repeated applications of the fifth undemonstrated argument-schema when, arriving at a juncture of three paths, after sniffing at the two down which the quarry did not go, he rushes off on the third without stopping to sniff. For, says this ancient authority, the dog in effect reasons as follows: the animal either went this way or that way or the other; he did not go this way and he did not go that; therefore, he went the other. (Sextus Empiricus, *Pyrr. Hyp*. I, 69 in [61], 98)

The reference to "repeated applications" of the 5th syllogism (a.k.a. DjS) invokes the two-step reasoning by Chrysippus's dog:

(38) (i) p or (q or r), not-p; $\therefore$(q or r)
(ii) q or r, not-q; $\therefore$r

The successors of Sextus were fond of this story, although questions have been raised as to whether the dog is authentic or apocryphal and as to what (in either case) it says about dogs and about us. Here is St. Thomas's account:

> A hound in following a stag, on coming to a cross-road, tries by scent whether the stag has passed by the first or second road: and if he find that the stag has not passed there, being thus assured, takes to the third road without trying the scent; as though he were reasoning by exclusion, arguing that the stag must have passed by this way, since he did not pass by the others, and there is no other road. —St. Thomas Aquinas (1225–1274) [79], *Summa Theologica Pars II*, Q. 13, Art. 2

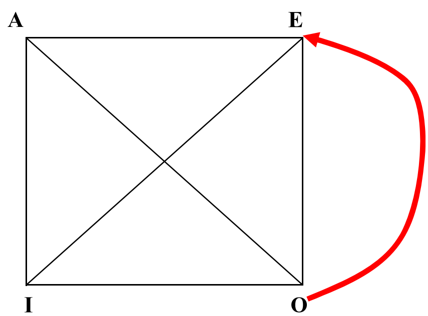
Aberdein [1] cites these and related passages from Plutarch, Montaigne, Coleridge, and others, largely endorsing canines' ability to draw DjS-based inferences.

Conversion to a Singular Square is straightforward. The dog in Chrysippus's (or Sextus's) anecdote—call him Argos—distinguishes three diverging paths down which the rabbit might have gone; let P1 designate the proposition that the rabbit went down the first path, and so on. Then we have Fig. 11.

In effect, Argos proceeds (à la Stalnaker [78]) by winnowing down his set of live rabbit-options by eliminating the possible worlds corresponding to P1 and P2, leaving him with just the P3-worlds—so why bother stopping to sniff?

6 Concluding Remarks

Expanding on Bosanquet's dictum that "The essence of formal negation is to invest the contrary with the character of the contradictory," we have reviewed multiple instantiations of the tendency by which apparent contradictory (formally wide-scope) negation is systematically strengthened (semantically or pragmatically) to contrary readings. This strengthening process begins by invoking the contextually licensed availability of a pragmatic disjunction between contraries as the first step in an application of the inference schema of disjunctive syllogism: from $\phi \vee \varphi$ and $\neg\phi$, infer φ. In this way, a range of apparently quite diverse phenomena in natural language can be collected under the umbrella the general preference for strengthening to contrariety via disjunctive syllogism and represented by the Singular Square of Aristotle's *Prior Analytics*. This general strengthening can be represented by the now familiar **E** > **O** shift along the "negative" side of the Square as in Fig. 11 and summarised by the syllogistic inference in (39).



(39) (i) "**O**" speaker's assertion
(ii) **A** ∨ **E** assumption of excluded middle (disjunction between contraries)
(iii) ¬**A** from (i) by definition of contradictory opposition
(iv) **E** from (ii), (iii) by DjS

We end this discussion by revealing the Ten Commandments of Contrariety, Disjunctive Syllogism, and Pragmatic Strengthening—or, invoking Tweedledee's remark in the epigraph, the Ten Pillars of Contrariwisdom. The first three of these directives are explicated in [40], a paper first presented at the First World Congress on the Square of Opposition, memorably conceived and staged by Jean-Yves Béziau in Montreux in 2007 (see also the related discussion, especially for lexical and double negation, in [41]). Seven years after that first Congress, we can now add seven additional commandments, the justification for which I have presented in the present contribution. We might conclude, with apologies to W.C. Fields, that any law of inference that generalises to children and dogs can't be all bad.

(40) **The Ten Commandments of Contrariety and Disjunctive Syllogism**

I. Thou shalt not lexically incorporate outer negations of intolerant values or inner negations of tolerant ones (see [35, 56] for terminology).
II. Thou shalt strengthen formal **O**-ish values into the **E** range.
III. Thou shalt interpret formal contradictories as virtual contraries.

IV. Thou shalt exclude the middle between contraries (when motivated).
V. Thou shalt raise thy embedded negations (when possible).
VI. Thy negated bare plural, definite plural, and mass predications shall be made strong.
VII. Thy weak implicatures shall be made strong (when epistemic considerations permit).
VIII. Thine "embedded implicatures" shall percolate globally as strong implicatures.
IX. Thy children and animal companions shall employ disjunctive syllogism for word learning.
X. Thou shalt chase thy rabbits and stags logically and with Stoicism.

Acknowledgements Some of this material was presented in other forms at previous occasions, including the first World Congress on the Square of Opposition in Montreux (June 2007), LNAT (Logic Now and Then) in Brussels (November 2008), and especially in the versions delivered at ESSLLI in Ljubljana (August 2011), SCLP in Santa Cruz (November 2011), CRISSP in Brussels (December 2011), and the Workshop on Logical Words at CIL 19 (Geneva 2013). I am grateful to Barbara Abbott, Donka Farkas, Bart Geurts, Susanne Grassmann, Elena Herburger, Dany Jaspers, Jacques Moeschler, Ben Russell, and Uli Sauerland for helpful discussions, with the usual disclaimers of responsibility.

References

1. Aberdein, A.: Logic for dogs. In: Hales, S. (ed.) What Philosophy Can Tell You About Your Dog, pp. 167–181. Open Court, Chicago (2008)
2. Ackrill, J.L. (ed. and trans.): Aristotle: Categories and De Interpretatione Clarendon, Oxford (1963)
3. Amis, K.: Jake's Thing. Viking, New York (1979)
4. Aust, U., Range, F., Steurer, M., Huber, L.: Inferential reasoning by exclusion in pigeons, dogs, and humans. Anim. Cogn. **11**, 587–597 (2008)
5. Bartsch, R.: 'Negative transportation' gibt es nicht. Linguist. Ber. **27**, 1–7 (1973)
6. Béziau, J.-Y., Payette, G. (eds.): The Square of Opposition: A General Framework for Cognition. Peter Lang, Bern (2012)
7. Blanché, R.: Structures Intellectuelles, 2nd edn. Vrin, Paris (1969)
8. Bloom, P.: How Children Learn the Meanings of Words. MIT Press, Cambridge (2000)
9. Bloom, P.: Can a dog learn a word? Science **304**, 1605–1606 (2004)
10. Bogen, J.: Aristotelian contraries. Topoi **10**, 53–66 (1991)
11. Bosanquet, B.: Logic, vol. 1. Clarendon, Oxford (1888)
12. Call, J.: Inferences by exclusion in the great apes: the effect of age and species. Anim. Cogn. **9**, 393–403 (2006)
13. Chierchia, G.: Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In: Belletti, A. (ed.) Structures and Beyond, pp. 39–103. Oxford U. Press, Oxford (2004)
14. Chierchia, G.: Broaden your views: implicatures of domain widening and the "locality" of language. Linguist. Inq. **37**, 535–590 (2006)
15. Clark, E.V.: The principle of contrast: a constraint on language acquisition. In: MacWhinney, B. (ed.) Mechanisms of Language Acquisition, pp. 1–33. Erlbaum, Hillsdale (1987)
16. Collins, C., Postal, P.: Classical NEG-Raising: An Essay in the Syntax of Negation. MIT Press, Cambridge (2014)
17. Deutscher, M.: A note on saying and believing. Analysis **105**, 53–57 (1965)
18. Erdöhegyi, Á., Topál, J., Virányi, Zs., Miklósi, Á.: Dog-logic: inferential reasoning in a two-way choice task and its restricted use. Anim. Behav. **74**, 725–737 (2008)
19. Fillmore, C.: The position of embedding transformations in a grammar. Word **19**, 208–231 (1963)
20. Fodor, J.D.: The linguistic description of opaque contexts. PhD dissertation, MIT (1970)
21. Gajewski, J.: Neg-raising and polarity. Linguist. Philos. **30**, 289–328 (2007)

22. Geach, P.T.: Logic Matters. University of California Press, Berkeley (1972)
23. Geurts, B.: Scalar implicature and local pragmatics. Mind Lang. **24**, 51–79 (2009)
24. Geurts, B.: Quantity Implicatures. Cambridge U. Press, Cambridge (2010)
25. Grassmann, S., Stracke, M., Tomasello, M.: Two-year-olds exclude novel objects as potential referents of novel words based on pragmatics. Cognition **112**, 488–493 (2009)
26. Grice, H.P.: Studies in the Way of Words. Harvard U. Press, Cambridge (1989)
27. Hamilton, W.: Lectures on Logic, vol. I. Blackwood, Edinburgh (1860)
28. Henry, D.P.: The Logic of St. Anselm. Clarendon, Oxford (1967)
29. Hintikka, J.: Knowledge and Belief. Cornell U. Press, Ithaca (1962)
30. Hirschberg, J.: A Theory of Scalar Implicature. Garland, New York (1991)
31. Hopkins, J.: A Companion to the Study of St. Anselm. U. of Minnesota Press, Minneapolis (1972)
32. Horn, L.R.: The semantic properties of logical operators in English. PhD dissertation, UCLA (1972)
33. Horn, L.R.: Remarks on neg-raising. In: Cole, P. (ed.) Syntax and Semantics 9: Pragmatics, pp. 129–220. Academic Press, New York (1978)
34. Horn, L.R.: Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In: Schiffrin, D. (ed.) Meaning, Form, and Use in Context (GURT '84), pp. 11–42. Georgetown University Press, Washington (1984)
35. Horn, L.R.: A Natural History of Negation. U. of Chicago Press, Chicago (1989). Reissue edition, Stanford: CSLI, 2001
36. Horn, L.R.: Hamburgers and truth: why Gricean inference is Gricean. In: Parasession on the Legacy of Grice. Proceedings of the Annual Meeting of the Berkeley Linguistics Society, vol. 16, pp. 454–471 (1990)
37. Horn, L.R.: Duplex negatio affirmat. . . : the economy of double negation. In: Papers from the Parasession on Negation, vol. 27, pp. 78–106. CLS, Chicago (1991)
38. Horn, L.R.: All John's children are as bald as the King of France: existential import and the geometry of opposition. In: Proceedings of the Thirty-Third Regional Meetings of the Chicago Linguistic Society, vol. 33, pp. 155–179 (1997)
39. Horn, L.R.: Contradiction. Revised entry. In: Zalta, E. (ed.) Stanford Encyclopedia of Philosophy (2010). http://plato.stanford.edu/entries/contradiction/
40. Horn, L.R.: Histoire d'*O: lexical pragmatics and the geometry of opposition. In: Béziau, J.-Y., Payette, G. (eds.) The Square of Opposition: A General Framework for Cognition, pp. 393–426 Peter Lang, Bern (2012)
41. Horn, L.R.: The singular square: contrariety and double negation from Aristotle to Homer. In: Blochowiak, J., et al. (eds.) (2014). http://www.unige.ch/lettres/linguistique/moeschler/Festschrift/Festschrift.php
42. Horn, L.R., Bayer, S.: Short-circuited implicature: a negative contribution. Linguist. Philos. **7**, 397–414 (2004)
43. Humberstone, L.: Extensionality in sentence position. J. Philos. Log. **15**, 27–54 (1986)
44. Jackendoff, R.: On some questionable arguments about quantifiers and negation. Language **47**, 282–297 (1971)
45. Jacoby, P.: A triangle of opposites in Aristotelian logic. New Scholast. **XXIV**, 32–56 (1950)
46. Jaspers, D.: Operators in the lexicon: on the negative logic of natural language. Universiteit Leiden dissertation. Utrecht: LOT (2005)
47. Jespersen, O.: Negation in English and Other Languages. A. F Høst, Copenhagen (1917)
48. Kaminski, J., Call, J., Fischer, J.: Word learning in a domestic dog: evidence for 'fast mapping'. Science **304**, 1682–1683 (2004)
49. Katzir, R.: Structurally defined alternatives. Linguist. Philos. **30**, 669–690 (2007)
50. Katzir, R., Singh, R.: Constraints on the lexicalization of logical operators. Linguist. Philos. **36**, 1–29 (2013)
51. Klooster, W.: Negative raising revisited. In: Koster, J., van Riemsdijk, H. (eds.) Germania et Alia: A Linguistic Festschrift for Hans den Besten (2003). Available at http://odur.let.rug.nl/~koster/DenBesten/contents.htm
52. Krifka, M.: Pragmatic strengthening in plural predications and donkey sentences. In: Proceedings of SALT, vol. 6 (1996)
53. Leslie, S.-J.: Generics: cognition and acquisition. Philos. Rev. **117**, 1–47 (2008)
54. Levinson, S.C.: Presumptive Meanings. MIT Press, Cambridge (2000)

55. Löbner, S.: Definites. J. Semant. **4**, 279–325 (1985)
56. Löbner, S.: Quantification as a major module of natural language semantics. In: Groenendijk, J., et al. (eds.) Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers, pp. 53–85. Foris, Dordrecht (1987)
57. Löbner, S.: Polarity and natural language: predication, quantification, and negation in particularizing and characterising sentences. Linguist. Philos. **23**, 213–308 (2000)
58. Markman, E., Abelev, M.: Word learning in dogs? Trends Cogn. Sci. **8**, 489–491 (2004)
59. Markman, E., Wachtel, G.: Children's use of mutual exclusivity to constrain the meanings of words. Cogn. Psychol. **20**, 121–157 (1988)
60. Mates, B.: Stoic Logic. University of California Press, Berkeley (1953)
61. Mates, B. (ed. and trans.): The Skeptic Way: Sextus Empiricus's Outlines of Pyrrhonism. Oxford U. Press, New York (1996)
62. McCall, S.: Contrariety. Notre Dame J. Form. Log. **8**, 121–138 (1967)
63. Mill, J.S.: An Examination of Sir William Hamilton's Philosophy, 2nd edn. Longman, Green, London (1865)
64. Oesterle, J. (ed. and trans.): Aristotle: On Interpretation. Commentary by St. Thomas and Cajetan. Marquette University Press, Milwaukee (1962)
65. Pilley, J., Reid, A.: Border collie comprehends object names as verbal referents. Behav. Process. **86**, 184–195 (2011)
66. Poutsma, H.: A Grammar of Late Modern English. Noordhoff, Groningen (1928)
67. Prince, E.F.: The syntax and semantics of neg-raising, with evidence from French. Language **52**, 404–426 (1976)
68. Quine, W.V.O.: Word and Object. MIT Press, Cambridge (1960)
69. Richter, F.K.: Hitler and words. Etc. **1**, 256–258 (1944)
70. Russell, B.: Against grammatical computation of scalar implicatures. J. Semant. **23**, 361–382 (2006)
71. Russell, B.: Topics in the computation of scalar implicatures. PhD dissertation, Brown University (2011)
72. Sapir, E.: Grading: a study in semantics (1944). Reprinted in P. Swiggers et al. (eds.), *The Collected Works of Edward Sapir, I*, 447–470. Berlin: de Gruyter, 2008
73. Sauerland, U.: Scalar implicatures in complex sentences. Linguist. Philos. **27**, 367–391 (2004)
74. Sesmat, A.: Logique II. Hermann, Paris (1951)
75. Sharvit, Y., Gajewski, J.: On the calculation of local implicatures. WCCFL **26**, 411–419 (2008)
76. Sigwart, Ch.: Logic, vol. I. (1885). H. Dendy, trans. New York: Macmillan, 1895
77. Stalnaker, R.: A defense of conditional excluded middle. In: Harper, W.L., et al. (eds.) Ifs, pp. 87–104. Reidel, Dordrecht (1981)
78. Stalnaker, R.: Context and Content. Oxford University Press, Oxford (1999)
79. Thomas Aquinas: The Summa Theologica, vol. I. Benzinger Brothers, New York (1947). Literally translated by Fathers of the English Dominican Province
80. van der Auwera, J.: Why languages prefer prohibitives. Wai Guo Yu, 1–25 (2006)
81. van der Auwera, J.: Prohibition: constructions and markers. In: Shu, D., et al. (eds.) Contrasting Meaning in Languages of the East and the West, pp. 443–475. Peter Lang, Bern (2010)
82. van Rooij, R., Schulz, K.: Exhaustive interpretation of complex sentences. J. Log. Lang. Inf. **13**, 491–513 (2006)
83. Vincent-Smith, L., Bricker, D., Bricker, W.: Acquisition of receptive vocabulary in the toddler-age child. Child Dev. **45**, 189–193 (1974)
84. Vizitelly, F.: A Desk-Book of Errors in English. Funk & Wagnalls, New York (1910)
85. von Fintel, K.: Bare plurals, bare conditionals, and only. J. Semant. **14**, 1–56 (2007)
86. Williams, C.J.F.: Saint Anselm and his biographers. Downside Rev. **82**, 124–140 (1964)
87. Williams, J.R.G.: Defending conditional excluded middle. Noûs **44**, 650–668 (2010)

L. Horn (✉)
Departments of Linguistics and Philosophy, Yale University, New Haven, USA
e-mail: laurence.horn@yale.edu

Aristotle on Language and Universal Proof

Jean-Louis Hudry

Abstract This paper shows that Aristotle's conception of language is incompatible with a divide between syntax and semantics. Language for Aristotle is not reducible to a formal syntax of names and sentences, since these are linguistic entities if and only if they have linguistic meanings by convention. Nor is there an abstract semantics, in so far as Aristotle's *De Interpretatione* resorts to a distinction between two kinds of meaning. On the one hand, linguistic meanings are by convention, so that names, verbs, and sentences are meaningful spoken sounds, and Aristotle concludes that linguistic meanings cannot be the same for all. On the other hand, non-linguistic thoughts are about mental contents, and when they are related to actual things, they are the same for all, because actual things are the same for all. For instance, the non-linguistic thought of snow is the same for all, but the linguistic meaning of the name 'snow' is not the same for all. While linguistic meanings rely on the learning of linguistic conventions, non-linguistic meanings are mental contents, derived from perceptions. In that respect, a mental conception of meaning is very different from a semantic theory of meaning. Nevertheless, the absence of syntax and semantics does not prevent Aristotle from developing a theory of deduction based on universal proofs, so that scientific explanations are made possible in the context of demonstrations. Accordingly, demonstrative knowledge is about explanatory middle terms, whose formalization allows Aristotle to establish a universal discourse for science.

Keywords Syntactic language · Linguistic meaning · Semantic proposition · Mental content · Demonstration · Middle term · Explanation

Mathematics Subject Classification Primary 03-02 · Secondary 03B65

Łukasiewicz [22] identifies Aristotle's deductions with universalized conditional propositions, so that the logic of deductions is a system of true propositions (cf. Miller [23], Bocheński [6], Patzig [25]). He translates "syllogistic necessity" into a universal quantifier: "the Aristotelian sign of syllogistic necessity represents a universal quantifier and may be omitted, since a universal quantifier may be omitted when it stands at the head of a true formula" [22, p. 11]. However, Aristotle cannot make sense of universal quantifiers. That is why Łukasiewicz immediately adds: "This, of course, is all known to students of modern formal logic, but some fifty years ago, it was certainly not known to philosophers" [22, p. 11].[1] This remark underlines Łukasiewicz's awareness that a modern conception

[1] Regarding the different conceptions of 'logic' and 'formal logic', see Beziau [4, 5]. See also Woods [34] and Hudry [16–19].

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_11

of logic uses tools that Aristotle could not have known. In that respect, Łukasiewicz suggests a reconstruction of Aristotle's logic as propositional logic, without implying that Aristotle himself elaborates a propositional logic. Nowadays, it is standard to identify the inferential necessity of Aristotle's deductions with logical validity. For instance, Keyt [20, p. 26] writes: "the conclusion of a syllogism follows 'of necessity' from its premises: only *valid* arguments are syllogisms" (cf. also Smiley [30]). An argument is valid if its premises logically entail its conclusion. This logical reconstruction is based on a criticism of Łukasiewicz, initiated by Corcoran [9, 10, 12, 14] and Smiley [29], so that the logical truth of the conditional proposition is replaced with the logical validity of the natural deduction (cf. also Smith [31]).

In both cases, to impose the logical concept of 'universal quantifier' or 'logical validity' on the *Prior Analytics* [28] would amount to neglecting the conceptual gaps between Aristotle and Frege in their respective understandings of logic. The main obstacle to the identification of Aristotle's system of deductions with modern logic is the absence of a divide between syntax and semantics. In other words, there is no way to infer the formal universality of a logical structure. The only universality introduced by Aristotle is related to his own conception of science, understood as universal explanation or proof in the context of demonstrative knowledge.

1 Linguistic Meaning vs. Syntactic Language

It is often said that Aristotle implicitly resorts to a metalinguistic rule, producing the abstract syntax of a generic sentence. For instance, Boger [7] writes:

> From his notion of sentence in *On Interpretation*, we can extract Aristotle's rule for the formation of a generic sentence in a natural language and express it as follows: a sentence in a given natural language consists in combining a noun and a verb (i.e., a predicate) in certain ways so as to produce a meaningful expression. This rule identifies the broadest *pattern* of a sentence in a natural language [7, p. 120, original emphasis].

This syntactic rule reduces a sentence to the combination of a syntactic name with a syntactic verb. Yet, *De Interpretatione* [24] never makes such a claim. Boger is ready to meet the objection, as he adds that Aristotle's arguments are too confusing, which prevents Aristotle himself from clearly distinguishing syntax from semantics.[2] In other words, we are asked to suppose that Aristotle genuinely aims to separate syntax from semantics, even though he is unable to spell out a syntactic rule explicitly. Boger reconstructs the argument in order to make it compatible with his main claim that there is a divide between syntax and semantics in Aristotle's underlying logic. Yet, is it really Aristotle's original position?

The view that Aristotle resorts to a syntactic rule faces a serious objection, since *De Interpretatione* has no reason to introduce a syntactic language. Indeed, Aristotle does not think of syntactic entities as belonging to linguistics. He deals with a spoken language,

[2] Boger [7] writes: "We might wish that Aristotle had expressed this rule with at least the modest precision here. However, Aristotle has neither a complete nor a complex set of syntax rules of sentence formation in *On Interpretation*. Still, it is evident from his treatment of this topic in his logical investigations that his understanding of the grammar of natural language is richer than his lack of rigorously stated rules would indicate. And while this syntax rule is mixed with semantic notions, he nevertheless has identified here the basic patterns of a sentence in natural language" (p. 120).

based on meaningful spoken sounds. A name (*onoma*) is a "meaningful spoken sound" (*phônê sêmantikê*) (2, 16a19), along with a verb (*rhêma*) that is also a meaningful spoken sound (3, 16b19–20). Names cannot be syntactic entities, since their linguistic status derives from their meaning by convention:

> A name is a spoken sound meaningful by convention (*kata sunthêkên*)... I say 'by convention' because no name is a name by nature (*phusei*) but only when it has become a symbol (*sumbolon*). (*De Interpretatione*, 2, 16a19, a26–28)[3]

Names are symbols if and only if they have meanings by linguistic conventions. In *Historia Animalium* (IV, 9) [3], Aristotle grants that many animals use spoken sounds, but only humans have the ability to produce linguistic conventions: "Viviparous quadrupeds utter spoken sounds (*phônên*) of different kinds, but they have no language (*dialekton*): this is particular to human being; for while whatever has language has spoken sounds, not everything that has spoken sounds has language" (536b32–537a3). Spoken sounds have natural meanings; for instance, the roar of the lion means something for other lions, or the barking of the dog means something for other dogs. Thus, *De Anima* asserts that a spoken sound (*phônê*) has a natural meaning conveyed by an act of imagination (for Aristotle, many animals have the ability to produce and to memorize images).[4] However, these natural meanings have no role to play in language. *De Interpretatione* says: "The inarticulate sounds (*agrammatoi psophoi*), e.g. of beasts, do indeed designate (*dêlousi*) something, yet none of them is a name" (2, 16a28–29). Only a linguistic convention is able to transform a spoken sound into a name, namely a spoken sound meaningful by convention. Not to understand a specific language amounts to hearing spoken sounds meaningless by convention. They become names and verbs as soon as the conventions of a given language are learnt. Aristotle then notices that the hearing of names, as meaningful spoken sounds, is connected with the growth of intelligence.[5] Aristotle also insists on the meaning of names by convention, when the *Poetics* (21, 1457b1–7) tells us that a name is either proper (*kurion*) or alien (*glôtta*). A proper name depends on a linguistic convention used in a given country, whereas an alien name refers to a different linguistic convention in use elsewhere. The essential claim is that a name can be both proper and alien, but it cannot be so in reference to the same speaker. For instance, the name '*sigunon*' ('spear') is proper for the Cypriots, but alien for the Greeks, meaning that the status of names varies in function of the speakers' ability to use linguistic conventions.

If we understand a name as a syntactic entity, it is impossible to grant meaning to a name, as implied by Boger's [7] interpretation: "Aristotle's syntax language specifies, ab-

---

[3] All translations in this paper are my own, unless otherwise indicated.

[4] *De Anima* asserts: "A spoken sound (*phônê*) is the impact of the inbreathed air against the windpipe [i.e. the organ of respiration], and the agent that produces the impact is the soul resident in these parts of the body. Not every sound (*psophos*), as we said, made by an animal is a spoken sound (even with the tongue we may merely make a sound which is not a spoken sound, or without the tongue as in coughing). What produces the impact must have soul in it (*empsuchon*) and must be accompanied by an act of imagination (*phantasias*), for the spoken sound is a sound with a meaning (*sêmantikos gar dê tis psophos estin hê phônê*), and is not the result of any impact of the breath as in coughing." (II, 8, 420b27–33).

[5] *De Sensu* states: "It is hearing that contributes most to intelligence (*phronêsin*). For the discourse (*logos*) is a cause of instruction (*mathêseôs*) in virtue of its being audible, which it is, not in its own right, but incidentally; for it is composed of names (*onomatôn*), and each name is a symbol (*sumbolon*)." (1, 437a11–15).

stractly, only nouns and verbs as its vocabulary, which are combined to form sentences according to this elementary rule" (p. 120). The identification of spoken sounds with syntactic elements is similar to Quine's formal grammar of syntactic phonemes.[6] Yet, Aristotle cannot allow syntactic phonemes to be parts of language, since to separate spoken sounds from their meanings would amount to denying their linguistic status. A spoken sound is a linguistic symbol only through its meaning by convention. Without such a meaning, a spoken sound is like the barking of a dog, and has nothing to do with language. Therefore, Aristotle cannot postulate a divide between syntax and semantics, in so far as such a divide threatens his conception of names and verbs as linguistically meaningful entities.

Likewise, a sentence (*logos*) is understood as a meaningful spoken sound, and results from the combination of a name with a verb, namely the combination of two meaningful spoken sounds:

> A sentence is a meaningful spoken sound (*phônê sêmantikê*), some part of which is meaningful in separation, as an assertion (*phasis*), not as an affirmation (*kataphasis*). Let me explain. 'Human', for instance, means something, but not that it is or is not (though it will be an affirmation or denial if something is added); but the single syllables of 'human' are meaningless. Nor is the 'ice' (*us*) in 'mice' (*mus*) meaningful; here it is a mere spoken sound (*phônê*) (*De Interpretatione*, 4, 16b26–32).

A syllable is a spoken sound without being a linguistically meaningful entity, as it has no meaning by convention. For instance, the meaningless syllable 'ice' in 'mice' is distinct from the meaningful spoken sound 'ice'.[7] Chapter 20 of the *Poetics* holds the same view; that is, a syllable is a meaningless spoken sound (*phônê asêmos*, 1456b35–36), unlike a name regarded as a meaningful spoken sound (*phônê sêmantikê*, 1457a11–12). Aristotle illustrates his position with another example: the name '*Theodôros*' includes the syllable *dôros*, which does not mean anything, as opposed to the name '*dôros*', which means a gift by convention. Therefore, a spoken sound is a linguistic entity, when meaningful by convention. If meaningless by convention, it is a mere spoken sound without linguistic status. In this framework, there is no reason for distinguishing syntax from semantics, since Aristotle does not accept syntactic symbols as linguistic entities, owing to the absence of meaning attached to them.

The second claim made by the above passage from *De Interpretatione* is that, while a name or a verb is an assertion (*phasis*), a sentence is either an affirmation (*kataphasis*) or a denial (*apophasis*), namely a meaningful spoken sound that either affirms or denies something of something else. The *Poetics* (20, 1457a23–26) also claims that sentences are meaningful spoken sounds; yet, it allows sentences to be "without a verb" (*aneu rhêmatôn*). Aristotle uses the account of man as an example, such that the expression 'two-footed terrestrial animal' is deemed to be a sentence, despite the absence of a verb. In this case, a sentence is reducible to a collection of names; but then, it is difficult to understand how such a sentence can be regarded as an affirmation. The ambiguity is removed, when *De Interpretatione* spells out that a sentence (*logos*) is not the same as a declarative sentence (*logos apophantikos*) (4, 17a2–7). That is, a sentence becomes declarative, when a verb is added:

---

[6] According to Quine's *Philosophy of Logic* [27], the grammarian has "to demarcate formally, in a reasonably simple and natural way, a class of strings of phonemes which will include practically all observed utterances and exclude as much as practicable of what will never be heard" (p. 22).

[7] The Greek name '*mus*' means a field-mouse, while the name (not the syllable) '*us*' means a pig.

> It is necessary that every declarative sentence (*panta logon apophantikon*) has as its part a verb or an inflection (*ptôseôs*). For even the definition of man is not yet a declarative sentence, unless 'is', 'will be', 'was' or something of this sort is added. (*De Interpretatione*, 5, 17a9–12)

Affirmation and denial are the two main declarative sentences (17a8–9), and both require the combination of a name with either a verb (in the present tense) or an inflection (i.e. a verb in the past or future tense). Aristotle is very clear on this: "without a verb there is neither affirmation nor denial" (*aneu de rhêmatos oudemia kataphasis oud' apophasis*) (10, 19b13). In other words, there is no discrepancy between the *Poetics* and *De Interpretatione*, so that 'two-footed terrestrial animal' is a sentence without a verb, which is not an affirmation, namely a declarative sentence.

A verb means something (*semainei ti*) and, on its own, is not different from a name (3, 16b19–21).[8] It acquires its linguistic specificity only when combined with a name, in so far as it "adds in meaning some combination" (*prossêmainei sunthesin tina*) (16b23–25). A verb adds not only time in meaning but also a relation of predication, such that a subject is predicated of something else. Hence, a verb is the linguistic sign (*sêmeion*) of a predication:

> A verb (*rhêma*) adds time in meaning (*prossêmainei chronon*), no part of it being meaningful separately; and it is a sign of things said of something else. I am saying that it adds time in meaning: 'recovery' (*hugieia*) is a name, but 'recovers' (*hugiainei*) is a verb, for it adds in meaning something's holding now (*to nun*). It is always a sign of what holds, that is, holds of a subject (*De Interpretatione*, 3, 16b6–10).

As the verb accounts for a relation of predication in time, an affirmation or a denial is always temporally determined. While verbs add in meaning the present, inflected verbs (*ptôsis rhêmatos*) add in meaning the past or future, which are tenses surrounding (*perix*) the present (3, 16b16–18). The *Poetics* is compatible with *De Interpretatione*, since it identifies a verb with a meaningful spoken sound, and while a name is without time (*aneu chronou*), a verb has to be connected with time (*meta chronou*) (20, 1457a14–19). Aristotle contrasts the name 'man' or 'white', which does not mean anything temporal (*ou sêmainei to pote*), with the verb 'walks' or 'has walked', which adds in meaning (*prossêmainei*) the present or the past. As for *De Interpretatione*, it concludes that a declarative sentence is a meaningful spoken sound about whether something holds or does not hold in accordance with temporal divisions (cf. 5, 17a23–24). In that respect, Aristotle could not have accepted tenseless verbs, as they would have violated the grammatical conventions of the Greek language. Therefore, his position contradicts the tenseless verbs of a formal grammar or syntax.[9]

To sum up, neither *De Interpretatione* nor the *Poetics* allow one to think of names, verbs, and sentences as syntactic entities, since meaningful spoken sounds are not syntactic phonemes. It is maybe tempting to reconstruct Aristotle's views by assuming a separation between syntax and semantics, but this reconstruction contradicts Aristotle's

[8] On the notion of a verb taken in isolation, i.e. not embedded in a sentence, see Wagner [32], Ax [2], and Weidemann [33].

[9] Quine [27] writes: "Our standard logical grammar is conspicuously untouched by the complications of *tense* which so dominate European languages. Logical grammar, like modern physics, is best served by treating time as a dimension coordinate with the spatial dimensions; treating dates, in other words, as just another determinable on a par with position. Verbs can then be taken as tenseless" (p. 30, original emphasis).

initial position. For instance, consider the view that *spoken sounds used in language* are initially meaningless. This claim seems to be sound and harmless. Yet, Aristotle would have rejected it as a contradiction in terms. If spoken sounds are used in language, they must have linguistic meanings by convention, but if they are assumed to be initially meaningless, they cannot pertain to linguistic conventions, and they cannot be part of language. Hence, the assumption that a language may contain syntactic phonemes makes no sense for Aristotle.

2 Mental Content vs. Semantic Proposition

The idea of a syntactic language in Aristotle leads some to acknowledge an abstract semantics in the form of non-linguistic propositions. For instance, Boger [7] understands Aristotle's notion of declarative sentence as a sentence expressing a non-linguistic proposition:

> Aristotle considers only those kinds of sentence that are either true (*alethês*) or false (*pseudês*); or, as we express this nowadays, he considers only those sentences that have a truth-value (16a9–13). His explicit interest is only with the kind of sentence that expresses a proposition, namely, with the declarative sentence [7, p. 120].

Boger is well aware that Aristotle does not speak of propositions as such; that is why he suggests that the terms '*kataphasis*' ('affirmation') and '*apophasis*' ('denial') may be interpreted as "sentences expressing propositions". In spite of this textual ambiguity, he wants to identify meaning, as expressed by a linguistic sentence, with a non-linguistic proposition:

> If we take 'proposition' more loosely to denote the *meaning* of a declarative sentence, we can easily see that Aristotle made and worked with a distinction between a sentence, which is a linguistic object, and the meaning or proposition it expresses, which is a non-linguistic object. Both of these, of course, he distinguished from what a sentence denotes, which is a state of affairs (*pragma*), that is, something that obtains or does not obtain in the world [7, p. 126, original emphasis].

According to this theory of meaning, natural language is about linguistic sentences (syntax) expressing non-linguistic propositions (semantics), and these sentences and propositions are distinct from states of affairs (denotations). If we agree with this view, we have no other choice than to conclude that Aristotle's position is very similar to Russell's, since both distinguish the syntax of linguistic sentences from the semantics of non-linguistic propositions.

We may object to the idea of an abstract semantics in Aristotle by connecting it with the absence of syntactic language in Aristotle. If spoken sounds cannot constitute a formal syntax of phonemes, there is no reason to assume an abstract semantics. Indeed, there cannot be an abstract semantics, if there is no syntactic language. What about an affirmation or a denial? Do these declarative sentences have a special linguistic status? It is true that Aristotle makes a clear distinction between declarative sentences (which are either true or false) and non-declarative sentences (which are neither true nor false):

> Every sentence is meaningful, not as a tool (*ouk hôs organon*), but as we have said, by convention (*kata sunthêkên*); and not every sentence is a declarative sentence, but only those in which there is truth or falsehood. We cannot say it about all sentences: a prayer is a sentence but is neither true nor false. Let us leave aside all the other sentences (since consideration of them belongs rather to the domain of rhetoric or poetry). Only the declarative sentence belongs to the present study (*De Interpretatione*, 4, 16b33–17a7).

Whether sentences are declarative or not, all of them are spoken sounds meaningful by convention. When the above passage asserts that the meaning of a sentence is not a tool, he implicitly criticizes Plato's position, which views names and sentences as ontological tools, mirroring the nature of things they signify. According to Plato, a thing may be designated by a plurality of names in different languages; yet, this linguistic fact does not conflict with the ontological principle that a name in a spoken language may refer to an intelligible Form (cf. *Cratylus*, 388b–c, 389d–390a). For instance, the English name 'snow' constitutes an ontological tool in relation to its intelligible definition, namely the Form of Snow. Aristotle rejects such a position by claiming that the meanings of names and sentences do not tell us anything about the nature of things, as they are meaningful only by linguistic convention. The name 'snow' has a linguistic meaning in the context of the English language. The ignorance of this language makes the spoken sound 'snow' meaningless. By rejecting the idea that the meanings of names and sentences may be ontological tools, Aristotle denies all references to ontology in the structure of language.

Consequently, Aristotle has no reason to postulate the existence of non-linguistic meanings or propositions, as expressed by syntactic sentences. Indeed, names, verbs, and sentences have linguistic meanings on their own, and the absence of syntactic sentences in Aristotle's conception of language prevents non-linguistic meanings or propositions to be expressed by such sentences. In other words, the truth or falsehood of a declarative sentence does not appeal to an abstract semantics of propositions. This view is confirmed in the *Categories*, which shows that the truth or falsehood of a declarative sentence exclusively depends upon actual things:

> Suppose, for example, that the sentence (*logos*) that somebody is sitting is true; after he has got up this same sentence will be false. Similarly with beliefs (*doxês*). Suppose you believe truly that somebody is sitting; after he has got up you will believe falsely if you hold the same belief about him. [. . .] Sentence and belief remain completely unchanged in every way; it is because of a change of the actual thing (*pragmatos*) that the contrary comes to belong to them. For the sentence that somebody is sitting remains the same; it is because of a change of the actual thing that it comes to be true at one time and false at another. Similarly with beliefs (*Categories*, 5, 4a24–28, 4a34–b1).

Truth and falsehood pertain to actual things, so that a same sentence may be true at one time, and false at another, depending upon a change of the actual thing to which the sentence refers. For instance, the meaningful spoken sound 'Somebody is sitting' is true at t_1 if and only if an individual is actually sitting at t_1, while the same meaningful spoken sound is false at t_2 if and only if this same individual is actually standing at t_2. As a result, only actual things (or states of affairs) make declarative sentences either true or false, without the need for Aristotle to rely on non-linguistic propositions.[10] The irrelevance of non-linguistic propositions may be illustrated by the following instance. 'Snow is white' and 'La neige est blanche' are two meaningful spoken sounds. Their linguistic meanings are not the same, since a speaker may understand one sentence without understanding the other. Even in the case in which we grasp the two different linguistic meanings, we are still unable to draw a synonymy relation between both sentences. Only the reference to an actual thing (i.e. actual snow) enables one to infer that both sentences are not only true but also synonymous, despite their different linguistic meanings. Therefore, actual things

[10] See Crivelli [15] analyzing Aristotle's concept of truth in relation to states of affairs.

provide an objective, ontological foundation, without implying an abstract semantics of non-linguistic propositions.[11]

According to some, the abstract semantics of propositions should provide Aristotle with non-linguistic meanings or thoughts, as expressed by spoken sounds understood as syntactic phonemes. For instance, Boger [7] identifies *phônê sêmantikê* (translated as 'meaningful expression') with a complete thought:

> Aristotle treats sentence formation in a natural language as essentially consisting in combining (*sunthesis*) a noun (*onoma*) and a verb (*rhêma*; i.e. a predicate [16a9–18]) so as to produce a meaningful expression (*phônê sêmantikê*), a complete thought [7, p. 119].

That is, a complete thought (abstract semantics) applies to a linguistic expression (syntactic language), so that the expression is said to be meaningful. Aristotle cannot agree with such a claim, since the linguistic meaning of *phônê sêmantikê* (which should be translated as 'meaningful spoken sound') derives from linguistic conventions. To identify *phônê sêmantikê* with a complete thought would indicate that thoughts are also by convention. Nobody could agree with such a claim, and it is certainly not what Boger intends to say, as he is well aware that thoughts for Aristotle are not linguistic, and cannot be by convention.

The crux of the problem lies in the interpretation of these so-called non-linguistic thoughts. To understand them as non-linguistic propositions means that there is an abstract semantics of propositions distinct from language itself, and this makes sense if language is regarded as syntactical. In contrast, we want to show that Aristotle identifies non-linguistic thoughts with mental contents, which are distinct from the linguistic meanings of spoken sounds. While spoken sounds are meaningful by convention, non-linguistic thoughts are meaningful by nature. Nevertheless, the linguistic meanings of spoken sounds are correlated with non-linguistic thoughts. Thus, Aristotle asserts in one of the most important claims in *De Interpretatione* that spoken sounds, which are not the same for all, are symbols of the affections of the soul, which are the same for all:

> Spoken sounds are the symbols of the affections of the soul (*têi psuchêi pathêmatôn*), and written marks (*graphomena*) are the symbols of spoken sounds. Just as letters (*grammata*) are not the same for all, neither are spoken sounds. But what these are in the first place signs of, i.e. affections of the soul, are the same for all; and what these affections are likenesses of, i.e. actual things (*pragmata*), are also the same. These matters have been discussed in the work on the soul (*peri psuchês*) and do not belong to the present subject (*De Interpretatione*, 1, 16a3–9).

A language for Aristotle is, above all, a meaningful spoken language, and this explains why written words are symbols of spoken sounds. The essential claim is that language by

[11] In fact, not all modern logicians hold the view that linguistic sentences express non-linguistic propositions. For instance, Quine [27] speaks of proposition as "a widespread myth of meaning" (p. 8). He criticizes propositions for the following reason: "Once a philosopher, whether through inattention to ambiguity or simply through an excess of hospitality, has admitted propositions to his ontology, he invariably proceeds to view propositions rather than sentences as the things that are true and false… My objection to recognizing propositions does not arise primarily from philosophical parsimony—from a desire to dream of no more things in heaven and earth than need be. Nor does it arise, more specifically, from particularism—from a disapproval of intangible or abstract entities. My objection is more urgent. If there were propositions, they would induce a certain relation of synonymy or equivalence between sentences themselves: those sentences would be equivalent that expressed the same proposition. Now my objection is going to be that the appropriate equivalence relation makes no objective sense at the level of sentences. This, if I succeed in making it plain, should spike the hypothesis of propositions." (pp. 2–3).

convention, whether the language is spoken or written, are not the same for all. However, a meaningful spoken sound (whether a name, a verb or a sentence) is the symbol (*sumbolon*) or sign (*sêmeion*) of an affection of the soul (symbol and sign being synonymous terms in this context).[12] Linguistic symbols are very different from the things they stand for. On the one hand, spoken sounds have linguistic meanings by convention, which are not the same for all. On the other hand, affections of the soul have non-linguistic meanings by nature, which are the same for all. Affections of the soul are understood as likenesses (*homoiômata*) of actual things, and they are the same for all because actual things are the same for all. We may illustrate Aristotle's position with the following instance. The word 'snow' is the written symbol of the meaningful spoken sound 'snow', and they are not the same for all, since only people knowledgeable in the English language grasp its linguistic meaning. Yet, the name 'snow' is the spoken symbol of an affection of the soul in relation to an actual thing, namely the non-linguistic thought of snow in relation to actual snow, and both the thought and the actual thing are the same for all. Accordingly, the linguistic meaning (which is not the same for all) of the name 'snow' is not to be confused with the non-linguistic thought (which is the same for all) in reference to actual snow.

It is always possible to reconstruct Aristotle's position in order to make it compatible with the logical investigations of a natural language. Thus, Aristotle's distinction between meaningful spoken sounds (which are not the same for all) and affections of the soul (which are the same for all) is reinterpreted by Boger [7] as a logical divide between the syntax of linguistic sentences and the semantics of non-linguistic propositions (thoughts):

> He [Aristotle] clearly indicates here [*De Interpretatione*, 16a3–9] his distinguishing very different linguistic objects as expressing the same meaning or expressing the same proposition—that peculiar, non-linguistic thing that is grasped by a human being in thought (p. 126).

An affection of the soul is identified with a non-linguistic thought or proposition, such that propositions correspond to an abstract semantics, revealing the logical structure of all natural languages. While a linguistic syntax is specific to a particular natural language, the non-linguistic semantics is common to all natural languages. It is interesting to note that Corcoran [13] does not go as far in his own reconstruction of Aristotelian logic, since he doubts that Aristotle may conceive of an abstract semantics:

> This passage [*De Interpretatione*, 16a3–9], one of the most important in the history of semantics, should probably not be construed as involving anticipations of modern abstracts, as opposed to mental conceptions of meaning, which began to emerge about the time of Boole [13, p. 266, footnote 14].

Corcoran endorses the view that non-linguistic meanings for Aristotle are nothing more than mental contents.[13] We must wait for the 19th century with philosophers like Boole [8] and Frege to see the mental conception of meaning challenged by a new philosophy of language. It is, therefore, quite anachronistic to identify Aristotle's affections of the

[12] Aristotle's notions of symbol and sign have received numerous interpretations since the Antiquity. See Ackrill [1] and Pépin [26]. For a carefully crafted interpretation, see Kretzmann [21].

[13] Corcoran [11] defends a similar view in relation to the *Prior Analytics*: "[Since] Aristotle nowhere makes specific reference to alternative interpretations… it seems that at every point he thought of his ideal language as interpreted in what we would call its intended interpretation. Moreover, it is doubtful that Aristotle ever conceived of a language apart from its intended interpretation. In other words, it seems that Aristotle did not separate logical syntax from semantics" (p. 104).

soul with an abstract semantics of non-linguistic propositions. In that respect, Corcoran [13] attempts to reassess the distinction between a linguistic sentence and a non-linguistic proposition:

> As a first approximation, we can think of a sentence as a series of inter-related written words and we can think of a proposition as a series of inter-connected meanings or concepts that may or may not have been expressed in words. Aristotle had already said that although sentences are not the same for all humans what they express is the same for all [13, p. 266].

Corcoran redefines the notion of proposition to make it compatible with Aristotle's mental conception of meaning. A proposition is said to be "a series of inter-connected meanings or concepts", which "may or may not have been expressed in words". Thus, if a proposition is not expressed by "a series of interrelated written words", it cannot be expressed by a sentence. A proposition may then be identified with a mental content, without reference to the abstract semantics of a syntactic language. Therefore, Corcoran's "approximation" allows one to identify Aristotle's mental conception of meaning with the notion of proposition. That is, non-linguistic propositions are mental contents, whereas linguistic sentences have meanings by linguistic convention. It is then possible to claim that *De Interpretatione* (cf. 16a3–9) distinguishes two kinds of meaning: while spoken languages are about linguistic meanings by convention (which are not the same for all), mental contents are about non-linguistic meanings (which are the same for all). It is also clear that such a distinction does not amount to the divide between syntax and semantics.

3 Middle Term as Universal Proof

It may be difficult to accept such a conception of language, which ignores the separation between syntax and semantics, since any introductory course of logic refers to this separation as an obvious principle accepted by all. Yet, there is nothing obvious about syntax and semantics. Boole, Frege, and others from the 19th century were struggling to understand these new concepts. Hence, from a historical standpoint, it should not be said that Aristotle was blindly and mysteriously able to foresee a divide between syntax and semantics, while he had no reason to postulate such a divide. Mainly, this modern reconstruction of Aristotle's logic hides the originality of his theory of deduction, which promotes a method aiming to produce universal explanation.

The *Prior Analytics* develops three deductive figures, but only the first one is deemed to be compatible with demonstrative knowledge. Indeed, demonstrations in the first figure theorize the reason why (*to dioti*) a thing is, so that they are able to explain what a thing is (*to ti esti*). Chapter 14 of *Posterior Analytics A* [28] asserts:

> Of the figures, the first is the most scientific. For the mathematics, among the sciences, produce demonstrations through this [figure], e.g. arithmetic, geometry, and optics, and so to speak almost all [sciences] that make investigation of the reason why (*tou dioti*); for the deduction of the reason why (*tou dioti*) is in general or for the most part, and in most cases, through this figure. Hence, because of this, the figure is the most scientific; for what is most important for knowledge is to theorize the reason why. Next, it is possible to pursue the science of what a thing is (*tou ti estin*) only through this figure. For in the middle [second] figure no predicative (*katêgorikos*) deduction comes to be, while the science of what a thing is (*tou ti estin*) is predicative; and in the last [third] figure what comes to be is otherwise not universal, while what a thing is (*to ti esti*) is of the universal; for it is not in some respect that man is a two-footed animal. Furthermore, this [first]

> figure has no need of the others, while these [second and third figures] are thickened and increased through this [first figure] until they come to the immediates. Hence, it is evident that the first figure is the most important for scientific understanding (*An. Post. A*, 14, 79a17–32).

Demonstrations in the first figure amount to demonstrative knowledge, because only the first figure provides demonstrations that are also explanations. While the deduced conclusion asserts that something is (*to hoti*), the middle term explains the reason why (*to dioti*) this thing is, and from this explanation we know what this thing is (*to ti esti*). Among the four kinds of demonstration in the first figure, only *Barbara* and *Darii* are predicative (*katêgorikos*) deductions, meaning that their conclusions are affirmations about predicates:

(Barbara) *Every middle term B is the major term A.*
Every minor term C is the middle term B.
It results of necessity that every minor term C is the major term A.

(Darii) *Every middle term B is the major term A.*
Some minor term C is the middle term B.
It results of necessity that some minor term C is the major term A.

On the other hand, the two other deductions of the first figure are not predicative (i.e. privative), so that their deduced conclusions are denials about predicates:

(Celarent) *No middle term B is the major term A.*
Every minor term C is the middle term B.
It results of necessity that no minor term C is the major term A.

(Ferio) *No middle term B is the major term A.*
Some minor term C is the middle term B.
It results of necessity that not every minor term C is the major term A.

When Aristotle speaks of demonstrations in the first figure, and that they are meant to account for the reason why (*to dioti*), he refers only to *Barbara* and *Darii*, since *Celarent* and *Ferio* do not say anything positive about the predicate A in relation to the subject C.

Aristotle uses the same justification to reject all the deductions in the second figure, since the deduced conclusions of *Cesare*, *Camestres*, *Festino*, and *Baroco* are not predicative (affirmative). Likewise, the deductions in the third figure cannot produce demonstrative knowledge. Indeed, some have deduced conclusions that are not predicative: *Felapton*, *Bocardo*, *Ferison*. Others have predicative conclusions that are not universal: *Darapti*, *Disamis*, and *Datisi*. The investigation of what a thing is (*tou ti esti*) implies definitions that have to be universals. For instance, the definition of man, as a two-footed animal, cannot be in some respect. This view indicates that, in the first figure, *Barbara* is a better demonstration than *Darii*, since only the predicative conclusion of *Barbara* is universal (cf. above). In Chap. 24 of the *Prior Analytics*, Aristotle acknowledges that the universal demonstration is better:

> Besides, if demonstration is a deduction that proves an explanation (*aitias*), i.e. the reason why (*tou dia ti*), and as the universal is more explanatory (*aitiôteron*)... then the universal demonstration is better; for it is more about the explanation, i.e. the reason why (*An. Post. A*, 24, 85b23–27).

The universal deduction, owing to its universal conclusion (cf. *Barbara*), is better than the particular deduction, relying on its particular conclusion (cf. *Darii*). Yet, in both kinds of deduction, the reason why (*to dioti*) is about a middle term, which is universal in the major premise of either *Barbara* or *Darii*:

(Barbara) *Every middle term B is the major term A.*
(*Darii*) *Every middle term B is the major term A.*

The major term A is predicated of every middle term B, so that A is a *per se* predicate of B. This means that there cannot be some B without A; e.g. 'every man is an animal' means that there cannot be an explanation of man without involving an explanation of animal. The distinction between *Barbara* and *Darii* is explained by the mediation introduced by the minor term C in relation to the middle term B, so that there are two distinct kinds of minor premise:

(*Barbara*) *Every minor term C is the middle term B.*
(*Darii*) *Some minor term C is the middle term B.*

In *Barbara*, the middle term B is the predicate of a universal minor term C. In *Darii*, the middle term B is the predicate of a particular minor term C. In both cases, the explanation B in relation to the major term A is universal, but the universality or the particularity of the minor term C makes the demonstration either universal or particular. In that respect, the universal demonstration *Barbara* is more explanatory than the particular demonstration *Darii*, since the explanation applies to all minor terms C in *Barbara*, but only to some minor terms C in *Darii*.

Explanation has to be based on universals, namely particulars universalized through induction. Chapter 31 of *Posterior Analytics A* claims that explanation is impossible, if confined to the perception of a particular:

> Hence, since demonstrations are universal, and these are not perceived, it is evident that there is no scientific understanding (*epistasthai*) through perception either, but it is clear that, even if we perceived the triangle as having its angles equal to two right angles, we would search for a demonstration and would not understand it, as some say; for we necessarily perceive particulars, but knowledge (*epistêmê*) amounts to cognizing (*gnôrizein*) the universal. That is why, if we were on the moon and saw the earth screening it, we would not know the explanation (*aitian*) of the eclipse. For we would perceive the reason why now (*dioti nun*) it was eclipsed, and not the reason why in general (*dioti holôs*) it was [eclipsed]; for there is no perception of the universal. Yet if, from observing that this happens many times, we hunted the universal, we would have a demonstration; for from several particulars the universal is clear. The universal is valuable because it makes clear the explanation (*aition*); so that, with respect to things whose explanation is different, the universal demonstration is more valuable than perception and intellection (*noêseôs*); but with respect to major premises, there is another account (*An. Post. A*, 31, 87b33–88a8).

Starting from the end of this passage, Aristotle distinguishes the major premise A from the middle term B, namely the *explanandum* as distinct from the *explanans*. The immediate major premise is an indemonstrable principle of demonstration. It depends upon induction and intellection. It is better cognized than demonstration, as intellection is prior to, and truer than, demonstrative knowledge.[14] Therefore, intellection accounts for an indemonstrable major premise (asserting that the major term A is predicated of every middle term

[14]Chapter 19 of *Posterior Analytics B* asserts: "Since of the thinking states (*tên dianoian hêxêôn*) by which we grasp truth, some are always true and some admit falsehood, e.g. opinion (*doxa*) and computation (*logismos*), but knowledge (*epistêmê*) and intellection (*nous*) are always true; and since no kind other than intellection is more precise (*akribesteron*) than knowledge, and since the principles are better cognized (*gnôrimôterai*) than demonstrations, and since all knowledge involves an account (*logos*), then there will

B), which is a principle of demonstration, whereas demonstration itself pertains to an explanation (*aition*), which is the reason why (*dioti*) the major term A is predicated of the minor term C. In other words, the immediate (*ameson*) premise, as a principle, has to be contrasted with the middle term (*meson*), as an explanation. The above passage underlines the universality of explanation, so that only a universal explanation plays the role of a proof. In that respect, the perception of a particular is powerless to explain anything. Aristotle suggests two instances. First, we may easily perceive that a triangle has angles equal to two right angles, and yet we do not prove anything, as the object of perception is a particular triangle. Second, we may imagine that, being on the moon, we perceive the interposition of the sun by the earth. In this hypothetical case, we perceive the reason why now there is a lunar eclipse, and yet no explanation can be inferred from this perception. Indeed, "the reason why now" (*dioti nun*) amounts to the perception of a particular event at an instant of time, which cannot constitute scientific knowledge. Only a universal explanation provides "the reason why in general" (*dioti holôs*). This implies a process of induction, based on the perception of a particular, so that we infer 'every interposition of the sun by the earth' from 'some interposition of the sun by the earth'. This induced fact is a universal, as it is based on the many perceptions of a particular. Aristotle concludes: "from several particulars, the universal is clear" (cf. above). When a universal is induced from many particulars, it is then possible to provide a universal explanation.

The *Posterior Analytics* precisely describes how the search for the middle term amounts to a scientific investigation, in so far as the deduced conclusion of the demonstration is knowledge if and only if there is an explanation for this conclusion:

> When we search for the fact that (*to hoti*) or if a thing is *simpliciter* (*to ei estin haplôs*), we search whether there is or is not a middle term (*meson*) for it; and when we have cognized (*gnontes*) either the fact that or if a thing is, either in part (*epi merous*) or *simpliciter*, next we search for the reason why a thing is (*to dia ti*) or what a thing is (*to ti esti*), we then search for what the middle term is (*ti to meson*). Here is what I mean by the fact that (*to hoti*) is in part and *simpliciter*, in part means 'is the moon eclipsed' or 'is it increasing'? For if the moon is something or is not something, we search for it in such cases; but *simpliciter* means whether there is or not the moon or night (*An. Post. B*, 2, 89b37–90a5).

The first step is to search for a fact or a thing to explain, and if there is one, it means that there is a middle term. The second step is to search for the reason why, i.e. what the middle term is, providing we have cognized the fact that there is a middle term. The affirmative demonstration in the first figure deals with four epistemological questions. Chapter 1 of *Posterior Analytics B* (cf. 89b23–25) defines the four components of demonstrative knowledge as whether a thing is (*ei esti*), the fact that (*hoti*), what a thing is (*ti estin*), and the reason why (*dioti*). To ask whether a thing is aims to search for something *simpliciter* as the thing to be investigated. The answer is to be found in the minor term C (e.g. the moon). The fact that (*to hoti*) is about something predicated of something else, and the answer lies with the deduced conclusion, in which the major term A is predicated of the minor term C (e.g. the eclipse is predicated of the moon). The reason why (*to dioti*) explains the deduced conclusion, namely the fact that the major term A is predicated of

---

not be knowledge of the principles, and since it is not possible for anything to be truer than knowledge, except intellection, then there will be intellection of the principles; and it results from these considerations that demonstration is not a principle of demonstration, nor knowledge a principle of knowledge. If then there is no other true kind apart from knowledge, intellection will be the principle of knowledge" (*An. Post. B*, 19, 100b5–15).

the minor term C. The answer resides in the explanatory middle term B in relation to the minor premise, in which the middle term B is predicated of the minor term C (e.g. the interposition of the earth is predicated of the moon). Finally, what a thing is (*to ti estin*) explains what it is to be this thing (*to ti ên einai*), and the answer is in the immediate (indemonstrable) major premise, in which the major term A is predicated of every middle term B (e.g. the eclipse is predicated of every interposition of the earth). Therefore, whether there is (*ei esti*) a minor term C and the fact that (*hoti*) a major term A is predicated of the minor term C answer the question whether or not there is a middle term B. On the other hand, the reason why (*to dioti*) the major term A is predicated of the minor term C and what the major term A (in relation to C) is (*to ti estin*) answers the question regarding what the middle term B is.

Consequently, Aristotle constructs a theory of deduction in the *Prior* and the *Posterior Analytics* in order to provide a method aiming at the discovery of universal proofs. This proof or explanation is called the middle term, and no demonstration is deemed to be knowledge if it does not make explicit the proof of the proved thing. While the deduced conclusion (asserting the fact that the major term A is predicated of the minor term C) is about the proved thing, the middle term B is about the proof or explanation. In that respect, universality is in the formalization of a proof, and not in the formalization of a language. Aristotle is silent about the syntactical structure of a language, so that any divide between syntax and semantics would be incomprehensible for him. In contrast, he insists on the idea of a formal proof, which takes the form of a universal middle term. This is the formal tool justifying the universality of scientific knowledge.

References

1. Ackrill, J.L.: Aristotle's Categories and De Interpretatione. Oxford (1963)
2. Ax, W.: Zum isolierten rhéma in Aristoteles' De interpretatione 16b19–25. Arch. Gesch. Philos. **61**, 271–279 (1979)
3. Balme, D.M.: Aristotle: Historia Animalium, Volume I, Books I–X. Cambridge (2002)
4. Beziau, J.-Y.: What is 'formal logic'? In: Lee, M.-H. (ed.) Proceedings of the XXII World Congress of Philosophy, vol. 13, pp. 9–22. Korean Philosophical Association, Seoul (2008)
5. Beziau, J.-Y.: Logic is not logic. Abstracta **6**, 73–102 (2010)
6. Bocheński, I.M.: Ancient Formal Logic. North Holland, Amsterdam (1951)
7. Boger, G.: Aristotle's underlying logic. In: Gabbay, D., Woods, J. (eds.) Handbook of the History of Logic, pp. 101–246. Amsterdam (2004)
8. Boole, G.: Laws of Thought. Cambridge (1854/2003)
9. Corcoran, J.: Completeness of an ancient logic. J. Symb. Log. **37**, 696–702 (1972)
10. Corcoran, J.: A mathematical model of Aristotle's syllogistic. Arch. Gesch. Philos. **55**, 191–219 (1973)
11. Corcoran, J.: Aristotle's natural deduction system. In: Corcoran, J. (ed.) Ancient Logic and Its Modern Interpretations, pp. 85–131. Kluwer, Dordrecht (1974)
12. Corcoran, J.: The founding of logic. Anc. Philos. **14**, 9–24 (1994)
13. Corcoran, J.: Aristotle's prior analytics and Boole's laws of thought. Hist. Philos. Logic **24**, 261–288 (2003)
14. Corcoran, J.: Aristotle's demonstrative logic. Hist. Philos. Logic **30**, 1–20 (2009)
15. Crivelli, P.: Aristotle on Truth. Cambridge (2004)
16. Hudry, J.-L.: Aristotle on meaning. Arch. Gesch. Philos. **93**, 253–280 (2011)
17. Hudry, J.-L.: Aristotle on deduction and inferential necessity. Rev. Metaphys. **67**, 3–30 (2013)
18. Hudry, J.-L.: Aristotle on modality and predicative necessity. Int. Philos. Q. **53**, 5–21 (2013)
19. Hudry, J.-L.: Aristotle on non-contradiction: philosophers vs. non-philosophers. J. Anc. Philos. **7**, 51–74 (2013)

20. Keyt, D.: Deductive logic. In: Anagnostopoulos, G. (ed.) A Companion to Aristotle. Wiley-Blackwell, Oxford (2009)
21. Kretzmann, N.: Aristotle on spoken sound significant by convention. In: Corcoran, J. (ed.) Ancient Logic and Its Modern Interpretations, pp. 1–21. Reidel, Dordrecht (1974)
22. Łukasiewicz, J.: Aristotle's Syllogistic from the Standpoint of Modern Formal Logic. Clarendon Press, Oxford (1920/1957)
23. Miller, J.W.: The Structure of Aristotelian Logic. Kegan Paul, Trench, Trubner, London (1938)
24. Minio-Paluello, L.: Aristotle Categoriae et Liber de Interpretatione. Oxford (1963)
25. Patzig, G.: Aristotle's Theory of the Syllogism: A Logico-Philological Study of Book a of the Prior Analytics. Reidel, Dordrecht (1968)
26. Pépin, J.: Sumbola, sèmeia, homoiomata. A propos de De interpretatione 1, 16a3–8 et Politique VIII 5, 1340a6–39. In: Wiesner, J. (ed.) Aristoteles Werk und Wirkung. Paul Moraux gewidmet. Band I: Aristoteles und seine Schule, pp. 22–44. Berlin (1985)
27. Quine, W.: Philosophy of Logic, Cambridge (1970/1986)
28. Ross, W.D.: Aristotle's Prior and Posterior Analytics. Clarendon Press, Oxford (1949)
29. Smiley, T.: What is a syllogism? J. Philos. Log. **2**, 136–154 (1973)
30. Smiley, T.: Aristotle's completeness proof. Anc. Philos. **14**, 25–38 (1994)
31. Smith, R.: Aristotle: Prior Analytics. Hackett, Indianapolis (1989)
32. Wagner, H.: Aristoteles, de interpretatione 3, 16b19–25. In: Palmer, R.B., Hamerton-Kelly, R. (eds.) Philomathes. Studies and Essays in the Humanities in Memory of Philip Merlan, pp. 95–115. The Hague (1971)
33. Weidemann, H.: Aristoteles über das isolierte Aussagewort: De int. 3, 16b19-25. Arch. Gesch. Philos. **64**, 239–256 (1982)
34. Woods, J.: Aristotle's earlier logic. In: Gabbay, D., Woods, J. (eds.) Handbook of the History of Logic, pp. 27–99. Amsterdam (2001)

J.-L. Hudry (✉)
Departamento de Filosofia, Universidade de São Paulo (USP), Av. Prof. Luciano Gualberto, 315, sala 1005, São Paulo, SP, 05508-900, Brazil
e-mail: jl.hudry@gmail.com

Béziau on *And* and *Or*

Lloyd Humberstone

Abstract We offer a commentary on aspects of Béziau's work on combining the logic of conjunction with that of disjunction, concentrating on the most recent incarnation of this work in a paper ('To Distribute or Not to Distribute?') co-authored with M. Coniglio. After some opening remarks we consider the issues of distribution and absorption (Sects. 2 and 3), a discussion naturally leading to the question of whether in the weaker of two senses of 'combined logic'—roughly, combined consequence relation rather than combined proof system based on the standard rules—the combined logic of conjunction and disjunction has a finite (strictly) characteristic matrix (Sect. 4). Section 5 takes a brief look at this consequence relation from the perspective of bivalent valuational semantics.

Keywords Connectives · Consequence relations · Rules · Combining logics

Mathematics Subject Classification (2000) Primary 03B62 · Secondary 03G12

1 Opening Remarks

Jean-Yves Béziau has a knack for noticing areas of logical theory where we find ourselves with intuitions pulling in different directions and are in need of theoretical illumination to organize our ideas and restore stability. His thought-provoking work in [1] draws attention to the way our intuitions about when one logic is weaker than another can come into conflict with each other when brought together with some considerations about translating one logic into another. The thoughts it provoked in this author are assembled in [25] (overflowing into [26]), and in what follows I want to provide a critical commentary of a similar kind—broadly sympathetic though occasionally corrective, and with some indication of connections with pertinent material in the literature—on the culmination (so far, at least), in Béziau and Coniglio [6], of another line of inquiry, initiated by Jean-Yves in [4] and continued in the intermediate [5]. This time the question is what it means to *combine* two logics governing different items of logical vocabulary, and again the discussion represented by all three papers concentrates on tensions between different aspects of what we might expect from such a combination.[1]

[1] The interest in different notions of combining of logics was originally sparked by work of Dov Gabbay (see [18], cited also in footnote 12 below); a reference in [6], namely Carnielli et al. [10], supplies much of that background. In the comments that follow I do not take up aspects of the discussion in [6] cast specifically in the terminology—fib(e)ring, etc.—of this heritage from Gabbay.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_12

Béziau and Coniglio call the pair consisting of a language together with a finitary substitution invariant consequence relation a structural Tarskian consequence system. They denote by $\mathsf{For}_\wedge$ the language with the binary connective $\wedge$ as its sole primitive connective, taking this to be the absolutely free algebra on a fixed (and for definiteness and familiarity, let us specify: denumerable) set of free generators $p_1, \ldots, p_n, \ldots$ or—when convenient—to be the universe of this algebra, whose elements are called formulas of the language and over which we lower case Greek letters are used as variables, with capital Greek letters ranging over sets of formulas. The free generators, [6] calls atomic formulas, but I will call sentence letters (to sidestep controversy over the question of whether, in the variant cases in which there are nullary connectives, these do or do not deserve to be described as atomic formulas[2]). For brevity we write p_1, p_2, p_3 as p, q, r.

Section 2 of [6] remarks that $\langle \mathsf{For}_\wedge, \models_\wedge \rangle$ is a structural Tarskian consequence system, where $\Gamma \models_\wedge \varphi$ is defined to hold just in case for every homomorphism v from $\mathsf{For}_\wedge$ to the two-element algebra with elements 1 and 0 (often written as T and F) and the usual bivalent truth-function associated with $\wedge$, we have: $v(\Gamma) \subseteq \{1\}$ implies $v(\varphi) = 1$.[3] Béziau and Coniglio also mention various syntactic characterizations of this same consequence relation (as we may as well say, since the language is fixed once the consequence relation of a Tarskian consequence system is given), for instance as the least consequence relation $\vdash$ satisfying for all formulas φ, ψ: $\varphi \wedge \psi \vdash \varphi$; $\varphi \wedge \psi \vdash \psi$; and $\varphi, \psi \vdash \varphi \wedge \psi$. (Note that we use semicolons where commas would be normally appear so as to avoid confusion with the commas internal to $\vdash$-statements—in which, of course, such things as "Γ, φ" mean "$\Gamma \cup \{\varphi\}$", etc.) Or again we the $\vdash$ on $\mathsf{For}_\wedge$ defined by: $\Gamma \vdash \varphi$ if and only if $Var(\varphi) \subseteq Var(\Gamma)$, where $Var(\Gamma)$ is the set of sentence letters occurring in at least one formula in Γ.[4] Having treated conjunction in isolation in Sect. 2, our authors do something roughly similar for the case of disjunction in Sect. 3 of [6]; here the language is called $\mathsf{For}_\vee$ (and later, the language with both connectives will be called simply For). Of course, we have the truth-functional treatment associating the Boolean disjunction function with $\vee$, and there is also a "*Var*" formulation with $\Gamma \vdash \varphi$ iff $Var(\psi) \subseteq Var(\varphi)$ for some $\psi \in \Gamma$.[5] But

[2]Similarly, I prefer to say "substitution invariant" rather than "structural" wherever possible, to avoid confusion with the notion of structural-as-opposed-to-operational rules.

[3]Such bivalent valuations v Béziau—in [6], as elsewhere—calls 'bivaluations', a usage not followed here. The notation "$v(\Gamma)$" denotes $\{v(\psi) \mid \psi \in \Gamma\}$. The "$v$" notation is used here to remain close to Béziau's, though my own preference (in [23], [28]) is to reserve this for bivalent valuations and use something like "h"—suggesting homomorphism—for matrix (or more generally algebraic) evaluations. Compositional semantics then proceeds at the "h" level, while considerations of validity are worked in terms of the—typically noncompositional—induced valuations v_h with $v_h(\varphi){=}T$ when $h(\varphi)$ is a designated element. In the case of a two-element matrix with exactly one designated value, the v/h contrast evaporates but in principle it is best registered (which is why T, F, may be preferable—and will be used again in Sect. 5—to 1, 0). It is also better not to call (as [6] does) designated elements 'distinguished' elements, since this confuses matrix-specific terminology and terminology apt for the underlying algebras; for example, a Boolean algebra as frequently conceived has distinguished elements 1 and 0 (nullary fundamental operations, singled out by constants of the equational language) but a Boolean matrix on such an algebra takes only the first of these as a designated value.

[4]And $Var(\varphi)$ is $Var(\{\varphi\})$. Béziau and Coniglio [6, p. 568] add after the "if and only if" the further condition that $\Gamma \neq \varnothing$, but this is redundant.

[5]From which Béziau and Coniglio correctly note the corollary this $\vdash$ has the unusual property—called left-primeness in [28]—that whenever $\Gamma \vdash \varphi$ there is some $\gamma \in \Gamma$ for which $\gamma \vdash \varphi$ (or $\{\gamma\} \vdash \varphi$, in a more pedantic notation).

the correspondence is not exact because the project of dualizing the threefold condition $\varphi \wedge \psi \vdash \varphi$; $\varphi \wedge \psi \vdash \psi$; and $\varphi, \psi \vdash \varphi \wedge \psi$, falters after turning the first two subconditions into $\varphi \vdash \varphi \vee \psi$ and $\psi \vdash \varphi \vee \psi$, since the best we can do for the third subcondition is to turn it into a *conditional* $\vdash$-condition (appearing with this label p. 569 of [6]):

($\vee_3$) If $\Gamma, \varphi \vdash \delta$ and $\Delta, \psi \vdash \delta$, then $\Gamma, \Delta, \varphi \vee \psi \vdash \delta$.

Equivalents of the first two appear there as ($\vee_1$) and ($\vee_2$) (see footnote 6). Béziau and Coniglio rightly say, when it comes to the question of providing a combined logic of conjunction and disjunction: "[t]he problem with the laws for conjunction and disjunction given in Sects. 2 and 3 resides exclusively in law ($\vee_3$)". What cannot be over-emphasized is the absence of a corresponding unconditional form for this case, by contrast with ($\wedge_1$), ($\wedge_2$), ($\wedge_3$), ($\vee_1$) and ($\vee_2$). Its importance lies in the fact that if $\vdash$ and $\vdash^+$ are substitution invariant consequence relations, not necessarily on the same language, but with $\vdash \subseteq \vdash^+$, then any *unconditional* condition (such as the first two for $\vee$ here) satisfied by $\vdash$ is automatically satisfied by $\vdash^+$, while this need not be so for conditional conditions (such as ($\vee_3$)). Indeed, it will not be so for any such conditions as lack unconditional forms equivalent to them as constraints on consequence relations.[6]

Here we use "$\Vdash$" rather than "$\vdash$" to indicate a generalized (or 'multiple conclusion') consequence relation assuming familiarity with the defining conditions for this notion—more or less obviously symmetrical versions of the defining conditions for consequence relations—and recall that it is with such considerations as have just been aired that these are sometimes urged as an improvement over consequence relations, since we can give straightforwardly dual conditions to all three of the proof-theoretically familiar $\wedge$-conditions when we work with these and want to treat $\vee$: we require that $\varphi \Vdash \varphi \vee \psi$; $\psi \Vdash \varphi \vee \psi$; and (the formerly unsayable) $\varphi \vee \psi \Vdash \varphi, \psi$. (The connection with valuations is, of course, that we are ruling out the possibility that a valuation should verify all formulas on the left of the "$\Vdash$" while falsifying all those on the right.) Béziau and Coniglio touch on this issue at the end of Sect. 9 of [6]:

> There is still another solution in the case of the logic of disjunction, which is to consider logics as multiple-conclusion consequence relations. In this case we can state the law ($\vee_3$) without going to the meta-level. But then we can reformulate the result of Sect. 6 and show that the combination of the logics of disjunction and conjunction is necessarily distributive.[7]

The cross-references in this passage will be clarified in the following section. For the moment I would like to object to the talk of a meta-level with its suggestion of a metalinguistic or meta-logical transition. The claim that for a particular consequence relation $\vdash$ or generalized consequence relation $\Vdash$, if $\varphi \vdash \delta$ and $\psi \vdash \delta$ then $\varphi \vee \psi \vdash \delta$ (or again with "$\Vdash$" at each occurrence of "$\vdash$") is no more metalinguistic than the claim that $\varphi \vee \psi \vdash \delta$ (or $\varphi \vee \psi \Vdash \delta$), for instance: both are equally metalinguistic—it's just that one is conditional and the other unconditional, as in the case that we have this generally, for arbitrary

[6]The already mentioned ($\vee_1$) and ($\vee_2$) are given in conditional form in [6], namely as "If $\Gamma \vdash \varphi$ then $\Gamma \vdash \varphi \vee \psi$" and "If $\Gamma \vdash \psi$ then $\Gamma \vdash \varphi \vee \psi$", but evidently these are respectively equivalent, as conditions on a consequence relation $\vdash$, to the unconditional forms "$\varphi \vdash \varphi \vee \psi$" and "$\psi \vdash \varphi \vee \psi$".

[7]This passage appears at p. 577 of [6]. The opening phrase "still another solution" makes sense in the light of the inset quotation from the paper given in Sect. 2, beginning "One solution…"; see further footnote 16 below (and the text to which it is appended). In giving the quotation here, I have changed "going at" to "going to" in the hope of capturing the authors' intentions. Some further typographical corrections to [6] appear in footnote 23 below.

φ, ψ, δ (in the language of $\vdash$), a condition on $\vdash$ called $(\vee_{3w})$ in [6] because it weakens the condition $(\vee_{3w})$ to the special case in which $\Gamma = \Delta = \varnothing$:

$(\vee_{3w})$ If $\varphi \vdash \delta$ and $\psi \vdash \delta$, then $\varphi \vee \psi \vdash \delta$.

We can, if we wish, discuss object-level analogues of these claims and conditions and consider sequents and sequent-to-sequent rules, for example the rule whose admissibility for $\vdash$ constitutes $\vdash$'s satisfying the condition $(\vee_{3w})$; here we use $\succ$ as a sequent-separator:

$$\frac{\varphi \succ \delta \quad \psi \succ \delta}{\varphi \vee \psi \succ \delta}$$

in which again the (proper) rules are not 'meta-level' versions of their zero-premiss cousins ($\varphi \succ \varphi \vee \psi$, for instance, "zero-premiss" alluding to *sequent* premisses, not formulas on the left of the "$\succ$" in a given sequent) or particular applications thereof (specific sequents such as $p \succ p \vee q$), though the *claim* that a certain rule is derivable or admissible, no less than the claim that a certain sequent is provable, will of course be a metalinguistic or meta-logical claim.[8]

The question of whether the $\vee$ rule above—or the version with additional side formulas on the left of both premiss-sequents (mentioned in footnote 8)—is interderivable with a one-premiss rule in the presence of the all zero-premiss SET-FMLA rules holding for classical $\vee$ ($\varphi \succ \varphi \vee \psi$; $\varphi \vee \varphi \succ \varphi$; etc.) is answered affirmatively by Rautenberg: for this purpose one can take the rule with premiss $\psi \succ \delta$ and conclusion $\varphi \vee \psi \succ \varphi \vee \delta$. (Conjecture: no such one-premiss sequent-to-sequent rule whose schematic formulation uses a single occurrence of $\vee$ can be found.) Actually, Rautenberg [32] treats the corresponding condition on consequence relations—or rather the version with side-formulas, which he labels $(\mathrm{m}^{\vee})$:[9]

$$\text{If } \Gamma, \psi \vdash \delta \quad \text{then } \Gamma, \varphi \vee \psi \vdash \varphi \vee \delta,$$

and his interest in it is in providing a purely unconditional set of conditions on $\vdash$ (essentially: zero-premiss sequent-to-sequent rules) such that $\vdash_{\vee}$ is the least consequence

---

[8]The contrast between conditional and unconditional $\vdash$/$\Vdash$ conditions I first encountered in lectures in the 1970s from Dana Scott, but it also emphasized in Rautenberg [32], and specifically in the case of consequence relations and $(\vee_3)$ type conditions on them; see Example 1.1 at the end of this section. The distinction between sequents, as formal analogues of arguments (in the premisses and conclusion sense—or set-of-conclusions $\Vdash$), and claims about sequents (or arguments) is stressed in [28]. Restoring the side-formulas Γ, Δ in the rule inset above here gives Gentzen's LJ sequent calculus rule for $\vee$-on-the-Left; the rule without side-formulas has a corresponding natural deduction rule called restricted $\vee$-Elimination or $(\vee\mathrm{E})_{res}$ at p. 299 of [28] (q.v. for references to the literature, which include p. 205f. of Dummett [12]), with premisses $\Gamma \succ \varphi \vee \psi$ and $\varphi \succ \delta$ and $\psi \succ \delta$ and conclusion $\Gamma \succ \delta$. (Note that this is the sequent-to-sequent description of a natural deduction rule, not a sequent calculus rule *sensu stricto*, in that is neither a structural rule—one involving no specific logical vocabulary—not a rule inserting the principal connective or quantifier in a formula on the left or on the right of the $\succ$.) Can we not, in that case, do multiple-conclusion logic with in a sequent calculus for quantum logic, still avoiding the unwanted distribution laws, and accepting $\varphi \vee \psi \succ \varphi, \psi$? Yes we can—but the Cut rule must be seriously restricted so that the associated relation holding between Γ and Δ when $\Gamma \succ \Delta$ is provable will no longer be a generalized consequence relation: again see the references in [28], this time on p. 301.

[9]And calls *monotonicity*; in the terminology of [28] this says that $\vee$ is monotone *with side formulas* in its second position according to $\vdash$, reserving the terminology without the added "with side formulas" for the corresponding condition with "Γ" and "Δ" absent. Of course, we pick up monotonicity in the first position from this since we have the unconditional principle: $\varphi \vee \psi \vdash \psi \vee \varphi$.

relation satisfying them. These appear as (1)–(4) on p. 323 of [32].[10] Rautenberg shows that ($m^{\vee}$) is satisfied by the consequence relation so described, and hence that, in Béziau and Coniglio's terms, this consequence relation satisfies their ($\vee_3$). To conclude these opening remarks, we repeat an example from Rautenberg [32, p. 326], illustrating the possibility of a consequence relation $\vdash \supseteq \vdash_{\vee}$ and not satisfying ($\vee_3$):

Example 1.1 Take the consequence relation on a language with (binary) $\vee$ and (nullary $\bot$) determined by the matrix whose $\vee$-table is that of Fig. 1 below and with $\bot$ as a constant assigned the value 2 (i.e. on all matrix evaluations it is to receive this value). This consequence relation, $\vdash$, extends the consequence relation $\vdash_{\vee}$ of the pure disjunction fragment of classical logic but does not satisfy ($\vee_3$) because $p \vee \bot \nvdash p$ (assign the value 3 to p), even though $p \vdash p$ and $\bot \vdash p$.

2 Distribution

As has already been foreshadowed by some of the excerpts from [6] given in the previous section, the main issue under discussion there is how $\wedge$ and $\vee$ interact when we combine the logics given for them separately. Béziau and Coniglio consider two examples of interaction, the main one—as indicated by the title of the paper—being the question of the distribution laws, and secondarily, the absorption laws. The latter will occupy us in the following section. As above, we are mainly concerned with 'single-conclusion logic' and reproducing the distribution laws under discussion as they appear in [6, p. 571], the question concerns whether the envisaged combined logic should satisfy (D_1) and (D_2), for all φ, ψ, δ, where "$\alpha \dashv\vdash \beta$" means "$\alpha \vdash \beta$ and $\beta \vdash \alpha$":

(D_1) $\varphi \wedge (\psi \vee \delta) \dashv\vdash (\varphi \wedge \psi) \vee (\varphi \wedge \delta)$,
(D_2) $\varphi \vee (\psi \wedge \delta) \dashv\vdash (\varphi \vee \psi) \wedge (\varphi \vee \delta)$.

And, of course, the answer depends on what we take to be "the combined logic" of the pure conjunction and pure disjunction logics. Béziau and Coniglio [6] recall some pertinent facts: if $\vdash$, in the language with connectives $\wedge$ and $\vee$, satisfies ($\wedge_i$) for $i = 1, 2, 3$ and ($\vee_j$) for $j = 1, 2, 3$, then $\vdash$ satisfies (D_1) and (D_2), whereas if $\vdash$ satisfies ($\wedge_i$) for $i = 1, 2, 3$, and ($\vee_j$) for $j = 1, 2$, and also ($\vee_{3w}$) from Sect. 1, then $\vdash$ satisfies the $\dashv$-half of (D_1) and the $\vdash$-half of (D_2), satisfying the converses of these conditions only if the full unweakened condition ($\vee_3$) is satisfied. There is very little discussion of quantum logic in [6],[11] so let it be noted here that this is the reason natural deduction systems for quantum logic use the restricted form of $\vee$-elimination mentioned in footnote 8.

If we agree that combination of two logics should be the smallest logic of which both are sublogics,[12] the question of whether the combined logic of (classical—though we

[10] For the reader's convenience, we list them here: (i) $\varphi \vdash \varphi \vee \psi$, (ii) $\varphi \vee \varphi \vdash \varphi$, (iii) $\varphi \vee \psi \vdash \psi \vee \varphi$, (iv) $\varphi \vee (\psi \vee \delta) \vdash (\varphi \vee \psi) \vee \delta$. Finitely many such unconditional $\vdash$-conditions—including some redundant ones—are also given for the $\{\wedge, \vee\}$-fragment of the classical (or intuitionistic) consequence relation in Dyrda and Prucnal [13], as well as by Rautenberg himself ([32] last four lines of p. 324).

[11] The only reference to it there is in the brief remark on p. 571 to the effect that "non-distributivity is a real phenomenon, especially known in quantum logic and lattice theory."

[12] This seems to be the cleanest policy, rather than to build into the definition further, that the combined logic is extends each of the logics combined conservatively, an idea Béziau [4] takes from Gabbay [18]

could equally well say, for example, intuitionistic) conjunction and disjunction becomes the question of what a logic is, and Béziau and Coniglio consider three answers to this question which we might describe briefly and no doubt somewhat misleadingly as semantic, consequence-theoretic and proof-theoretic.[13] The first of these identifies a logic with a favored account of validity, which is typically defined in terms of truth-preservation over a range of entities—valuations or more structured truth-conferring entities (models)—which are subjected to various constraints (for example, that a conjunction is only to count as true relative to such an entity when each conjunct is). In this case the natural notion of combination of logics is obtained by combining the constraints for the different connectives in the logics to be combined. While this has led to interesting developments in modal logic (see Sects. 1.2.1 and 1.2.2 in [10]) in the current more fundamental setting it rapidly leads to the delicate area of sensitivity to the metalogic, and we prefer to steer clear of it after doing no more than indicating the area in question.[14] Here imposing the joint conditions on bivalent valuations associated with $\models_\wedge$ and $\models_\vee$—namely that such valuations should verify conjunctions iff they verify both conjuncts and falsify disjunctions iff they falsify both disjuncts—Béziau and Coniglio [6, p. 571] write that "it is straightforward to prove that the obtained logic (...) satisfies the following laws", where the laws in question are (D_1) and (D_2) above. It is straightforward indeed if one makes use of the distributive laws in the metalanguage in which the proof is conducted. One might try to sidestep this dependence on the meta-logic, which, if such a proof is offered in the dialectical situation of trying to persuade someone sceptical about distribution, presents itself as a circularity, by doing a verification by truth-tables. However, as Dummett has pointed out on several occasions (usually crediting the observation to Hilary Putnam), this simply conceals the circularity: the truth-table method itself passes from what bivalence gives us: p is either true or false, as are q and r, to the metalinguistic disjunction whose disjuncts correspond to the eight lines of the truth-table for a consequence-statement involving formulas in at most p, q, r: either p and q and r are true, or p and q are true while r is false or ...—and this very transition represents an appeal to distributivity.[15] Béziau and Coniglio [6] intro-

but has dropped by the time we reach [6]. The latter policy would mean that for some pairs of logics there is no combined logic. It is better to use the simple definition, which leaves it an open question in any given case whether the combined logic is conservative over each of the originals. Nonconservativity is not an issue for any of the candidates for combining the logic of $\wedge$ and $\vee$ under consideration in what follows.

[13] These are not alternative answers to a question of fact, but different options as to how to individuate logics for the purposes at hand. Several formulations in [6] sound a note of quasi-religious zealotry, such as in part of the heading of Sect. 12 (p. 581), which includes the words "What is the right way to combine logics?", and in—what are inappropriately called—Conjectures 1 and 2 on p. 571, which speak of having applied the right or the wrong way to combine LC and LD. The interest should be in exploring and contrasting different modes of combination of logics rather than trying to outlaw all but one of them some as some illegitimate. These rather fevered formulations are best regarded as rhetorical devices (to attract attention to the issues), though they may have misled Béziau into thinking that the existence of considerations favoring one rather than another and other considerations tending in the opposite direction somehow amount to a paradox—see the title of [4]—rather than a plurality of legitimate notions of combination.

[14] "Fundamental" is perhaps not the right word: the issue concerns the degree to which the semantic treatment is homophonic. The remarks just made about modal logic concerned the Kripke semantics; compare [21].

[15] See, for example, the top paragraph of p. 55 of Dummett [12]; the next paragraph on the page discusses this issue of sensitivity to the meta-logic more generally.

duce syntactically characterized consequence relations $\vdash_\wedge$ and $\vdash_\vee$ and shows quickly that they coincide respectively with the semantically characterized $\models_\wedge$ and $\models_\vee$; from now on, we use the "$\vdash$" rather than the "$\models$" notation in connection with these two.

The second of the three options alluded to would have us take logics to be (substitution invariant) consequence relations.[16] In that case the combined logic is the least such consequence relation—in fact, in the formulation of [6], the least structural Tarskian consequence system—in the language with connectives $\wedge$ and $\vee$, extending the pure conjunction consequence relation and the pure disjunction consequence relation. The two single-connective Tarskian consequence systems are in [6] LC and LD, respectively, with associated consequence relations $\vdash_\wedge$ and $\vdash_\vee$; their combined Tarskian consequence system is referred to as CD, which on the current conception of combination of logics means that the associated consequence relation, $\vdash_{\mathsf{CD}}$, is the join (or least upper bound, or supremum) of $\vdash_\wedge$ and $\vdash_\vee$ in the lattice of all substitution invariant consequence relations on the combined $\wedge/\vee$ language (with partial order $\subseteq$);[17] this involves a slight reconstrual of what $\vdash_\wedge$ and $\vdash_\vee$ amount to, recasting them in a common language—see Sect. 5 below for further details. Béziau and Coniglio show in their Theorem 8.1 [6, p. 575] that $\vdash_{\mathsf{CD}}$ does not satisfy the distributive laws (D_1) and (D_2), so on this way of understanding logics and their combination, the combined logic of conjunction and disjunction is not distributive. We return to the proof of Theorem 8.1 below, after looking at the alternative way of taking this terminology.

The second option—or perhaps better, second strategy, since its introduction concludes with a question as to how exactly to implement the idea is mooted in the following passage [6, p. 576]:

> One solution is to consider that a logical structure is not completely given by the consequence relation but must be further specified by some meta-properties, but then how to define the set of relevant meta-properties?

The obvious property bearing on the distribution issue is of course $(\vee_3)$, and setting aside some proposed candidates (in terms of valuational semantics and also translations between logics: p. 576f. of [6]), let us immediately pass to what seems the more promising suggestion: individuate logics by means of their rules. Of course, the idea of attending to formula-to-formula rules is already embodied in the widespread acknowledgment that logics should be understood as (substitution invariant) consequence relations rather than (certain) set of formulas, but this does not go far enough, since we need also to capture sequent-to-sequent rules. Béziau and Coniglio [6, p. 577] cite Béziau [2] in this connection but it seems appropriate here to recall the birth of what came to be called inferentialism, taken as the view that full understanding of an item of logical vocabulary is a matter of reasoning in accordance with appropriate introduction and elimination rules, in Gentzen [19]. Gentzen himself stressed the meaning-conferring role of the introduction rules (p. 80 of the English translation in our bibliography), but this is not a compulsory

---

16 In an important variation, one would take here *generalized* consequence relations; this is the significance of the opening sentence of our inset quotation from [6] in Sect. 1: "There is still another solution…". Of course, the philosophical legitimacy of this move is perennially contested. See Steinberger [35], for example.

17 Some of these labels will have distracting associations that need to be ignored: the current LC, LD and CD have nothing to do with Dummett's intermediate logic LC, Curry's supra-minimal logic LD, or the intermediate predicate logic of constant domains, respectively.

component of the general view. Even with this component in force, consider the natural deduction introduction rule for implication (which coincides with the LJ sequent calculus rule for inserting the connective in question on the right, and is popularly known as 'conditional proof'). This is an assumption-discharging rule and so represents one-premiss sequent-to-sequent rule which cannot be replaced by an equivalent zero-premiss rule (i.e. sequent-schema), and so corresponds to a condition on consequence relations not automatically inherited by their extensions (for discussion of which, see [27]). From this perspective, it is natural to take the logic of implication to be given by this rule and the elimination rule, rather than to identify it simply with the least consequence relation (which will automatically be substitution invariant and finitary) on the language with this as its sole connective. This does not capture the idea of reasoning in accordance with the given rules come what may (by way of additional connectives and whatever rules they answer to). In the case of disjunction, it is, of course, the elimination rule that plays a similar role, which is tantamount to acknowledging as intelligible only such additions to the language as continue to respect the condition ($\vee_3$), which, is not itself guaranteed to survive such extensions. And indeed, this broadly inferentialist spirit is voiced by the Béziau and Coniglio when they raise the question [6, p. 576] of "whether a logic in which ($\vee_3$) does not hold can properly be called a logic of disjunction."

An interesting point made in [6, p. 579] is that in the description of LD (or, as I would prefer to put it, of $\vdash_\vee$), it makes no difference whether we describe the consequence relation involved using ($\vee_3$) or ($\vee_{3w}$) alongside ($\vee_1$) and ($\vee_2$): this follows from the fact that defining the consequence relation using ($\vee_3$) makes it left-prime in the sense of footnote 5 above. Thinking of codifying the consequence relation concerned using a proof system—a collection of rules, that is—we get the same consequence relation whether we include only the special ($\vee_{3w}$)-like rule inset in the final paragraph of Sect. 1, or instead the general version of this rule which has "Γ" on the left in the first premiss, "Δ" on the left in the second, and both on the left in the conclusion. About this, our authors remark that "[t]his example clearly shows that different presentations of the same logic produce different logics when the given logic is combined with other[s]," but while correct if *logic* is taken to mean *consequence relation* matters are less clear if the finer-grained individuation policy described above is followed. One reasonable proposal might be two regard two rule-based presentations (or proof systems) as presentations of the same logic if the same sequent-to-sequent rules are derivable in them (meaning by derivable: primitive or derived). By this criterion, the presentations yield different logics, since the general ($\vee_3$) style rule is not derivable—though it is, of course, *admissible* (or 'permissible' to use the language of Béziau [2])—in the proof system with rules corresponding to ($\vee_1$), ($\vee_2$) and ($\vee_{3w}$), alongside the usual structural rules.

Moving away from these somewhat philosophical considerations about the individuation of logics, I would like to raise three technical matters before passing from the discussion of distribution to that of absorption. The first concerns the matter of the unique characterization of connectives by rules governing them, on which for further details see §4.3 of [28]. As is well known, the usual rules (involving sequents with a set of formulas to the left and a single formula on the right) governing $\wedge$ in a natural deduction system or in a sequent calculus uniquely characterize that connective in the sense that if a new binary connective $\wedge'$, say, is added governed by analogous rules, then one can prove in the resulting system—the combined proof system, to use the above terminology—then the connectives yield equivalent compounds in the sense that for all φ, ψ, the sequents

$\varphi \succ \psi$ and $\psi \succ \varphi$ are provable. And the same goes for $\vee$. But there is a difference between the two cases. In the case of $\wedge$, the zero-premiss rules governing this connective (in, say, classical logic) uniquely characterize it, whereas in the case of $\vee$, this is not so, as Sect. 4.35 of [28] explains.[18] This is because the rule which reflects the constraint on valuations that a disjunction can only be receive the value 1 if at least one of its disjuncts does—$\vee$-elimination, ($\vee$ Left), corresponds to the condition ($\vee_3$) on the associated consequence relation, and, as Example 1.1, this condition is not in general inherited on passage to a more extensive consequence relation. This present case is a further illustration of the same phenomenon: $\vdash_\vee$ satisfies the condition but the join of this with $\vdash_{\vee'}$ in the lattice of all substitution invariant consequence relations[19] does not. (In fact, the 'quantum-logical' conditions ($\vee_1$), ($\vee_2$), and ($\vee_{3w}$)—or more precisely the corresponding rules already characterize $\vee$ uniquely—a consideration used in Sect. 4.32 of [28] in an *ad hominem* reply to the intuitionist complaining that the classical logician supplies $\neg$ with rules which are *stronger than needed* for unique characterization, the reply being that the intuitionistic logician, invoking ($\vee_1$), ($\vee_2$), and ($\vee_3$) is employing rules governing $\vee$ which are similarly stronger than needed for unique characterization.) Of course, if we allow multiple conclusion sequents (in the terminology of [28], pass from the logical framework SET–FMLA to the framework SET–SET) then this difference between $\wedge$ and $\vee$ disappears, because we have the zero-premiss rule $\varphi \vee \psi \succ \varphi, \psi$ to play with (and likewise with $\vee'$, $\vee$'s rule-reduplicated companion).

The second technical point is similarly related to this. The distribution principle (D_1) and (D_2) are entirely symmetrical in $\wedge$ and $\vee$, and yet when it comes to doing sentential logic, by striking contrast with case of working with equations in lattice theory, it seems that we are tracing the failure of the their distinctively distributive aspects (the $\vdash$-direction of (D_1) and the $\dashv$-direction, the other directions corresponding to inequalities holding in arbitrary lattices) specifically to an issue about $\vee$—whether to impose the condition ($\vee_3$), which yields these conditions, or only the condition ($\vee_{3w}$), which does not. Why all the fuss about conditions involving $\vee$, and nothing about $\wedge$, in the logical case? Predictably the answer is that "when it comes to doing sentential logic," it all depends on *how* we decide to do this. The whole asymmetry was created by the decision to look at consequence relations (which can be thought of as certain sets of SET–FMLA sequents) with their potentially multiple (or empty) left-hand sides but their doggedly singleton right-hand sides. One sees this from the passage to generalized consequence relations (sets of SET–SET sequents), as is evident from footnote 8. The point can be emphasized by a consideration of a notion dual to that of a consequence relation: think of a *dual-consequence relation* as a relation between formulas and sets of formulas answering to left-right reversed versions of the defining conditions on consequence relations; using "$\vdash$*" for such a relation—which is to be taken as an unstructured symbol, not thought of as the result of

[18] Unique characterization by zero-premiss rules of the present kind is called *functional dependence* in Smiley [34], as is mentioned in the notes and references to §4.3 of [28], where it is also mentioned that some think of unique characterization as a matter of implicit definition—though there is a potential risk of confusion here (occasioned by the interpretation of schematic letters used in the formulation of rules, which need to be instantiated by formulas of the combined language with both the given connective # and its reduplicated version #′).

[19] In [28], this join is given the ad hoc designation $\vdash_{DD}$—the subscript suggesting "double disjunction".

applying an operation ("*") to some consequence relation ($\vdash$).[20] For example, the condition called (EXT) in [6, p. 567] becomes the condition: $\varphi \vdash^* \Gamma$ whenever $\varphi \in \Gamma$. As with consequence relations, the semantic interpretation of dual-consequence relations is just the obvious restriction of the interpretation for generalized consequence relations. Now the full story of (classical) disjunction can be told in unconditional terms—we require that $\varphi \vdash^* \varphi \vee \psi$; $\psi \vdash^* \psi \vee \psi$; and $\varphi \vee \psi \vdash^* \varphi, \psi$. This time two thirds of the corresponding story for conjunction can be told unconditionally: $\varphi \wedge \psi \vdash^* \varphi$ and $\varphi \wedge \psi \vdash^* \psi$, but the final chapter has to go conditional: *if* $\delta \vdash^* \varphi, \Gamma$ *and* $\delta \vdash^* \psi, \Delta$ *then* $\delta \vdash^* \varphi \wedge \psi, \Gamma, \Delta$. Now (D_1) and (D_2)—with $\vdash$ rewritten as $\vdash^*$—follow, while, as above with $(\vee_3)$ and $(\vee_{3w})$, there is a weakening available for this last condition, setting $\Gamma = \Delta = \varnothing$ and with only the weakened version of the condition, the 'distinctively distributive' parts of (D_1) and (D_2) do not follow. While there may not be much intuitive appeal to the idea of dual-consequence relations (or the associated logical framework FMLA–SET, as it would be called in the nomenclature of [28]) and it is certainly not hard to see why it has not had the coverage that consequence relations have,[21] at least we can use it to appreciate how there is nothing inherent in the subject-matter which would locate the repercussions for distributivity in the behavior of $\vee$ rather than $\wedge$. (Notice, incidentally, that the pure logic of conjunction, taken as the pure $\wedge$-fragment of the classical dual-consequence relation is what, dualizing the definition in footnote 5, would be called *right*-prime.)

This brings us, finally, to the proof of Theorem 8.1 in Béziau and Coniglio [6], according to which the least substitution-invariant consequence relation, $\vdash_{\mathsf{CD}}$, in the language with $\wedge$ and $\vee$ which extends both $\vdash_\wedge$ and $\vdash_\vee$ does not satisfy the distribution laws (D_1) and (D_2). The main contrast is with the consequence relation the authors call $\vdash_{\wedge\vee}$, defined as the least consequence relation (on this same language) satisfying $(\wedge_1)$–$(\wedge_3)$ and $(\vee_1)$–$(\vee_3)$, but also with the consequence relation $\vdash_{\mathsf{Lat}}$ about to be introduced. The former is the conjunction–disjunction fragment of intuitionistic and classical logic; a syntactical presentation of the latter requires a weakening of $(\vee_3)$ to $(\vee_{3w})$. (In fact, the authors call the consequence relation $\vdash^{\wedge\vee}$, but for uniformity, here we keep all identifying tags on turnstiles in subscript position. Rather confusingly, Béziau and Coniglio pair this consequence relation with the label "LCD"—which goes back to Béziau [4]—when specifying the 'Tarski consequence system' involved, but let us avoid this as being far too close to "CD" which it is supposed to contrast with.)

The proof of Theorem 8.1 in [6] goes via a consideration of the consequence relation they call $\vdash_{\mathsf{Lat}}$ characterized semantically in the following terms: $\Gamma \vdash_{\mathsf{Lat}} \varphi$ if and only if exists a nonempty finite set $\{\gamma_1, \ldots, \gamma_n\} \subseteq \Gamma$ such that $v(\gamma_1) \wedge \cdots \wedge v(\gamma_n) \leq v(\varphi)$ holds in L, for every lattice L and every homomorphism $v : \mathsf{For} \longrightarrow L$. (That is a direct quotation from p. 575; For is the language—considered as an algebra—with connectives $\wedge$ and $\vee$, though the "$\wedge$" in the definition picks out the meet operation of L, whose partial ordering is the "$\leq$" of the definition. Béziau and Coniglio actually write "$\vdash_{\mathsf{Lat}}$", as did the

[20] Thus the present dual-consequence relations are not the dual consequence relations of Wójcicki, who actually prefers the 'consequence *operations*' conceptualization of the material—see [20] and references, which are dualized versions of particular consequence relations; likewise with the 'dual consequence relations' terminology in Gabbay [17], except that by 'consequence relation' Gabbay means *generalized* consequence relation.

[21] See Sect. 4, 'Apparent Left/Right Asymmetries Historically Explained', of [26].

present author in Sect. 2.14 of [28] for a closely related logic,[22] but here we are using the same font for all turnstile subscripts.) Since the definition says "for every lattice L" rather than "for every distributive lattice L", we see that $\vdash_{\mathsf{Lat}}$ does not satisfy the distribution conditions (D_1) and (D_2), but evidently $\vdash_{\mathsf{Lat}} \supseteq \vdash_{\wedge}, \vdash_{\vee}$, so $\vdash_{\mathsf{Lat}} \supseteq \vdash_{\mathsf{CD}}$, and therefore $\vdash_{\mathsf{CD}}$ does not satisfy (D_1), (D_2).

The proof is perfectly fine, and Béziau and Coniglio also supply a more syntactical characterization of the consequence relation $\vdash_{\mathsf{Lat}}$ figuring so prominently in it: it is the least consequence relation satisfying $(\wedge)_1$, $(\wedge)_2$, $(\wedge)_3$, $(\vee)_1$, $(\vee_2)$, and $(\vee_{3w})$. But the exact relationship of $\vdash_{\mathsf{CD}}$ to $\vdash_{\mathsf{Lat}}$ is left open: in particular, we are left with the question of whether $\vdash_{\mathsf{CD}} = \vdash_{\mathsf{Lat}}$ or $\vdash_{\mathsf{CD}} \subsetneq \vdash_{\mathsf{Lat}}$. What Theorem 8.1 tells about $\vdash_{\mathsf{CD}}$ is that it does not satisfy (D_1) and (D_2), and this leaves us in ignorance as to whether this is because is only distinctively distributive halves of these conditions that fail—in which case we should have $\vdash_{\mathsf{CD}} = \vdash_{\mathsf{Lat}}$, or whether even the other distributively innocent halves that fail, corresponding to inequalities holding in all lattices and meaning that $\vdash_{\mathsf{CD}}$ is properly weaker than $\vdash_{\mathsf{Lat}}$. Prima facie, one expects the latter, because just like $(\vee_3)$ itself, $(\vee_{3w})$ is a *conditional* condition on consequence relations and therefore prone to failing on passage from a consequence relation satisfying it (here $\vdash_{\vee}$) to an extension—even a substitution-invariant extension—thereof, such as $\vdash_{\mathsf{CD}}$, as we have seen with a couple of illustrations already (Rautenberg's Example 1.1 at the end of Sect. 1, and the 'double disjunction' example earlier in the present section). But this does not tell us that the general possibility of failure to be preserved is realized in the present case. We shall see that it is. (The discussion presupposes some familiarity with modal logic.)

Consider the following translation, τ from the language For, with its connectives $\wedge$ and $\vee$, to a language having in addition the 1-ary connective $\Box$, which we shall be interpreting in the usual manner in Kripke models $\langle W, R, V\rangle$ (where $\wedge$ and $\vee$ will also have their customary interpretation, and below, we use $\to$ and $\leftrightarrow$ for material implication and material equivalence, respectively):

- $\tau(p_i) = p_i$,
- $\tau(\varphi \vee \psi) = \tau(\varphi) \vee \tau(\psi)$,
- $\tau(\varphi \wedge \psi) = \Box(\tau(\varphi) \wedge \tau(\psi))$.

Now consider what in the nomenclature of Chellas [11] is the modal logic $\mathsf{K4!}$, i.e. the smallest normal modal logic—considered as a set of formulas—containing for every formula φ the formula $\Box\varphi \leftrightarrow \Box\Box\varphi$, a logic which is easily seen (and well known) to be sound and complete w.r.t. the class of models $\langle W, R, V\rangle$ in which $R = R^2$. We write "$\vdash_{\mathsf{K4!}} \varphi$" to mean that $\varphi \in \mathsf{K4!}$. We use this modal translation τ to define a consequence relation on the language For we will call $\vdash_{\mathsf{mod}}$ (the subscript suggesting "modal").

Definition 2.1 $\Gamma \vdash_{\mathsf{mod}} \psi$ if and only be for some $\varphi_1, \ldots, \varphi_n \in \Gamma$ we have:

$$\vdash_{\mathsf{K4!}} \Box\big(\tau(\varphi_1) \wedge \cdots \wedge \tau(\varphi_n)\big) \to \Box\tau(\psi).$$

[22] In fact, for a logic in Fmla–Fmla, in which what corresponds to the consequence relations, generalized consequence relations and dual-consequence relations of Set–Fmla, Set–Set, and Fmla–Set, respectively, are pre-orders on the set of formulas (called monomonoconsequence relations in [6]). The corresponding Fmla–Fmla fragment of $\vdash$ is treated in [28] under the name *DLat* rather than *Lat* (initial "D" for Distributive).

Proposition 2.2 $\vdash_{\mathsf{mod}}$ *is a substitution invariant consequence relation for which we have* $\vdash_{\mathsf{CD}} \subseteq \vdash_{\mathsf{mod}}$.

Proof The first claim, that $\vdash_{\mathsf{mod}}$ substitution invariant consequence relation, is left to the reader to check. Since $\vdash_{\mathsf{CD}}$ is the least substitution-invariant consequence relation in the language For extending $\vdash_\wedge$ and $\vdash_\vee$, we must show that if $\Gamma \vdash_\vee \psi$ or $\Gamma \vdash_\wedge \psi$, then $\Gamma \vdash_{\mathsf{mod}} \psi$. For the case of $\vdash_\vee$, this is simple because this is a consequence relation on the language $\mathsf{For}_\vee$, restricted to which language, τ is the identity map, so if $\varphi_1, \ldots, \varphi_n \vdash_\vee \psi$, then $(\varphi_1 \wedge \cdots \wedge \varphi_n) \to \psi$, alias $(\tau(\varphi_1) \wedge \cdots \wedge \tau(\varphi_n)) \to \tau(\psi)$, is a truth-functional tautology and thus, $\mathsf{K4!}$ being a normal modal logic, we have $\vdash_{\mathsf{K4!}} \Box(\tau(\varphi_1) \wedge \cdots \wedge \tau(\varphi_n)) \to \Box\tau(\psi)$, which is to say: $\varphi_1, \ldots, \varphi_n \vdash_{\mathsf{mod}} \psi$. Next the case of $\vdash_\vee$. Since this is the least consequence relation according to which for any φ and ψ, we have each of φ, ψ a consequence of $\varphi \wedge \psi$, and $\varphi \wedge \psi$ a consequence of $\{\varphi, \psi\}$, we check these properties in turn. For the first, we must have $\vdash_{\mathsf{K4!}} \Box\tau(\varphi \wedge \psi) \to \Box\tau(\varphi)$, which is to say: $\vdash_{\mathsf{K4!}} \Box\Box(\tau(\varphi) \wedge \tau(\psi)) \to \Box\tau(\varphi)$. But this is clearly the case since for any formulas α, β, we have $\vdash_{\mathsf{K4!}} \Box\Box(\alpha \wedge \beta) \to \Box\alpha$; thus $\varphi \wedge \psi \vdash_{\mathsf{mod}} \varphi$. The second case, with ψ a consequence of $\varphi \wedge \psi$ works similarly, bringing us to the final case, for which we must show that $\varphi, \psi \vdash_{\mathsf{mod}} \varphi \wedge \psi$, which amounts to: $\vdash_{\mathsf{K4!}} \Box(\tau(\varphi) \wedge \tau(\varphi)) \to \Box(\tau(\varphi \wedge \psi))$, i.e. $\vdash_{\mathsf{K4!}} \Box(\tau(\varphi) \wedge \tau(\varphi)) \to \Box\Box(\tau(\varphi) \wedge \tau(\psi))$. But we always have $\vdash_{\mathsf{K4!}} \Box\alpha \to \Box\Box\alpha$. $\Box$

Let us observe that the consequence relation $\vdash_{\mathsf{mod}}$ does not satisfy the condition $(\vee_{3w})$. First notice that $p \vdash_{\mathsf{mod}} p \wedge p$, since $\vdash_{\mathsf{K4!}} \Box p \to \Box\Box(p \wedge p)$, or, more simply put, $\vdash_{\mathsf{K4!}} \Box p \to \Box\Box p$. Similarly, $q \vdash_{\mathsf{mod}} q \wedge q$. Weakening both claims, we get that $p \vdash_{\mathsf{mod}} (p \wedge p) \vee (q \wedge q)$ and $q \vdash_{\mathsf{mod}} (p \wedge p) \vee (q \wedge q)$. Thus $(\vee_{3w})$ would require that $p \vee q \vdash_{\mathsf{mod}} (p \wedge p) \vee (q \wedge q)$, which is to say:

$$\vdash_{\mathsf{K4!}} \Box(p \vee q) \to \Box(\Box p \vee \Box q).$$

But this is not the case, as is most easily seen by substituting the negation of p for q, giving an $\mathsf{K4!}$-provable antecedent with a more obviously unprovable consequent. (Alternatively, consider the model $\langle W, R, V\rangle$ with $W = \{w_0, w_1\}$, where $w_0 \neq w_1$, $R = \{\langle w_0, w_0\rangle, \langle w_0, w_1\rangle\}$ and $V(p) = \{w_0\}$, $V(q) = \{w_1\}$, noting that $R^2 = R$: the inset conditional formula is then false at w_0 in the model $\langle W, R, V\rangle$.)

Now attending to the distribution issue, we may split (D_1) and (D_2) into their separate directions, at the same time for definiteness instantiating the schematic letters (φ, ψ, δ) to distinct sentence letters:

$(D_1^{\vdash})$ $p \wedge (q \vee r) \vdash (p \wedge q) \vee (p \wedge r)$,
$(D_1^{\dashv})$ $(p \wedge q) \vee (p \wedge r) \vdash p \wedge (q \vee r)$,
$(D_2^{\vdash})$ $p \vee (q \wedge r) \vdash (p \vee q) \wedge (p \vee r)$,
$(D_2^{\dashv})$ $(p \vee q) \wedge (p \vee r) \vdash p \vee (q \wedge r)$.

The first and the fourth of these conditions are what we have been calling distinctively distributive, while the second and third ('distributively innocent' conditions, as it was put above) are analogues of inequalities holding in arbitrary lattices. To see whether they are satisfied when $\vdash$ is taken as $\vdash_{\mathsf{mod}}$, we spell them out explicitly in turn:

For $(D_1^{\vdash})$: $\vdash_{\mathsf{K4!}} \Box\Box(p \wedge (q \vee r)) \to \Box(\Box(p \wedge q) \vee \Box(p \wedge r))$,
For $(D_1^{\dashv})$: $\vdash_{\mathsf{K4!}} \Box(\Box(p \wedge q) \vee \Box(p \wedge r)) \to \Box\Box(p \wedge (q \vee r))$,

For $(D_2^{\vdash})$: $\vdash_{\mathsf{K4!}} \Box(p \vee \Box(q \wedge r)) \to \Box\Box((p \vee q) \wedge (p \vee r))$,
For $(D_2^{\dashv})$: $\vdash_{\mathsf{K4!}} \Box\Box((p \vee q) \wedge (p \vee r)) \to \Box(p \vee \Box(q \wedge r))$.

Let us simplify the formulas involved by collapsing any "□□" to "□", since what is at issue is their provability in K4!, where these two prefixes are equivalent:

$$\Box\big(p \wedge (q \vee r)\big) \to \Box\big(\Box(p \wedge q) \vee \Box(p \wedge r)\big);$$
$$\Box\big(\Box(p \wedge q) \vee \Box(p \wedge r)\big) \to \Box\big(p \wedge (q \vee r)\big);$$
$$\Box\big(p \vee \Box(q \wedge r)\big) \to \Box\big((p \vee q) \wedge (p \vee r)\big);$$
$$\Box\big((p \vee q) \wedge (p \vee r)\big) \to \Box\big(p \vee \Box(q \wedge r)\big).$$

With very little effort one sees that of these formulas, only the second is K4!-provable. Thus $\vdash_{\mathsf{mod}}$ fails not only the distinctively distributive conditions $(D_1^{\vdash})$ and $(D_2^{\dashv})$ but also one of the basic lattice-theoretic conditions, namely $(D_2^{\vdash})$. This helps us with the query raised above apropos of the proof of Theorem 8.1 in [6], as to whether the inclusion appealed to in that proof—$\vdash_{\mathsf{CD}} \subseteq \vdash_{\mathsf{Lat}}$—was proper. Since $\vdash_{\mathsf{mod}}$ does not satisfy the condition $(D_2^{\vdash})$, we may conclude that $\vdash_{\mathsf{mod}} \subsetneq \vdash_{\mathsf{Lat}}$, so in the light of Proposition 2.2 we have the following:

Corollary 2.3 $\vdash_{\mathsf{CD}} \subsetneq \vdash_{\mathsf{Lat}}$.

This in turn leaves us with a residual question concerning $\vdash_{\mathsf{mod}}$ itself, concerning which what we know so far is only that $\vdash_{\mathsf{CD}} \subseteq \vdash_{\mathsf{mod}}$: so again the question arises as to whether *this* inclusion is proper. On the face of it, one would expect an affirmative answer here, since in view the K4!-provability of the second formula in the above list, $\vdash_{\mathsf{mod}}$ satisfies the condition $(D_1^{\dashv})$ which there is no reason to think that $\vdash_{\mathsf{CD}}$ satisfies. A more definitive resolution of this question is deferred to the following section (see Proposition 3.1).

3 Absorption

While the distribution equivalences—specifically least in what we called their distinctively distributive directions—have, whether rightly or wrong, been hotly contested (quantum logic and so on) the absorption equivalences have had an easier ride, raising only the occasional eyebrow (as in [24]). For continuity with the preceding section we call the equivalences concerned (A_1) and (A_2):

(A_1) $\varphi \vee (\varphi \wedge \psi) \dashv\vdash \varphi$,
(A_2) $\varphi \wedge (\varphi \vee \psi) \dashv\vdash \varphi$,

though in [6] it is only the backward ("⊣") direction of (A_1) that is given that name, and only the forward ("⊢") direction of (A_2) that is so-called. The other directions of these two equivalences correspond to special cases of inequalities holding for arbitrary join semilattices and meet semilattices, respectively (thinking of "⊢" when there is only one formula to its left, as "≤"). We make a decomposition of the conditions like that given for the distribution laws, in the course of which we again switch from schemata to specific formulas:

$(A_1^{\vdash})$ $p \vee (p \wedge q) \vdash p$,
$(A_1^{\dashv})$ $p \vdash p \vee (p \wedge q)$,
$(A_2^{\vdash})$ $p \wedge (p \vee q) \vdash p$,
$(A_2^{\dashv})$ $p \vdash p \wedge (p \vee q)$.

When Béziau and Coniglio speak of absorption laws, the topic of their Sect. 11,[23] they have in mind the 'distinctively absorptive' conditions $(A_1^{\vdash})$ and $(A_2^{\dashv})$, which give us the logic of lattices proper rather than of mere bisemilattices.[24] But what do they mean when, as the content of Theorem 11.1 [6, p. 580] they say: "The logic CD of conjunction and disjunction does not satisfy the absorption laws"? Do they mean that $\vdash_{\mathsf{CD}}$ does not satisfy *both* of $(A_1^{\vdash})$, $(A_2^{\dashv})$, or do they mean that $\vdash_{\mathsf{CD}}$ does not satisfy *either* of $(A_1^{\vdash})$, $(A_2^{\dashv})$? Since, as we shall see, $\vdash_{\mathsf{CD}}$ fares differently w.r.t. these two conditions, we must interpret the claim of Theorem 11.1 in the former way to make it come out true: that is, in fact one of the two conditions is satisfied but not the other. It turns out, however, that the *proof* offered to justify the claim picks the wrong one as the one that fails. (One might simply assume that the fact that the two conditions are dual rules out this possibility—in which case the present situation serves as a reminder how the left–right asymmetry inherent in working with consequence relations thwarts duality-generated expectations. Recall the discussion of the oppositely lopsided dual-consequence relations from the previous section.)

The proof in question directs us to the algebra of the same signature as our formula algebra For, whose universe is [0, 1], the closed real interval between 0 and 1, and whose binary operations, corresponding to the connectives $\wedge$ and $\vee$ of For, are—now quoting directly (from [6, p. 580]) for the rest of this paragraph—the "usual product $\cdot$ and the truncated sum $\oplus$ such that

$$a \oplus b = \begin{cases} a + b & \text{if } a + b \leq 1, \\ 1 & \text{otherwise.} \end{cases}$$

Consider homomorphisms $v : \mathsf{For} \longrightarrow [0, 1]$ such that $v(\varphi \wedge \psi) = v(\varphi) \cdot v(\psi)$ and $v(\varphi \vee \psi) = v(\varphi) \oplus v(\psi)$. Define a relation $\vdash_{[0,1]}$ as follows: given $\Gamma \cup \{\varphi\} \subseteq \mathsf{For}$, $\Gamma \vdash_{[0,1]} \varphi$ iff there exists a finite nonempty set $\{\gamma_1, \dots, \gamma_n\} \subseteq \Gamma$ such that, for every homomorphism $v : \mathsf{For} \longrightarrow [0, 1]$, $v(\gamma_1) \cdots v(\gamma_n) \leq v(\varphi)$."

The authors go on to show that $\vdash_{\mathsf{CD}} \subseteq \vdash_{[0,1]}$, and then conclude that $\vdash_{\mathsf{CD}}$ does not satisfy the absorption laws on the ground that $\vdash_{[0,1]}$ does not, the justification for the latter claim being: "for instance, $\frac{1}{2} \not\leq \frac{1}{2} \cdot (\frac{1}{2} \oplus \frac{1}{4})$." This strongly suggests that the target is $(A_2^{\dashv})$:

---

[23] This section (p. 579) is headed "Another bridge principles". Whether the authors meant "Another bridge principle", counting the absorption laws collectively as a single principle or "Other (better: "Further") Bridge Principles", counting them separately, is not clear. (The subject of what which conditions involving the interaction of two connectives (here $\wedge$ and $\vee$) constitute *bridge principles*—terminology taken in [6] from [9]—is interesting in its own right, but for reasons of space is not taken up here.) In line 10 of p. 578 of [6], it is clear that for "another connectives" the authors intended "other connectives". Numerous other linguistic and typographical mishaps are equally easily corrected, such as in the last paragraph of Sect. 1 (p. 567) for "necessary distributive" read "necessarily distributive", as in line 8 of Sect. 6 (p. 573), and in Conjecture 2 (p. 571) for "necessary leads", read "necessarily leads".

[24] For definitions, some discussion, and further references on bisemilattices see p. 633 and elsewhere in [24].

we put $v(p) = \frac{1}{2}$, $v(q) = \frac{1}{4}$ and now we have $v(p \wedge (p \vee q)) = \frac{1}{2} \cdot (\frac{1}{2} \oplus \frac{1}{4}) = \frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}$, so certainly $v(p) \nleq v(p \wedge (p \vee q))$. But this does not show that $p \nvdash_{[0,1]} p \wedge (p \vee q)$, as is apparently being assumed. The words "there exists a finite nonempty set $\{\gamma_1, \ldots, \gamma_n\} \subseteq \Gamma$ such that, for every homomorphism $v : \mathsf{For} \longrightarrow [0, 1]$, $v(\gamma_1) \cdots v(\gamma_n) \leq v(\varphi)$," in the quoted passage must allow us to choose $\gamma_i = \gamma_j$ even when $i \neq j$, or else this same choice of $v(p)$ would similarly show that $p \nvdash_{[0,1]} p \wedge p$ since $\frac{1}{2} \nleq (\frac{1}{2} \cdot \frac{1}{2})$, contradicting the claim that $\vdash_{\mathsf{CD}} \subseteq \vdash_{[0,1]}$. (That is, we must allow $\{\gamma_1, \gamma_2\} \subseteq \Gamma = \{p\}$ with $\gamma_1 = \gamma_2 = p$, since for all v, $v(\gamma_1) \cdot v(\gamma_2) \leq v(p \wedge p)$.)

We thus lose any reason for doubting that $\vdash_{\mathsf{CD}}$ satisfies $(A_2^{\dashv})$, which is just as well, since it plainly does: we have $p \vdash_{\mathsf{CD}} p$ and $p \vdash_{\mathsf{CD}} p \vee q$ since these $\vdash$-claims hold for $\vdash_{\vee}$, and then using substitution invariance and the fact that we have $\varphi, \psi \vdash_{\mathsf{CD}} \varphi \wedge \psi$ since this holds for $\vdash_{\wedge}$, taking φ as p and ψ as $p \vee q$, we get $p \vdash_{\mathsf{CD}} p \wedge (p \vee q)$.[25]

Despite this misstep in the proof, Theorem 11.1 of [6], as we are interpreting it, is perfectly correct: $\vdash_{\mathsf{CD}}$ does *not* satisfy the absorption laws in their entirety. This is what one would expect from the fact that while, as in the argument of the previous paragraph, we can give an unconditional formulation of the ($\wedge$-introduction or '$\wedge$ Right') condition called in [6] ($\wedge_3$)—namely as $\varphi, \psi \vdash_{\mathsf{CD}} \varphi \wedge \psi$—what would be needed for a similar treatment of the other distinctively absorptive condition $(A_1^{\vdash})$ would be the irreducibly conditional condition ($\vee_3$) (or the similarly conditional, weakened form, ($\vee_{3w}$)). To show that $\vdash_{\mathsf{CD}}$ does not satisfy $(A_1^{\vdash})$, we can use the modal resources of the preceding section, as it suffices to show that $\vdash_{\mathsf{mod}}$ does not satisfy $(A_1^{\vdash})$, since $\vdash_{\mathsf{CD}} \subseteq \vdash_{\mathsf{mod}}$. That is, it suffices to observe that $\nvdash_{\mathsf{K4!}} \Box(p \vee \Box(p \wedge q)) \to \Box p$, since this formula is false at w_1 in the same two-element model as was mentioned after the proof of Proposition 2.2.[26]

For those preferring a more traditional matrix-based argument to this same conclusion, consider the (logical) matrix depicted in Fig. 1, which follows the convention of using an asterisk to indicate the first occurrence of the designated value(s) in the tables describing the algebra of the matrix. We can see that the consequence relation determined by this matrix[27] is a substitution invariant extension of $\vdash_{\wedge}$ and $\vdash_{\vee}$, and thus of $\vdash_{\mathsf{CD}}$, because the reducts of the algebra to $\wedge$ and to $\vee$ are both semilattices, depicting greatest lower bound in the first case w.r.t. the order $1 \geq 2 \geq 4$ and, in the second case, least upper bounds w.r.t. the order $1 \geq 2, 3$ (2, 3 incomparable). Because the orders are different, this bisemilattice is not a lattice, and we can expect trouble on the absorption front. In particular, the consequence relation determined by the matrix does not satisfy $(A_1^{\vdash})$, since

[25] Here we are using the transitivity or Cut property of consequence relations—"CUT" on p. 567 of [6] (or what is called 'Cut for sets' at p. 15 of Shoesmith and Smiley [33], in the special case with the set Z a singleton); of course, we do not have to write "$p, p \vdash_{\mathsf{CD}} p \wedge (p \vee q)$" and apply some contraction-like condition to get to our final destination, since the formulas on the left of any "$\vdash$" are taken as comprising a set (not a multiset or a sequence) of formulas.

[26] This is one place—and not the only one—where the choice in Definition 2.1 of K4! as the pertinent modal logic, as opposed to, e.g. the stronger S4 (or 5), earns its keep: the implicational formula just mentioned is S4-provable. Incidentally, in the Kripke countermodel just cited, while $V(p)$ must be chosen as in the earlier discussion, the choice of $V(q)$ is immaterial for falsifying the current implicational formula at w_0.

[27] The consequence relation determined by a matrix holds between Γ and φ when on every matrix evaluation (homomorphism) on which all formulas in Γ receive a designated value, so does φ. If we were conducting the discussion in terms of (SET–FMLA) sequents, this is what it is for the sequent $\Gamma \succ \varphi$ to be valid in the matrix.

Fig. 1 A matrix counterexample for $(A_1^{\vdash})$

$\wedge$	1	2	3
*1	1	2	3
2	2	2	3
3	3	3	3

$\vee$	1	2	3
1	1	1	1
2	1	2	1
3	1	1	3

if we assign to the sentence letters p and q the values 2 and 3, $p \wedge q$ receives the value $2 \wedge 3 = 3$ and so $p \vee (p \wedge q)$ receives the value $2 \vee 3 = 1$, so these values we pass from a designated value on the left to an undesignated value on the right in considering whether $p \vee (p \wedge q) \vdash p$ for the consequence relation $\vdash$ determined by the matrix of Fig. 1, showing again that $\vdash_{\mathsf{CD}}$ does not satisfy $(A_1^{\vdash})$ (as this $\vdash \supseteq \vdash_{\mathsf{CD}}$).

This matrix has been introduced into the discussion not simply as an alternative to the use of Kripke models, but to settle a question left over from Sect. 2 which they could not settle there—at least not via $\vdash_{\mathsf{mod}}$. As noted at the end of that section, $\vdash_{\mathsf{mod}}$ satisfies the distribution condition $(D_1^{\dashv})$, making it unlikely that $\vdash_{\mathsf{mod}}$ coincided with $\vdash_{\mathsf{CD}}$. (Recall that we already knew that $\vdash_{\mathsf{CD}} \subseteq \vdash_{\mathsf{mod}}$.) We can now settle this question, using the fact that $\vdash_{\mathsf{CD}}$ is included in the consequence relation determined by the matrix of Fig. 1—or to put it more simply, using the soundness of $\vdash_{\mathsf{CD}}$ w.r.t. that matrix. Thus it suffices to find an invalidating matrix evaluation, as we do here, successively calculating values of subformulas from the sentence letters, from since the left-hand formula end up with a designated and the right-hand formula an undesignated value:

$$
\begin{array}{ccccccccccc}
(p & \wedge & q) & \vee & (p & \wedge & r) & \vdash & p & \wedge & (q \vee r) \\
2 & & 2 & & 2 & & 3 & & 2 & & 2 \quad\; 3 \\
& 2 & & & & 3 & & & 2 & & 1 \\
& & & 1 & & & & & & 2 &
\end{array}
$$

Thus we conclude that, as suspected, $\vdash_{\mathsf{CD}}$ does not satisfy $(D_1^{\dashv})$, which allows us to supplement Corollary 2.3 to give a fuller description of the relations among the consequence relations in play in Sect. 2:

Proposition 3.1 $\vdash_{\mathsf{CD}} \subsetneq \vdash_{\mathsf{mod}} \subsetneq \vdash_{\mathsf{Lat}}$.

Digression Before returning to $\vdash_{\mathsf{CD}}$ itself, a word might be said about the mysteriously titled Sect. 7 of [6], 'No Semantics for Non-Distributive logics', which promises "a very general result about the impossibility of finding a semantics for a non-distributive logic of conjunction and disjunction taken as a consequence relation. To prove this result, we use a very general definition of semantics. (...) The following result really shows that there are no semantics for non-distributive logics within the Tarskian paradigm."[28] In the ellipsed passage what is given is a description of matrix semantics, so the claim appears to be that no consequence relation in the language For not satisfying the distribution conditions can be many-valued, in the sense of being determined by a matrix (see footnote 27);

[28] This passage appears on p. 574 of [6]; I have changed "impossibility to find a semantics" to "impossibility of finding a semantics".

recall that such a matrix may have 1, 2, or any finite number, or be of any infinite cardinality. Since we have just seen that the consequence relation determined by the matrix of Fig. 1, which by definition is a many-valued consequence relation, the implied claim is evidently false. What is the result, then, to which the remarks just quoted are a prelude? It (= Theorem 7.1, [6, p. 575]) reads "The laws of distributivity hold in every semantics validating the laws $(\wedge_1)$–$(\wedge_3)$ and $(\vee_1)$–$(\vee_3)$." But of course they do (at least on the most reasonable decipherment of talk of—possibly *conditional*—laws holding in a semantics): we already know that any consequence relation satisfying the cited conditions, because of the presence of $(\vee_3)$ in particular, satisfies the distribution conditions. This is no way shows that matrix semantics cannot exist "for non-distributive logics within the Tarskian paradigm," at least to the extend that what the Tarskian paradigm might be from the perspective of anything said in [6]—Tarskian consequence systems and all that. It is true that Tarski also gave what have subsequently been called Tarski-style conditions (see the index entry under 'Tarski-style conditions' in [28] for examples and references) on pairs comprising a consequence relation and a connective in its language to justify the connective's being called conjunction or disjunction (or negation or implication), with special reference to the behavior of these connectives in intuitionistic and classical logic, and for the cases of conjunction and disjunction these amount to $(\wedge_1)$–$(\wedge_3)$ and $(\vee_1)$–$(\vee_3)$, but this shows nothing about the impossibility of finding "a semantics for a non-distributive logic of conjunction and disjunction taken as a consequence relation," even when attention is restricted to a semantic account in terms of single matrices (let alone collections thereof or "$\leq$-based" semantics[29] as it is called in Sect. 2.14 of [28]—of the kind deployed by Béziau and Coniglio, for instance, in describing $\vdash_{\mathsf{Lat}}$ semantically). So quite what the authors are up to in Sect. 7 of [6] escapes me. **End of Digression.**

With the inclusion relations among between $\vdash_{\mathsf{CD}}$, $\vdash_{\mathsf{mod}}$ and $\vdash_{\mathsf{Lat}}$ having been sorted out in Proposition 3.1, in part with the aid of a fourth consequence relation, namely that determined by the matrix of Fig. 1, the question arises as to whether this is indeed a fourth consequence relation or instead coincides with $\vdash_{\mathsf{CD}}$ itself. If that were, we would have a simple semantic description of $\vdash_{\mathsf{CD}}$ as a three-valued logic.[30] A glance at the $\wedge$ table in the matrix of Fig. 1 shows us that this is not so, however, when taken together with some simple considerations about congruentiality, for which we need to recall some terminology (from [28], adapted from the usage of David Makinson, Krister Segerberg and others):

[29]This phrase seems preferable to talk—as in Bou et al. [8]—of preserving degrees of truth, because, in the first place, even for applications for which degrees of truth have been postulated by some (such as vagueness) their postulation is highly contentious, and secondly, because for other applications degrees of truth are evidently beside the point: for example, the local consequence relation for modal logic, where "$\leq$" amounts to inclusion of sets of verifying worlds rather than anything one could regard as comparing degrees of truth. The terminology of $\leq$-based semantics has its own disadvantages, though—in particular, in the risk of confusion it invites with *order algebraizability* as in Raftery [31]. Incidentally, despite our earlier casual talk of "the logic of lattices proper rather than of mere bisemilattices", a $\leq$-based semantic account of the logic of bisemilattices faces an immediate difficulty: in the general case (of bisemilattices which are not lattices) there are two candidate choices of $\leq$ to choose from.

[30]A consequence relation is n-valued if it is determined by a matrix with (i.e. the universe of whose algebra has) n elements but by no matrix with fewer than n elements.

Definition 3.2 An n-ary connective # in the language of a consequence relation $\vdash$ is *congruential in its* ith *position* (where $1 \leq i \leq n$) according to $\vdash$ when for all formulas $\varphi_1, \ldots, \varphi_n, \psi$ of the language of $\vdash$, if $\varphi_i \dashv\vdash \psi$ then $\#(\varphi_1, \ldots, \varphi_n) \dashv\vdash \#(\varphi_1, \ldots, \varphi_{i-1}, \psi, \varphi_{i+1}, \ldots, \varphi_n)$; # itself is *congruential* according to $\vdash$ if # is congruential in each of its n positions according to $\vdash$; and $\vdash$ is *congruential* if every primitive connective in the language of $\vdash$ is congruential according to $\vdash$.

A conspicuous respect in which $\vdash_{\mathsf{CD}}$ contrasts with $\vdash_\wedge$ and $\vdash_\vee$ is that, unlike them, it is not congruential, specifically because although $\wedge$ is congruential according to $\vdash_{\mathsf{CD}}$, $\vee$ is not (in either of its two positions). For example, while $q \dashv\vdash_{\mathsf{CV}} q \wedge q$, we do not have $p \vee q \dashv\vdash_{\mathsf{CV}} p \vee (q \wedge q)$. The easiest way to see this quickly, recalling Definition 2.1, is to use the fact that $\vdash_{\mathsf{CD}} \subseteq \vdash_{\mathsf{mod}}$, since $\vdash_{\mathsf{K4!}} \Box q \leftrightarrow \Box\Box(q \wedge q)$ while $\nvdash_{\mathsf{K4!}} \Box(p \vee q) \leftrightarrow \Box(p \vee \Box(q \wedge q))$. (We could of course simplify these occurrences of "$\Box(q \wedge q)$" to "$\Box q$".[31]) The general point here in the terminology of Smiley [34] is that while φ and $\varphi \wedge \varphi$ are, according to $\vdash_{\mathsf{CD}}$, equivalent, they are not *synonymous* (freely interreplaceable in all contexts);[32] one might put this by saying that $\wedge$ is idempotent though not *pervasively idempotent* according to $\vdash_{\mathsf{CD}}$.[33] However, according to the consequence relation determined by the matrix in Fig. 1, the φ and $\varphi \wedge \varphi$ must be synonymous (for all φ), since the operation associated with conjunction, for which, as in Fig. 1, we use the same symbol, $\wedge$, here, we have $x \wedge x = x$ for each matrix element x. Thus $p \vee q \dashv\vdash p \vee (q \wedge q)$ where $\vdash$ is the consequence relation determined by our matrix, so this $\vdash \supsetneq \vdash_{\mathsf{CD}}$: alternatively put, $\vdash_{\mathsf{CD}}$ is sound, but not complete, w.r.t. this matrix. (The failure of congruentiality here is not surprising, since one expects connectives written as $\wedge$ and $\vee$ to be monotonic according a consequence relation in whose language they appear, and *a fortiori* congruential. But in Sect. 1, we recalled the close connection Rautenberg [32] observed between monotonicity for $\vee$ and the condition $(\vee_3)$—which $\vdash_{\mathsf{CD}}$ does not satisfy.[34])

A different 3-element matrix emerges from the consideration of $\vdash_{\mathsf{mod}}$ in Sect. 2, since every appeal there to K4! was to a specific frame $\langle W, R\rangle$ for with two elements, which can accordingly be traded in for the four-element algebra of subsets of W, though for our purposes we can ignore $\varnothing$ and re-name $\{w_0, w_1\}$, $\{w_0\}$ and $\{w_1\}$, "1", "2" and "3", calculating the effect of the $\vdash_{\mathsf{CD}}$ connectives $\wedge$ and $\vee$, recalling that $\wedge$ amounts to a $\Box$-ed version of K4!'s own $\wedge$ (via the translation τ given in that section). The effect of this is to change the entries in the $\wedge$ table by replacing all 2s with 3s, and leaving the $\vee$ table as it is. This means that we no longer have $x \wedge x$ always equal to x, while still equi-designated with x, making room for $\wedge$ to be idempotent but not pervasively idempotent

[31] This is because the so-called local consequence relation associated with K4!, as with any normal modal logic, is itself congruential.

[32] In the terminology of Humberstone [29], they are equivalent according to $\vdash_{\mathsf{CD}}$ but not according to the replacement core, $\vdash^{\mathsf{rep}}_{\mathsf{CD}}$, of that consequence relation.

[33] Although pervasive idempotence is a stronger property than idempotence, we avoid the phrase "strongly idempotent" because this is used for a different property in [28] and [30].

[34] I am skirting over the "with/without side formulas issue here for a quick formulation. For monotonicity without side formulas, the correspondence is with $(\vee_{3w})$. Similarly, while the word *congruential* as introduced in Definition 3.2 refers to a "without side formulas" condition (sometimes called self-extensionality in the literature), what Rautenberg means by congruentiality is the "with side formulas" analogue—called *extensionality* in [28] and elsewhere.

according to the consequence relation determined by the altered matrix. One complication stops this from being the consequence relation $\vdash_{\mathsf{mod}}$, however, and that is that while we have used τ to guide in the construction of the matrix, Definition 2.1 inserts a further initial occurrence of $\Box$ to the τ-translations of formulas over and above any arising in the translation of conjunctions.[35]

Let us not pursue that complication further here, however, since the interest of $\vdash_{\mathsf{mod}}$ here lies in the way it enables us to prove things about $\vdash_{\mathsf{CD}}$ and in our final section we shall see that restricting attention to any single finite (e.g. 3-element) matrix is not going to help with that project. This is not to suggest that any deep insight into $\vdash_{\mathsf{CD}}$ is likely to come from matrix methodology, with its emphasis on designated values, or even its transformation into the algebraic logic of Blok and Pigozzi [7], with algebraizability's equally voracious appetite for equivalence formulas—demands with which an atheorematic (or 'purely inferential') consequence relation like $\vdash_{\mathsf{CD}}$ is not in a position to satisfy, any more than $\vdash_{\wedge\vee}$, the full conjunction–disjunction fragment of classical consequence; indeed this applies to the broader class of protoalgebraic logics rather than just the algebraizable ones.[36] (More promising is a description in terms of bivalent valuational semantics, which we will see briefly in Sect. 5.) However, for the record as well as because of the independent interest of showing that neither $\vdash_{\mathsf{CD}}$ nor $\vdash_{\mathsf{mod}}$ is locally tabular (as finitely many non-equivalent formulas in any finitely number of sentence letters), we address the existence of a finite determining (or 'strictly characteristic') matrix for $\vdash_{\mathsf{CD}}$. According to Shoesmith and Smiley [33] (independently reworking a correction by R. Wójcicki to a not-quite-correct earlier proof by Łoś and Suszko—for references, see [33, p. 278]) any substitution invariant consequence relation satisfying a certain (as they call it) *Cancellation Condition* is many-valued in the sense of being determined (or 'characterized') by some matrix or other, a condition I assume $\vdash_{\mathsf{CD}}$ does satisfy,[37] with so the most obvious question remaining is as to whether such a matrix can be found which is finite. Readers in whom this question arouses no curiosity should pass directly to Sect. 5.

4 Finite Matrices and $\vdash_{\mathsf{CD}}$

We are interested in the set of formulas of the language for K4! of the form $\Box\tau(\varphi)$ with φ an $\{\wedge, \vee\}$-formula and τ the translation used to specify $\vdash_{\mathsf{mod}}$ in Definition 2.1, or else—to avoid gratuitously long modal formulas—the result of replacing subformulas $\Box(\delta \wedge \delta)$ in such formulas by $\Box\delta$. The set of these formulas, we call the $\vdash_{\mathsf{mod}}$-fragment of K4!, in connection with which we say formulas φ and ψ are equivalent when $\vdash_{\mathsf{K4!}} \varphi \leftrightarrow \psi$.

There are infinitely many pairwise non-equivalent formulas in a single propositional variable in the $\vdash_{\mathsf{mod}}$-fragment of K4!.

Consider the following sequence of formulas from the $\vdash_{\mathsf{mod}}$-fragment of K4!.

- $\varphi_1 = \Box p$,
- $\varphi_{n+1} = \Box(\varphi_n \vee p)$.

---

[35] Without this, we would not have, for example, $p, q \vdash_{\mathsf{mod}} p \wedge q$.

[36] See Font et al. [14, 16], and [15] for these problems and a suggested resolution.

[37] In the case of finitary $\vdash$ the condition in question simplifies to: if $\Gamma, \Delta \vdash \varphi$ and the formulas in Δ have sentence letters in common with those of $\Gamma \cup \{\varphi\}$, then either $\Gamma \vdash \varphi$ or else for all ψ, we have $\Delta \vdash \psi$.

The corresponding sequence of $\vdash_{\mathsf{mod}}$-formulas themselves—i.e. with φ_i being a the result of prefixing a $\Box$ to the (simplified) τ-translation of ψ_i—is as follows:

- $\psi_1 = p$,
- $\psi_{n+1} = ((\psi_n \wedge \psi_n) \vee p)$.

Note that $\vdash_{\mathsf{K4!}} \varphi_n \to \varphi_{n+1}$; in $\vdash_{\mathsf{mod}}$ terms: $\psi_n \vdash_{\mathsf{mod}} \psi_{n+1}$. We claim that the converse implication of the former implication is not K4! provable—and thus that $\psi_{n+1} \nvdash_{\mathsf{mod}} \psi_n$—for any n, and will substantiate this claim by working through a representative case giving the general idea, showing by appeal to the semantic characterization of K4! as the logic determined by the class of models (on frames) whose accessibility relations R satisfy $R^2 = R$ ('dense transitive frames'), that $\nvdash_{\mathsf{K4!}} \varphi_4 \to \varphi_3$. Thus the formula needing to be invalidated on a frame of the kind described is:

$$\Box\big(\Box\big(\Box(\Box p \vee p) \vee p\big) \vee p\big) \to \Box\big(\Box(\Box p \vee p) \vee p\big),$$

whose K4!-unprovability shows that for $\psi_3 = (((p \wedge p) \vee p) \wedge ((p \wedge p) \vee p)) \vee p$, and $\psi_4 = (\psi_3 \wedge \psi_3) \vee p$, we have $\psi_4 \nvdash_{\mathsf{mod}} \psi_3$. Now in fact rather than working with the modal formula inset above it is easier to work with its dual form, whose equi-provability with the original in K4! will show the original to be unprovable:

$$\Diamond\big(\Diamond(\Diamond p \wedge p) \wedge p\big) \to \Diamond\big(\Diamond\big(\Diamond(\Diamond p \wedge p) \wedge p\big) \wedge p\big).$$

Consider the model $\mathcal{M} = \langle W, R, V\rangle$ with $W = \{w_0, w_1, \ldots, w_6\}$ and Rw_iw_j just in case i is odd and $i < j$ or i is even and $i \leq j$. Finally let $V(p) = \{w_i \in W \mid i \text{ is odd}\}$. Thus our frame consists of a string of alternating reflexive and irreflexive points, each having everything later than it accessible to it, and we are considering an arbitrary model on this frame in which p is true at the precisely the irreflexive points, the last of which (w_6) has nothing accessible to it. (In fact, whether we put w_0 into $V(p)$ as here, or not, makes no difference.) The reflexive points are interpolated between the irreflexive ones in order to secure that $R \subseteq R^2$ (the converse inclusion—transitivity—being immediate from the way R was defined).

Claim The implication above is false at w_0 in $\mathcal{M}$, having a true antecedent and a false consequent, or, using some standard notation:

$$\text{(a) } \mathcal{M} \models_{w_0} \Diamond\big(\Diamond(\Diamond p \wedge p) \wedge p\big) \quad \text{and} \quad \text{(b) } \mathcal{M} \not\models_{w_0} \Diamond\big(\Diamond\big(\Diamond(\Diamond p \wedge p) \wedge p\big) \wedge p\big).$$

Checking (a) first: $\mathcal{M} \models_{w_6} p$, as $w_6 \in V(p)$. Since Rw_4w_6 and also $w_4 \in V(p)$, $\mathcal{M} \models_{w_4} \Diamond p \wedge p$. So, again, since Rw_2w_4 and $w_2 \in V(p)$, $\mathcal{M} \models_{w_2} \Diamond(\Diamond p \wedge p) \wedge p$. But Rw_0w_2, so $\mathcal{M} \models_{w_0} \Diamond(\Diamond(\Diamond p \wedge p) \wedge p)$.

Turning to (b), we need to check that $\Diamond(\Diamond(\Diamond p \wedge p) \wedge p) \wedge p$ is false at all points other than w_0, since these are the points R-accessible to w_0. Since the second conjunct is false at w_1, w_3, w_5, this leaves only w_2, w_4 and w_6 to check as possible counterexamples where the formula might be true. But if its first conjunct is true at any of them, then by the transitivity of R, it, and therefore the conjunction, is true at w_2, so we only have to rule that out. Suppose $\mathcal{M} \models_{w_2} \Diamond(\Diamond(\Diamond p \wedge p) \wedge p)$. This would mean that $\Diamond(\Diamond p \wedge p) \wedge p$ is true at some point accessible to w_2, which because of the second conjunct, means that that point would have to be w_4 or w_6. Since the first conjunct begins with $\Diamond$, the conjunction is certainly false at w_6 as there is no x with Rw_6x. Thus the only remaining option is that $\mathcal{M} \models_{w_4} \Diamond(\Diamond p \wedge p)$, which would imply that either w_5 or w_6 verifies $\Diamond p \wedge p$, which is not the case for w_6 because of the first conjunct, and not the case for w_5 because of the second. So (b) is correct after all.

The above argument establishes, by considering the dual form with $\Diamond$, that specifically $\nvdash_{\mathsf{K4!}} \varphi_4 \to \varphi_3$, but suggests its own adaptation to the case of arbitrary n—making the string of alternating irreflexive and reflexive points of length $2n + 1$ in the model construction—showing that $\nvdash_{K4!} \varphi_{n+1} \to \varphi_n$. This implies that for no distinct m, n, are φ_m and φ_n K4!-equivalent, and thus finally establishes:

Proposition 4.1 *There are infinitely many pairwise non-equivalent formulas in a single propositional variable, taking equivalence as mutual consequence for* $\vdash_{\mathsf{mod}}$.

Proof By the reasoning above, whenever $m \neq n$, we have $\psi_m \nvdash_{\mathsf{mod}} \psi_n$ or else $\psi_n \nvdash_{\mathsf{mod}} \psi_m$. □

Since $\vdash_{\mathsf{CD}} \subseteq \vdash_{\mathsf{mod}}$ (Proposition 2.2), we have:

Corollary 4.2 *There are infinitely many pairwise non-equivalent formulas in a single propositional variable, when equivalence is taken mutual consequence by* $\vdash_{\mathsf{CD}}$.

Finally, using the fact (in the literature, "tabular implies locally tabular") that there are only finitely many non-equivalent formulas the logic determined by a finite matrix, we have our main result:

Corollary 4.3 $\vdash_{\mathsf{CD}}$ *is not determined by any finite matrix.*

5 Coda on Valuational Semantics for $\vdash_{\mathsf{CD}}$

Despite its not being two-valued, $\vdash_{\mathsf{CD}}$ can be described informatively in terms of bivalent valuations; though not to the fore in [6], this mode of description is much in evidence in other publications of Béziau's (e.g. [3]), and it will guide us in what follows. Since, as have just seen, $\vdash_{\mathsf{CD}}$ is far from being two-valued in the sense of footnote 30, the valuations concerned do not associate with each connective $\wedge$, $\vee$, a two-valued truth-function for computing the values of compounds in a locally compositional manner;[38] it will come as no surprise that $\vee$, in particular, is not amenable to such a treatment. We recall a few definitions and facts for semantics in terms of (bivalent) valuations. (Further details, relevant proofs, etc. can be found in [22] taken together with Sects. 1.12–1.15 of [28].) For this discussion we use T and F for the truth-values rather than 1 and 0 as in echoing Béziau above (see footnote 3). A valuation for a language is for present purposes a function from the set of formulas to the set $\{T, F\}$. Such a valuation (for the language of a

[38]In terms of the h/v_h distinction (for matrix semantics) of footnote 3, compositionality reigns at the h level rather than the bivalent v_h level. The contrast has been variously stressed in work of Suszko, Dummett and Scott.

consequence relation $\vdash$) is consistent with $\vdash$ if there is no set of $\Gamma \cup \{\psi\}$ formulas of the language for which $v(\psi) = T$ for all $\varphi \in \Gamma$ while $v(\psi) = F$; the set of all valuations consistent with $\vdash$ is denoted by $Val(\vdash)$. $\vdash$ is *determined by* a set V of valuations when for all Γ, ψ, we have $\Gamma \vdash \psi$ if and only if every $v \in V$ assigning T to each formula in Γ assigns T to ψ.[39] Every consequence relation is determined by $Val(\vdash)$, and typically by many proper subsets $Val(\vdash)$. The *conjunctive combination* $u \cdot v$ of valuations (for the same language) u and v is the unique valuation satisfying: for all formulas (of this language) φ, $u \cdot v(\varphi) = T$ iff $u(\varphi) = T$ and $v(\varphi) = T$; their *disjunctive combination* is the valuation $u + v$ satisfying (again, for all φ): φ, $u + v(\varphi) = T$ iff $u(\varphi) = T$ or $v(\varphi) = T$ (or both). These two binary operations can be understood as lifted to applying to arbitrary collections of valuations in the obvious way. If V is a class of valuations then V^{Π} and V^{Σ} are the classes, respectively, of all conjunctive combinations of valuations in V and of all disjunctive combinations in V. A valuation v for a language with $\vee$ (resp., $\wedge$) is $\vee$-boolean (resp., $\wedge$-boolean) iff for all formulas φ, ψ, of the language, $v(\varphi \vee \psi) = T$ iff $v(\varphi) = T$ or $v(\psi) = T$ (resp., iff for all φ, ψ, we have $v(\varphi \wedge \psi) = T$ iff $v(\varphi) = T$ and $v(\psi) = T$). The class of all $\vee$-boolean valuations (for an unspecified but fixed language with $\vee$) as a connective we denote by $V_{\vee}$; similar the class of $\wedge$-boolean valuations is $V_{\wedge}$. Finally, $\vdash_0 \dot{\cup} \vdash_1$ is the join of consequence relations $\vdash_0$ and $\vdash_1$ on the same language in the lattice (with ordering $\subseteq$) of consequence relations on that language. The basic facts that need to be recalled here are the following:

1. For any consequence relations $\vdash_0$ and $\vdash_1$ on the same language,

$$Val(\vdash_0 \dot{\cup} \vdash_1) = Val(\vdash_0) \cap Val(\vdash_1).$$

2. For any consequence relation $\vdash$, $(Val(\vdash))^{\Pi} \subseteq Val(\vdash)$.
3. $(V_{\wedge})^{\Pi} \subseteq V_{\wedge}$.
4. $(V_{\vee})^{\Pi} \nsubseteq V_{\vee}$ (on the other hand: $(V_{\vee})^{\Sigma} \subseteq V_{\vee}$).
5. $Val(\vdash_{\#}) = (V_{\#})^{\Pi}$, for $\# = \wedge, \vee$.

To bring Fact 1 to bear on the case of $\vdash_{\mathsf{CD}}$ it will help to reconstrue the labels $\vdash_{\wedge}$ and $\vdash_{\vee}$ as denoting substitution invariant consequence relations on the same language, having as primitive connectives both $\wedge$ and $\vee$, respectively determined by $V_{\wedge}$ and $V_{\vee}$, taken as the classes of $\wedge$-boolean and $\vee$-boolean valuations for this language.[40] Then we can identify $\vdash_{\mathsf{CD}}$ as $\vdash_{\wedge} \dot{\cup} \vdash_{\vee}$. Invoking Fact 1, we accordingly have:

$$Val(\vdash_{\mathsf{CD}}) = Val(\vdash_{\wedge}) \cap Val(\vdash_{\vee}).$$

Simplifying further with the aid of Fact 5, we have:

$$Val(\vdash_{\mathsf{CD}}) = (V_{\wedge})^{\Pi} \cap (V_{\vee})^{\Pi},$$

and further, using Fact 3:

$$Val(\vdash_{\mathsf{CD}}) = V_{\wedge} \cap (V_{\vee})^{\Pi}. \qquad (*)$$

[39] Thus $\vdash$'s being determined by a matrix in the sense of footnote 27 is a matter of its being determined (in the current sense) by the class of induced valuations v_h with h a matrix evaluation, as explained in footnote 3.

[40] Thus distinct $\vee$-compounds ($\wedge$-compounds) simply behave like distinct sentence letters as far as $\vdash_{\wedge}$ (resp., $\vdash_{\vee}$) is concerned.

We cannot remove the reference ("Π") to taking conjunctive combinations in the second $\cap$-term on the right-hand side of (*), in view of Fact 4.[41] In the case of $\vdash_\vee$ we have this consequence relation is determined by the set of all its consistent valuations (i.e. the conjunctive combinations of $\vee$-boolean valuations) as well as by the proper subset comprising just the $\vee$-boolean valuations themselves—given the latter result, the former follows by appeal to Fact 2 above—but we cannot say similarly that just as $\vdash_{\mathsf{CD}}$ is determined by (the class of all its consistent valuations) $V_\wedge \cap (V_\vee)^\Pi$ so it is also determined by $V_\wedge \cap V_\vee$, since of course the consequence relation determined by this class is the stronger $\vdash_{\wedge\vee}$ (which $Val(\cdot)$ maps to $(V_\wedge \cap V_\vee)^\Pi$).

To conclude this discussion, we examine the asymmetry observed in Sect. 3 between $(A_1^\vdash)$ and the dual absorption principle $(A_2^\dashv)$:

$$(A_1^\vdash)\ p \vee (p \wedge q) \vdash p \qquad (A_2^\dashv)\ p \vdash p \wedge (p \vee q).$$

Recall that we found the second but not the first of these conditions to be satisfied for $\vdash = \vdash_{\mathsf{CD}}$. The present aim is illuminate this contrast with the aid of (*) above. Let us begin with $(A_2^\dashv)$. Take a valuation v consistent with $\vdash_{\mathsf{CD}}$ and verifying the left-hand formula in $(A_2^\dashv)$: $v(p) = T$. According to (*), v is both $\wedge$-boolean and also a conjunctive combination of $\vee$-boolean valuations u_i say (where $i \in I$, some index set); since $v(p) = T$, for each u_i, we have $u_i(p) = T$, and so, u_i being $\vee$-boolean, $u_i(p) = T$. As this is so for each $i \in I$ and v is the conjunctive combination of the u_i, we have $v(p \vee q) = T$. So now we have $v(p) = T$ and $v(p \vee q) = T$. But v is $\wedge$-boolean, giving us the desired conclusion that $v(p \wedge (p \vee q)) = T$.

By contrast, suppose a valuation v consistent with $\vdash_{\mathsf{CD}}$ verifies the left-hand side of $(A_1^\vdash)$. As before, v is $\wedge$-boolean and is the conjunctive combination of $\vee$-boolean valuations u_i $(i \in I)$. For each such u_i, then we have $u_i(p) = T$ or $u_i(p \wedge q) = T$. One might think that since in either case we have $u_i(p) = T$, we may after all conclude that $v(p) = T$, which would show that $(A_1^\vdash)$ was correct for $\vdash = \vdash_{\mathsf{CD}}$. But in fact there is no justification for thinking that in the case in which $u_i(p \wedge q) = T$, we have $u_i(p) = T$. That would follow if u_i were $\wedge$-boolean, but it is v, not the various u_i whose conjunctive combination coincides with v, which we are given is $\wedge$-boolean. The simpler case—with fewer sentence letters, that is—in which the "q" in $(A_1^\vdash)$ is replaced by another occurrence of "p" (*cf.* the one-variable formulas ψ_i from Sect. 4) fails on the same grounds: the $\vee$-boolean valuations u_i discern no connection between the formulas p and $p \wedge p$. The more complex (3-variable) case of $(D_1^\dashv)$:

$$(p \wedge q) \vee (p \wedge r) \vdash p \wedge (q \vee r),$$

meets a similar fate. (Recall that in Sect. 2 we could not use $\vdash_{\mathsf{mod}}$ to show $\vdash_{\mathsf{CD}}$ failed to satisfy this condition, which was settled instead by appeal to a matrix argument in Sect. 3.) In fact, as the reader may have surmised, the $\vdash_{\mathsf{mod}}$ was inspired by the considerations of the present section, thinking of the worlds in the Kripke semantics as playing

[41] The first claim under Fact 4, that is. The second—parenthetical—claim is included as a semantic gloss on the left-primeness of $\vdash_\vee$: suppose $\gamma \nvdash_\vee \varphi$ for each $\gamma \in \Gamma$. Thus we have $\vee$-boolean valuations (one for each γ) v_γ with $v_\gamma(\gamma) = T$, $v_\gamma(\varphi) = F$. According to the parenthetical part of Fact 4, the disjunctive combination of all the v_γ is itself $\vee$-boolean. But this valuation verifies each $\gamma \in \Gamma$ while falsifying φ, so $\Gamma \nvdash_\vee \varphi$.

the role of the $\vee$-boolean valuations u_i in our recent calculations, with $\Box$ attached to the τ-translations for evaluating $\vdash_{\mathsf{mod}}$-claims corresponding to the consideration of conjunctive combinations of such valuations, and the insertion of $\Box$ for the τ-translation of conjunctions corresponding to the fact that the conjunctive combination of $\wedge$-boolean valuations, unlike that of $\vee$-boolean valuations, is another such valuation (Facts 3 and 4 above). But this modal treatment did not quite match $\vdash_{\mathsf{CD}}$ because at no point did it invoke anything corresponding to valuations that failed to be $\wedge$-boolean.

No doubt there is more to be said about $\vdash_{\mathsf{CD}}$, but that is all that will be said here. Full marks to Béziau for directing attention in [4, 5], and [6], to this surprisingly unfamiliar consequence relation.

References

1. Béziau, J.-Y.: Classical negation can be expressed by one of its halves. Log. J. IGPL **7**, 145–151 (1999)
2. Béziau, J.-Y.: Rules, derived rules, permissible rules and the various types of systems of deduction. In: Pereira, L.C., Hauesler, E.H. (eds.) Pratica—Proof, Types and Categories, pp. 159–184. PUC (Pontifical Catholic University), Rio de Janeiro (1999)
3. Béziau, J.-Y.: Sequents and bivaluations. Log. Anal. **44**, 373–394 (2001)
4. Béziau, J.-Y.: A paradox in the combination of logics. In: Carnielli, W.A., Dionísio, F.M., Mateus, P. (eds.) Proceedings of Comblog'04, pp. 87–92. IST, Lisbon (2004)
5. Béziau, J.-Y., Coniglio, M.E.: Combining conjunction with disjunction. In: Prasad, B. (ed.) Proceedings of the 2nd Indian International Conference on Artificial Intelligence (IICAI 2005), Pune (India), pp. 1648–1658 (2005)
6. Béziau, J.-Y., Coniglio, M.E.: To distribute or not to distribute? Log. J. IGPL **19**, 566–583 (2010)
7. Blok, W.J., Pigozzi, D.: Algebraizable logics. Mem. Am. Math. Soc. **77**, 396 (1987)
8. Bou, F., Esteva, F., Font, J.M., Gil, A.J., Godo, L., Torrens, A., Verdú', V.: Logics preserving degrees of truth from varieties of residuated lattices. J. Log. Comput. **19**, 1031–1069 (2009). (Correction: ibid. **22**, 661–665 (2012))
9. Carnielli, W.A., Coniglio, M.E.: Bridge principles and combined reasoning. In: Müller, T., Newen, A. (eds.) Logik, Begriffe, Prinzipien des Handelns/Logic, Concepts, Principles of Action, pp. 32–48. Mentis, Paderborn (2007)
10. Carnielli, W.A., Coniglio, M.E., Gabbay, D., Gouveia, P., Sernadas, C.: Analysis and Synthesis of Logics—How to Cut and Paste Reasoning Systems. Applied Logic Series, vol. 35. Springer, Amsterdam (2008)
11. Chellas, B.F.: Modal Logic: An Introduction. Cambridge University Press, Cambridge (1980)
12. Dummett, M.: The Logical Basis of Metaphysics. Harvard University Press, Cambridge (1991)
13. Dyrda, K., Prucnal, T.: On finitely based consequence determined by a distributive lattice. Bull. Sect. Log. **9**, 60–64 (1980)
14. Font, J.M., Jansana, R.: A General Algebraic Semantics for Sentential Logics. Springer, Berlin (1996)
15. Font, J.M., Verdú, V.: Algebraic logic for classical conjunction and disjunction. Stud. Log. **50**, 391–419 (1991)
16. Font, J.M., Guzmán, F., Verdú, V.: Characterization of the reduced matrices for the $\{\wedge, \vee\}$-fragment of classical logic. Bull. Sect. Log. **20**, 124–126 (1991)
17. Gabbay, D.M.: Semantical Investigations in Heyting's Intuitionistic Logic. Reidel, Dordrecht (1981)
18. Gabbay, D.M.: Fibring Logics. Oxford University Press, Oxford (1999)
19. Gentzen, G.: Untersuchungen uber das Logische Schliessen. Math. Z. **39**, 176–210, 405–431 (1934) (English translation in The Collected Papers of Gerhard Gentzen, ed. M. Szabo, North-Holland, Amsterdam 1969)
20. Górnicka, A.: Dual consequence operations associated with a certain class of logical matrices. Bull. Sect. Log. **29**, 143–150 (2000)

21. Humberstone, L.: Homophony, validity, modality. In: Copeland, B.J. (ed.) Logic and Reality: Essays on the Legacy of A.N. Prior, pp. 215–236. Clarendon Press, Oxford (1996)
22. Humberstone, L.: Classes of valuations closed under operations Galois-dual to Boolean sentence connectives. Publ. Res. Inst. Math. Sci. **32**, 9–84 (1996)
23. Humberstone, L.: Philosophical issues in many-valued logics. In: Craig, E. (ed.) Routledge Encyclopedia of Philosophy, vol. 6, pp. 84–91. Routledge, London (1998)
24. Humberstone, L.: False though partly true—an experiment in logic. J. Philos. Log. **32**, 613–665 (2003)
25. Humberstone, L.: Béziau's translation paradox. Theoria **71**, 138–181 (2005)
26. Humberstone, L.: Contrariety and subcontrariety: the anatomy of negation (with special reference to an example of J.-Y. Béziau). Theoria **71**, 241–262 (2005)
27. Humberstone, L.: Extensions of intuitionistic logic without the deduction theorem. Rep. Math. Log. **40**, 45–82 (2006)
28. Humberstone, L.: The Connectives. MIT Press, Cambridge (2011)
29. Humberstone, L.: Replacement in logic. J. Philos. Log. **42**, 49–89 (2013)
30. Humberstone, L.: Aggregation and idempotence. Rev. Symb. Log. **6**, 680–708 (2013)
31. Raftery, J.G.: Order algebraizable logics. Ann. Pure Appl. Log. **164**, 251–283 (2013)
32. Rautenberg, W.: 2-element matrices. Stud. Log. **40**, 315–353 (1981)
33. Shoesmith, D.J., Smiley, T.: Multiple-Conclusion Logic. Cambridge University Press, Cambridge (1978)
34. Smiley, T.: The independence of connectives. J. Symb. Log. **27**, 426–436 (1962)
35. Steinberger, F.: Why conclusions should remain single. J. Philos. Log. **40**, 333–355 (2011)

L. Humberstone (✉)
Department of Philosophy, Monash University, Wellington Rd., Clayton, Victoria 3800, Australia
e-mail: lloyd.humberstone@monash.edu

Universal Logic or Logics in Resemblance Families

Dale Jacquette

Abstract It is a momentous and as yet unsolved, perhaps unsolvable, question in the philosophy of logic, as to whether there is a single universal logic. The alternative is to maintain that there are only fundamentally distinct logics, some similar to others in some but not other ways, and each reflecting another logical dimension of what for convenience can be considered an island with particular dependencies in a sea of logics. The first question this essay considers, on the event of honoring Jean-Yves Beziau for his accomplishments and contributions to the program of Universal Logic, is whether there can be a universal logic, or whether a family resemblance model for overlapping different kinds of logically irreducible similarities between putatively disparate formal logical languages provides a more plausible and explanatorily fruitful model for understanding the proliferation of logics, especially since the formalization of modal and nonclassical systems. The second question is whether it makes any difference to Beziau's Universal Logic whether there can really be a universal logic in the sense prescribed. Here the conclusion is that Beziau's Universal Logic research program is unaffected by the unattainability of a universal logic, construed either as an ideal of reasoning or ideal theory of reasoning. Beziau's explanations of what he means by 'universal logic' are sampled from both the Preface to his 2005 edited volume *Logica universalis: towards a general theory of logic*, and his 2014 *Synthese* essay, The relativity and universality of logic. The concept of universal logic and Universal Logic are critically evaluated, with the consequence that an alternative and in many ways preferable family resemblance model of similarities of different kinds selecting different logics by virtue of different partially overlapping shared properties is not seriously challenged by Beziau's defense of logical universalism. It is one thing to recognize that reasoning is in some sense unitary, whereas theories about reasoning are legion. It is another thing to ask why there are so many logics, and consider that the reason may be that reasoning itself, though in some sense unitary, has as many different logical dimensions as there are philosophically motivated formal systems of logic. If reasoning has the loose unity of a family rather than the tight unity of a single abstract universal entity or actual dynamic psychological occurrence, then to capture the expressive and inferential structures of a selected part of thought and discourse in the entertainment and expression of which requires its own particular kind of logical reasoning.

Keywords Beziau, Jean-Yves · Family resemblance predications · Logic · Relative logic · Universal Logic · Universal logic · Wittgenstein, Ludwig

Mathematics Subject Classification 03A05 · 033B · 03C62

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_13

1 Metalogical Models

Anyone who considers the proliferation of formal logical languages since the turn of the previous century cannot fail to wonder whether all of these systems are aspects of or participants in a single archetypal universal logic, or whether they are only related one to another in the way that Wittgenstein in *Philosophical Investigations* [17, §§66–67], characterizes explicitly as non-universalist family resemblance relations. The question here, responding to Jean-Yves Beziau's impressive program in Universal Logic, concerns the underlying rationale for thinking that there is or could be *a universal logic*. If this is just a name for an ideal that underwrites Jean-Yves' work in logic, then while acknowledging his many contributions, we cannot proceed in a program of Universal Logic unless there could be a universal logic.

What is wanted is a good reason to expect that there can be a universal logic, or to prefer or disprefer the contrary metalogical characterization, that there is instead at most a family of mutually irreducible logics. Mutually irreducible distinct logical formalizations and formal and informal metalogics, often reflecting very different philosophical and methodological presuppositions, would then be as deserving of the ostensibly universal name 'logic', as Wittgenstein would allow in the case of the ostensibly universal term 'game'. The purpose of this essay is to consider both sides of this intrinsically interesting question as fairly and open-mindedly as possible, letting the chips fall where they may. At the end of the process, to preview, the conclusion will be reached that there is better reason to regard the blossoming of formal logics since the late 19th century as related by Wittgensteinian family resemblances, rather than to hold that there could be a single abstract Platonic archetypal or other ideal universal logic. The consequences for Beziau's program in Universal Logic may be nugatory nonetheless, and an argument is made for the survival and thriving of good work in Universal Logic, even without the possibility of there existing a universal logic.

2 Beziau's Explanation of Universal Logic

We first consider how Beziau describes the concept of universal logic in a relatively early source, the Preface to his 2005 edited volume, *Logica Universalis: Towards a General Theory of Logic*:

> *Logica Universalis* (or *Universal Logic*, *Logique Universelle*, *Universelle Logik*, in vernacular languages) is not a new logic, but a general theory of logics, considered as mathematical structures. The name was introduced about ten years ago, but the subject is as old as the beginning of modern logic: Alfred Tarski and other Polish logicians such as Adolf Lindenbaum developed a general theory of logics at the end of the 1920s based on consequence operations and logical matrices. Talking about the papers of Tarski dealing with this topic, John Etchemendy says: "What is most striking about these early papers, especially against their historical backdrop, is the extraordinary generality and abstractness of the perspective adopted" [6, p. vii].

Beziau refers to Etchemendy [14]. Etchemendy might be pointing in the right direction, as far as Beziau is concerned, but he reasonably explains that 'extraordinary generality and abstractness' does not yet amount to universality among the logical formalisms that need to be considered, if a universal logic is to be possible. It is around this core idea that Beziau has organized his expansive program of Universal Logic studies, with many different supported applications of rigorous logical and mathematical research. What, then, and once again, is Beziau's idea of Universal Logic?

There is something charmingly retro about the idea of a universal logic. Leibniz spoke of a *Characteristica universalis*, but that was in the halcyon days before any single such formalism was actually developed, and long before the sprawling multiplication of *characteristicae* visited upon the formal symbolic (mathematical) logical literature, after the mid-20th century. Beziau remarks in his [11, p. 8]: 'This approach is different from the later idea of Leibniz of a characteristic universalis, a universal language which, associated with a calculus ratiocinator, would be a machine more powerful that [than] human thought. As people have rightly pointed out, Leibniz was anticipating computation'.

The ideal undeniably has an intuitive appeal. Why should there not be a universal logic, *the* logic that transcends all of the particular apparently disparate developments in formal symbolic logic? Universal logic would then describe what is common to all logical formalisms, deservedly so-called. The relation in question might meaningfully be compared with Wittgenstein's *Tractatus* 3.326 distinction [16], between perceivable conventional *Zeichen* [*sign*] and corresponding transcendental *Symbol*. Different symbolic logics are analogized with different colloquial languages, by virtue of comparable expressive differences reflected in their respective grammars. The individual formal symbolic logics scattered thick on the ground are transcended by a single archetypal universal logic, in consequence of which all distinct formalisms are rightly called logics.

Beziau takes note of the conspicuous multiplication of logical systems, beginning from the timeframe he mentions, near the beginning of the 20th century:

> Logics were proliferating: each day a new logic was born. By the mid eighties, there were more logics on earth than atoms in the universe. People began to develop general tools for a systematic study of this huge amount of logics, trying to put some order in this chaotic multiplicity. Old tools such as consequence operations, logical matrices, sequent calculus, Kripke structures, were revived and reshaped to meet this new goal. For example sequent calculus was the unifying instrument for substructural logics. New powerful tools were also activated, such as labeled deductive systems by Dov Gabbay [11, p. vii].

There is more than hinted exaggeration in Beziau's thumbnail portrait of the times. His main point is nevertheless clear and reasonable. What follows, however, is Beziau's more questionable interpretation of the response that the explosion of logical systems in this era occasioned among several logicians singled out for mention. Beziau now adds:

> Amazingly, many different people in many different places around the world, quite independently, started to work in this new perspective of a general theory of logics, writing different monographs, each one presenting his own way to treat the problem [11, p. vii].

Beziau refers in particular to Norman Martin, Newton da Costa, John Cleave, Arnold Koslow, Ryszard Wojcicki, and Dov Gabbay. The question is whether these authors understood their research as contributing to a universal logic in anything like Beziau's sense, or conceived of themselves as adding in new preferred ways to the rife proliferation of inherently mutually logically irreducible distinct logical systems. Suppose that these logicians had precursor inklings of Beziau's program for a Universal Logic. The philosophical question that remains is whether the imaginary ideal logicians are correct to think of their work in logic as steps along the way toward a universal logic, rather than as new additions to the greater expanding family of logics. Beziau concludes in this part of his Preface: 'Through all these publications, the generality and abstractness of Tarski's early work was being recovered. It is surrounded by this atmosphere that I was doing my PhD and that I coined in the middle of a winter in Poland the expression "universal logic", by analogy to the expression "universal algebra" ' [11, p. viii].

Whatever the merits of Tarski's contributions to mathematical logic, whatever our deservedly high regard for his work, Tarski develops only one kind of logic or family of logics and formal semantics, and not a system of universal logic. Tarski's logic could only attain that status by default if, unlike all other logics, it alone were to survive criticism on logical or philosophical grounds, if there were good reasons for outright rejecting every other logic that does not conform to the Tarskian model. This appears very much not to be the case, in a marketplace of ideas that considers Tarski as important, and perhaps even supreme, without eliminating all competition from the choice of logic as specific mode of reasoning and formal theory of exactly that mode of reasoning.

There are rules for Tarski's logic, as there are for any logic, and the rules permit some things and forbid others, just as there are other logics that permit what Tarski's logical rules forbid, and forbid what Tarski's logical rules permit. Given these facts about the multiplicity of formal logical systems, however impressive Tarski's logic, we could not consider Tarski's a *universal* logic, if it did not triumph over all alternative formalism in all categories of theoretical modeling in practical application to all aspects of enlanguaged thought and discourse. If we ever need to rely on classical logic and at other times on dialethic or paraconsistent logic, intuitionistic logic, gap-valued or many-valued, and if these axiom systems are distinct logics that cannot be reconciled to one another's assumed truth, then the prospects for identifying a universal logic that underlies or transcends all of these seemingly diametrically opposed logics are not encouraged. These are family resemblance differences among logical formalisms, that, precisely because the systems at issue are *logics*, are certain to be different from one another in more fundamental ways than family resemblances among games or shades of red.

3 Logic in the Universal Ideal

The image of a single, essential, underlying logic is enormously appealing. Leibniz's programmatic references to a *Characteristica universalis* have resonated historically with the intuitions and ambitions of many logicians. Ironically, the dispersion of classical and nonclassical logics can be understood as the consequence of individual logicians, each trying in different ways to arrive at the best logic, at least for certain formalization and inference purposes.

Each different distinct formal system of logic can be understood as a quest to arrive at the single correct and fully comprehensive logic, with mechanical decision methods and deductively valid rule-governed inferential derivation structures. What would be more remarkable, and the strongest imaginable empirical support for a widely shared concept of universal logic, would happen instead, if so many different logicians active in the wake of the great innovations of the late nineteenth century had largely agreed and converged on a common logic. Since just the opposite has occurred, the question persists whether there can be a universal logic, or whether there can be nothing more generalizable than a *family* of distinct overlappingly resemblant logics that are mutually logically irreducible by any single underlying or transcendent all-embracing common logic.

We shall not expect Beziau to say that Universal Logic is just the banner of his brand and way of doing logic that he thinks deserves to be regarded as the rightful heir to Leibniz's *Characteristica universalis*, Tarskian in some laudable sense. If there are other non-Tarskian logics, and if Beziau's Universal Logic does not provide a universal logic in the sense of transcending or underwriting all other distinct kinds of legitimately so-called logics, then there can be no logical basis for a universal logic underwriting Beziau's branded Universal Logic or *Logica universalis*. We are disappointed if there is no universal logic that is not *the* logic of both classical and paraconsistent or dialethic, or bivalent and gapped or >2 many-valued logics, monotonic versus nonmonotonic, to mention just a few obvious examples along the spectra of formalisms in many-dimensionally rhizome-branching configurations that the proliferation of formal logics, of which Beziau rightly speaks, has assumed in fractal aetiolation, from beneath the surface of which all other logics have emerged as outgrowths.

Universal logic is not supposed to be a logic that you can merely apply to everything, to every proposition in a language. That ideal is already achieved in Aristotelian categorical syllogistic logic, and in the subsyllogistic logical relations menomically represented in the categorical square of opposition. We are properly dissatisfied for similar reasons if Beziau's Universal Logic is merely the Tarskian logic that Beziau prefers, or the logic that seems to underwrite perhaps many, but still not all other, logics. If there is a universal logic, to speak Hegelianese, it must be common to all Tarskian and all non-Tarskian logics, which implies once again that a universal logic cannot be either Tarskian or non-Tarskian, and in particular it cannot be Tarskian.

The first issue that needs to be addressed, in order to decide whether logic can be universal, is the relation between logic and philosophy. I think that while philosophy of logic is manifestly a branch of philosophy, logic plays a complex role both in and out of the general study of philosophy, and more especially in the philosophy of logic. It is standard to treat logic either as ancillary or foundational to philosophy. After all, philosophy involves reasoning and argument at every level, and logic is in some sense the theory of good reasoning, and especially of logical inference in argument. We need logical principles for proper inference in argument contexts in philosophy. While this is true, we equally need logic in every other discipline, all of which in one way or another also depend on correct reasoning in order to advance their conclusions. Such correct reasoning is not all necessarily governed by the same logic, nor by multiple outgrowths of a single nurturing underground *Urlogik*. Thought might draw upon different kinds of family resemblance related logics for different tasks, just as a tradesperson seldom has a single universal do-all tool.

The generality of 'logic' suggests its standing not only outside but above all other fields, including philosophy and the philosophy of logic. To a certain extent, this categorization seems obviously right. The fact that logic is required to serve this role undoubtedly contributes, moreover, to the force of the idea of a universal logic. There is one logic, just as there is one universal human ideal of reasoning. Like other ideals, that of universal logic is shattered upon contact with the rich diversity of human logical reasoning itself, and the assumptions needed in order to adequately formalize the expression of and inference among the relevant propositions. Where history, law, the special sciences and everyday reasoning are concerned, wherever judgment is reached by argument and justification, rather than unsupported intuition, revelation or authority, logic is needed for the acceptance of inferential connections among propositions. Wherever, one might say, propositions and the relations of their truth values and the inclusion of objects in or their exclusion from predicate extensions provide the basis for thinking, logic is present and correct logical principles of reasoning are at work. *That* logic of reasoning is the universal logic thematized by Beziau's Universal Logic. The situation in philosophy of logic is inherently more complex, because in that exact application, we are using logic to reason about logic. This necessity can easily be made to appear as bootstrapping, paradoxical in a self-referential or self-justifying way, or as an instance of hermeneutic circularity, of an innocuous or vicious type, depending again on one's point of view.

4 Universal and Family Resemblance Metalogic

The dilemmas that arise in considering the relation of classical to nonclassical logics are typified by the following. If we are unhappy with classical logic CL, then we may want to replace it with a nonclassical logic NCL1, NCL2, etc. If we rely intuitively on CL to support the development of any nonclassical logics, then it must seem that, of the two, CL is the real, fundamental, or more deeply underlying logic. The undergirding of NCL1 by CL, or in support of NCL2, can be seen as showing that nonclassical logics are not indispensable to all reasoning, but are in some way derivative, and that NCL1, NCL2, etc. are not as basic to reasoning generally as CL.

There is a great difference between supplementing CL with notations for modalities, times, places, persons, actions, or for imperatives and interrogatives, and many other things whose logical structures are interesting or philosophically essential to consider, and replacing CL with an NCL that is logically inconsistent with CL, as when NCL logically implies for all propositions p, that p and not-p is true. If we try to use a nonclassical logic NCL2 in support of the original nonclassical logic NCL1, then there is no independent logical ground for the nonclassical formalism. In effect, there is no dialectical motivation for turning away from rather than supplementing CL. It is simply presupposed from the outset, rather than established by reasoning in response to problems, paradoxes, and the other sorts of cognitive challenges for which a logical language is usually invented or devised, that CL is too limited for expressive purposes, and needs to be replaced. The possibilities of building on CL are often overlooked nowadays, when the trend is more toward replacing CL at the drop of a hat with something interestingly extravagant.

To argue in favor of CL by means of CL is viciously circular, and to argue in favor CL by means of an NCL is unsound. To develop an NCL1 from CL is not only unsound but deductively invalid, whereas to develop NCL1 from NCL1 itself is as viciously circular as to argue in favor of CL by means of CL. To develop NCL1 from another NCL2 only pushes the problem back one further step, as we ask how NCL2 is uniquely described and its expressive and inferential advantages justified. How does NCL2 relate to CL in supporting NCL1?

We cannot proceed in philosophy of logic without good answers to these fundamental theory-building metalogical metaphilosophical questions, necessarily involving or relating a new NCL formalism to CL. CL is always available for building rather than razing and replacing. The choice is that of theorizers, theory-builders, in the philosophy of logic. If we supplant CL out of more than idle curiosity, with some NCL, then we must try to justify the decision in terms of what we believe we need our best logics to do. The formalization of propositional expression and the licensing of inferential relations primarily involves indefinitely large sets of semantically pre-evaluated propositions, sentential or abstract, under any combination of univocally assigned truth values. Are we to preserve CL and set it alongside all the NCLs? Or do we demolish CL in order to make room for preferred NCLs? The philosophy of logic is called upon precisely to answer these kinds of fundamental questions that may be vital to symbolic logic's development and its philosophical approval rating.

5 Metalogical Dilemmas for Universal Logic

What, then, to do? How do we surmount this fine-honed two-edged objection? We encounter here a difficulty in the philosophy of logic that has seldom been remarked or investigated, let alone resolved. The existence of the problem already marks a significant difference in the relation between logic and other extra-logical, and even extra-philosophical disciplines, as compared with the relation between one logic and another, or between logic and philosophy, with logic considered as a part, or even as the most fundamental underlying foundation, of philosophy.

It is not good enough, when trying to avoid the dilemma, merely to say that logic is *discovered* rather than *invented*. That distinction also touches deeply on one of the most important and troublesome choices in the philosophy of logic. It encourages logicians to conceptualize the purpose of logical formalization in abstract terms either as altogether independent of natural reasoning, or as regimenting the reasoning patterns that actually prevail at the root of all inference. Whatever its ultimate purpose, one can say that if a formal language does not map satisfactorily onto a large body of meaningful discourse, then we should not first think that there is anything wrong with the meanings that language and other made objects are used to express. We should rather look first to the logic, and inquire into its adequacy. It is by such pressures laid upon a classical logic that variant formalisms have developed in response to particular analytical challenges. It is conceptual pruning that keeps the thriving branches of formal symbolic logic vibrant and fruitful.

Complete with syntax, formal semantics, axioms and inference rules, or their natural deduction or graphic equivalents, CL and any NCL must earn their keep by providing satisfactory logical structures for the symbolization of individual logically complex expressions, and for what are considered to be deductively correct inference structures. The proposal is to comprehend CL and NCLs alike. We may take the meaning of any expression to be the propositional content intended by a subject in predicating a property of an object. Meaning in this sense is reflected in a clear understanding of the propositional content of any predication of any property to any object. Regardless of how thought and a natural language or logical notation expresses predications, we speak in this connection specifically of propositional meaning, where the meaning of an expression can in principle be correctly cashed out as a truth-professing proposition.

The same is true without prejudice, when logical reasoning is turned inwardly upon logic itself, just as when logic is applied in the sense of being satisfactorily mapped upto any other subject. The logic of logic, the logic that underwrites a given exposition of logic or theory or philosophy of logic, is itself not yet philosophy of logic. A philosophy of logic must say something contentful concerning the nature of logic and logical relations. And there are many things to say. Logic, even the logic of a philosophy of logic, does not recommend reference domain comprehension principles for a first-order CL (or

NCL). Satisfactory semantic connections beyond possible one-one correspondences between logic and meaningful expressions can obviously only be arrived at through a course of philosophical inquiry, as a specifically philosophical, rather than purely logical, conclusion. The same is true more generally of philosophy, assuming logic or philosophy of logic to be some kind of part or component of a complete broadly scientific philosophy. We test logic in applications other than logic first. When a formalism comes up shining, we may come to trust it also in thinking about logic itself, as just another application in its broad range to draw upon in developing, among other things, a philosophy of logic.

Such a proposal, promising as it may at first appear, even if construed as a last straw for the philosophy of logic, cannot possibly succeed. Systematized as a formal language or not, the conflict posed by the dilemma at bottom concerns *any* logic, regardless of how it is instantiated or expressed, required for whatever reasoning is supposed to govern thinking about logic. This includes the logic in question, insofar as it is involved in developing a philosophy *of* logic. The problem, so fundamental to logic, philosophy, and the philosophy of logic, is stubbornly persistent and not easily or obviously resolved, if in the end it is solvable satisfactorily at all. One possible answer is to suppose that logic is not a part of philosophy, but in some sense external, perhaps ancillary to or otherwise excluded from it. If so, then we can speak without circularity of logic as reflected theoretically in all disciplines, even in thinking about logic. We can choose in a principled way from among alternative preferred formalizations of logic, on the one hand, and, on the other, the applications of a preferred logic in other disciplines, including philosophy, more generally, and philosophy of logic in particular. Logic is then its own logically independent thing. It is foundational to every discipline including philosophy, mathematics, and hence, *a fortiori*, to philosophy of logic and philosophy of mathematics, not to mention philosophy of language. Could this so-described logic be something like the universal logic needed to give legs to Beziau's Universal Logic?

6 Implications for Beziau's Universal Logic Program

For this reason alone, the logicist program of trying rigorously to reduce mathematics to principles of pure logic should have been understood from the outset as hopeless. The task was misconceived, and logicism developed as a reduction of elementary mathematics to the principles of specific applied logics.

Mathematical concepts, unfortunately, are found among the arid logics available in formalizations. They are only applied logical terminologies that are as close to pure logic as $F = ma$ in applied mathematics of kinematics is to $1 + 1 = 2$ in pure elementary arithmetic. Later knockdown objections to logicism, such as Gödel's incompleteness proof, and Church's and Rosser's generalizations, are best interpreted as corroboration for the logicist effort that was logically, conceptually, and cognitively doomed to failure anyway, and whose undertaking would have been most wisely avoided in the first place, except perhaps as a philosophical experiment confirming exactly the predictable theoretical calamity.

Logic, then, is not systematically part of philosophy, even though the study of logic and its applications may be of special philosophical interest, theoretical importance, and practical utility. The study of logic contributes indirectly, in somewhat the way that empirical discoveries in the natural sciences often nourish and positively influence philosophical reflection in all fields. The fact that we speak intelligibly of philosophy of logic, as we do of philosophy of mathematics and philosophy of science, where mathematics and science are not generally considered part of philosophy, further confirms the intuitive pre-analytical distinction. Whereas, in contrast, we do not typically recognize as legitimate accepted categories, "The Philosophy of Ethics", "The Philosophy of Metaphysics", or "The Philosophy of Epistemology", as opposed to simply writings on ethics, metaphysics, or epistemology, as subdisciplines of philosophy. At least I have never seen a book or professional essay with such titles. We should accordingly already be alerted to the fact that logic is not part of philosophy by the fact, in contrast, that we *do* speak of philosophy of logic, just as we speak of the philosophy of language, philosophy of mathematics, philosophy of action, philosophy of literature, and the like.

The first principle in a philosophy of logic on which to agree might therefore be that logic is vitally important to, but not strictly a part of, philosophy. Logic must be independent of philosophy in order for there to be a noncircular philosophy of logic. An adequate philosophy of logic investigates all philosophically interesting aspects of logic, and in the process presupposes and puts to use specific syntactically formalizable expressive and inferential logical principles. With anything less substantial on board, there is no logic in the proper sense of the word about which to philosophize in the first place. How to pull off this balancing act of using logic to develop a philosophy of logic, without slipping into vicious circularity or deductive fallacy, is one of the main challenges for a methodologically circumspect philosophy of logic.

The best solution is (perhaps) first to develop as ontically neutral a base logic as possible, as an extension of CL, which might be considered an NCL. Second, to minimize the use and hence the ontic impact of CL or its specific NCL extension in expressing the arguments necessary to support all the usual commitments, or their opposites, of an independently appealing philosophy of logic. The less we rely on a specific logic, CL or NCL, to make progress in the philosophy of logic, the less likely we are to smuggle specific logical preconceptions surreptitiously into our most general thinking about the nature of logic, its expressive and inferential scope and limits.

The previously mentioned dilemma surfaces again in another way. If there is no single archetypal universal logic, no *universal language and formal system of logic*, in what seems to be Beziau's required sense, then the project of Universal Logic would appear to be seriously compromised, reduced to sloganeering about an unrealizable ideal. The ideal of Universal Logic would then be ineluctably unattainable. It would not describe a practically realizable ideal. If there *is* a universal logic in the universal language and formal system sense, then, whatever that sense turns out to be, it will surely be once again just another logic.

How, we must find ourselves asking with Wittgenstein, in [17, §§66–67], can there be anything common and essential to all the different logics available for modeling many specific different kinds of reasoning. As Wittgenstein remarks in the last sentence of [17, §66]: 'And the result of this examination is: we see a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail.' That is an inquiry in the philosophy of logic toward which the present efforts at inquiry are directed. The question itself is not sophisticated, which is taken as a virtue, and the answer, when it comes into focus, can only be either that what pass as distinct logical formalisms are all expressions of a perhaps deeper or more transcendent universal logic, or they are just similar kinds of formal languages with expressive and inferential rules and capabilities that have been designed for a variety of different purposes, with no single underlying logic to be unearthed in any or all of them taken collectively.

7 Beziau's Advanced Case for Universal Logic

In his recent [11] essay, 'The Relativity and Universality of Logic', Beziau distinguishes between relative and universal logics. He asks a particular question about one limited direction of opposition to a universal logic. Even if Beziau shows that the challenge he considers does not pose a decisive refutation of the possibility of a universal logic, it does not follow that Beziau thereby considers the most dangerous threat to the universal logic promised by a program of Universal Logic. Beziau writes:

> Is logic relative? Is logic universal? These are fundamental questions regarding the very nature and meaning of logic. To properly answer these questions it is important to make the distinction between logic as reasoning and logic as the theory of reasoning. This is a distinction that is very easy to understand but most often not explicitly made [11, p. 1].

Beziau is right that whether logic is relative or universal is a fundamental question in the philosophy of logic, but the distinction he then hoists into position between logic as reasoning and as the theory of reasoning is insufficiently explained, and relations between reasoning and the theory of reasoning are not considered. If there is no theory of a putative kind of reasoning, what reason do we have to think that the reasoning itself exists?

Can there be reasoning without a possible theory of reasoning? More emphatically, if there is a logical theory of some kind or aspect of reasoning, then the existence of the theory itself testifies adequately to the existence of precisely that corresponding kind of reasoning, at least on the part of the logician, later thematized and theorized, but originally independently of any developed theory of precisely that corresponding kind of reasoning. That there is a difference between reasoning and theory of reasoning, few would deny, although reasoning about the theory of reasoning is presumably theoretical in content. If these two lines of argument are collected, then there is a fact about some moment, part or aspect of reasoning if and only if there is a corresponding fact about the theory of exactly that species or specimen of reasoning. The question we shall want to ask, in working through Beziau's [11] clarification of the concept of Universal Logic, honoring the distinction between reasoning and the theory of reasoning, as far as reasoning and its best theory will permit, is whether the distinction between reasoning and theory of reasoning can possibly provide the lifeboat to which Beziau's Universal Logic might want desperately to cling.

If we are going to say anything about reasoning at all we shall be spouting theory, in at least a loose sense of the word. Otherwise, we must follow Wittgenstein's advice in *Tractatus* 7 [16]. There is reasoning and there is theory of reasoning, fair enough. There cannot fail to be an important structural feature of one that is not found or reflected in the other, and where their fundamental logical structures are our concern, we can as well direct attention to reasoning itself as to an ideal theory of reasoning. What good does it do to declare that we concern ourselves directly with reasoning and not the theory of reasoning, as we proceed immediately thereafter to articulate some of the propositions belonging to nothing other than a theory of reasoning?

Beziau diagnoses the source of relativism in philosophy of logic as a cause rather than a symptom or consequence of 'extreme' relativism. It sounds like radical politics or an advanced state of a disease. Certainly nothing good. Beziau strings together some remarkable claims, as he takes the reader from Aristotle to Kant to what he considers the present day's chaos in logical theory, always favoring the attractions of relativism and away from universalism in the philosophy of logic:

> Contemporary logic with its chaotic multiplicity of systems of logic seems to promote extreme relativism, but in the middle of this chaos there are still some individuals claiming that a given logic is the true one, in the sense that a particular system of logic is a true description of the true way of reasoning similarly to what Kant was claiming about Aristotle's logic more than two centuries ago in the preface of the second edition of the *Critique of Pure Reason* (1787) [11, pp. 2–3].

If I rightly understand Beziau's explicit statement, then he maintains that the fact of individuals *claiming* that *their* system of logic is *the true one*, is enough to hold alight the torch of *Universal Logic*. The problem is that when these logics are compared, first in this respect and then in that, they often have enthusiastic commitment and passionate devotion to the truth or correctness of the logic in question as a theory of reasoning that they each individually but mutually contradictorily accept. Nothing whatsoever is resolved as to which if any of these logical systems is correct, let alone which if any of them could be considered universal.

Kant famously says that Aristotle had so perfected the science of logic that no significant advances in or departures from it had been made since the time of ancient Greek philosophy. Without explanation, but with George Boole fixed at the crossroads, Beziau marks the turning in logic whereby many different kinds of systems are developed, and logic, the theory of reasoning itself, is frequently changing in the formal languages sense. From Aristotle-to-post-Kantian stagnation, suddenly there are more logics than we comfortably know philosophically what to do with, an embarrassment of riches. As Beziau reports:

> After 2000 years of immobility, the theory of reasoning, since Boole, like all other sciences, is in constant transformation. And a position of objective relativism in logic, according to which there is an objective reality of reasoning, but that any theory of it is relative seems quite natural. At some point some people had the idea that classical propositional logic and 1st order logic were a perfect description of our way of reasoning. But today there are many systems of logic and this seems as absurd as considering that syllogistic is the right description of reasoning. However, there are still some people claiming that they have found the right logic. The confusion is connected with two problems: the double meaning of logic we have already mentioned and the duality descriptive/normative. Let us consider the case of intuitionistic logic [11, p. 4].

Beziau acknowledges that there are many mutually logically irreducible logics, and that there is nevertheless a positive prospect for a universal logic of reasoning. Universal logic has its chance, even if there is no universal logical *theory* of reasoning on the horizon. *If* we get logic right, *then* we have before us a logical theory of *all* reasoning, which, in Beziau's nuanced sense, is a *universal* logic. It will be a theory of reasoning that applies to any and all reasoning. Historically, moreover, it is undeniably possible that Tarski pointed one way toward such a unique universal logic of all expression. This is how Beziau introduces the multiplicity of intuitionistic logics, which cannot even find a common denominator among themselves, and certainly no universal logic underlying or transcending them all, appearances notwithstanding. Beziau begins with this frank admission:

> First, it is important to note that there is not only one intuitionistic logic, but a whole variety of intuitionistic systems of logic. Even if someone thinks that intuitionistic Logic is the one true essence of mathematical reasoning, he may admit that there are various approximate theories of it. And there is another point: Brouwer claimed that intuitionistic Logic is the way we should reason, we should not follow classical Logic (see van Stigt 1990). There is something normative in this perspective that we don't have in physics, since it would be difficult to say that the universe should behave in such or such way [11, p. 4].

What is disconcerting about Beziau's forthright remarks, is that he openly acknowledges that there are multiple intuitionistic logics that cannot even be reconciled amongst themselves in terms of their distinct axiomatic contents, let alone externally with the axioms of classical non-intuitionistic logic, or any of the other multiple family members of a wide-spreading network of logical formalisms. Beziau admits all these facts while persistently maintaining that there must nevertheless be a universal logic, and just as persistently resisting offering any example that could seriously be considered as both universal and a logic.

Nor, if we accept Beziau's distinction between reasoning and theory of reasoning, do we have any compelling reason to suppose that there can be a universal logic for all natural reasoning, as opposed to that of any theory of natural reasoning, if natural reasoning from an ideal reasoner's phenomenological perspective includes all the kinds of reasoning that are modeled in different classical and nonclassical systems of logic, as different theories of reasoning. If someone can come up with a system of logic, a theory of reasoning, it means that thought and reasoning can be like that. If it is not, more often than among a handful of logicians, it is only because of accidental anthropological extra-logical factors. Under the right imaginable environmental pressures, theory could describe thought as exactly like that modeled in paraconsistent inference, or intuitionist of one type or another, or nonmonotonic doxastic logic, since everything in theory needs to be believed, or, etc.

Beziau offers yet another concept of universal logic, like moss-covered pathways in a Zen-garden. He develops an idea of Gary Birkhoff's [13] for universal algebra in order to explain what awaits at the end of the walk:

> What Birkhoff developed is not an axiomatic basis but a conceptual basis resting on abstract algebras as structures and morphisms. Birkhoff's conceptual foundation is not for all mathematics but for all algebraic structures, but it can be naturally extended and generalized to all mathematics in various ways (cf. Bourbaki, Category theory, Model Theory) including logic. That is the way to universal logic. Like universal algebra, universal logic can be seen as a unifying theory: a theory unifying the universe of the multiplicity of logics, unification based on a conceptual basis rather than on an axiomatic one (cf. Beziau 2010 [9]). A good title for a book presenting a general study of the different systems of logic could be The world of possible logics... [11, pp. 6–7].

Beziau is divided in his loyalty to Tarski as the founding spirit of Universal Logic, on the one hand, and his desire to set theory free from theory, and put it back in touch with its proper subject, which seems a hard go on Tarskian logical principles and methodology. With Tarski we are at once in the land of object and meta-theory, meta-theory and meta-meta-theory, and so on, merely to safeguard elementary propositional logic and truth functional semantics from the self-non-applicational treachery of the liar paradox. Since to speak of reasoning at all is to contribute to the theory of reasoning, what is wanted is not an iterative theory of the theory of reasoning, but simply a correct theory of reasoning.

Fair enough. Except that it is exceedingly hard to find anyone who is actually offering an iterative theory of the theory of reasoning, rather than precisely what each takes to be the best, even the only or only rational, preferred, theory of reasoning. That is what they think they are doing, and that is what they are actually doing, although recognizing these facts goes no distance at all toward deciding which theory of reasoning is more likely to be true, once theirs is thrown into the bull-pen with all the others. Beziau acknowledges the need to disregard Tarski's logical axioms, while not denying that Tarski has a logic. Such admissions already place enormous pressure on the idea that there can be a universal logic. Tarski thereby assumes only a ceremonial role in the ongoing work of Universal Logic. It was Tarski, the tablets on the statues will someday say, who had a similar idea for something like a universal logic. That he did not get things right does not detract from the importance of his early sense of what a universal logic could and should be like.

We are right back to our miserable starting place, pondering as before whether there is a single all-purpose all-occasion universal logic, transcending or below the root of more specialized logics that do not necessarily get along amicably together on the garden's surface. Beziau repeatedly projects a universal logic for Universal Logic, but does not deliver the system. It would need to be a theory of reasoning, if Beziau had it on offer, and not an iterative theory of a theory of reasoning. That is all in its favor, if Beziau's distinction is observed, but the same can be said of many different theories of reasoning, formal systems of logical expression and inference that on the same grounds could equally deserve to be called universal, and identified as the universal logic, although they are in other ways anti-Tarskian, the polar opposites ideologically of Tarski's logic and philosophy of logic. It is not unexpected that such philosophical oppositions should exist.

The only urgent question, for Beziau's logical universalism, is whether Tarski got it right. And that is where, to put it mildly, opinions begin to diverge. If everyone with a stake in the outcome would like to have their preferred logical formalism, granted its virtues, designated as *the universal* logic, then their collective will is obviously not going to advance the dialogue toward any more secure basis for considering any particular one of the multiple logically incompatible logics as *preferred*, and hence as *the universal logic*. Each side will wave its own flag. We have on the ground conflicts of incommensurable pros and cons, advantages and hard to compare disadvantages, among, by their own standards, alternative mutually incompatible competing logics. Everyone playing this game would like to have *their* logic thought of as *the universal logic*. Wanting and claiming do not make anything so, even as a logical possibility. If such efficacy existed, we should need to consider all these advantages countermanded by one another anyway, by wanting and claiming the opposite to be true on the part of partisans of the contrary anti-universal logic dialectical side, who may wish that a non-universal logic, relative or family resemblance, should triumph instead of any universal logic.

Beziau tries to deflect the objection that there can be no universal logic if there are logically incompatible systems of logical axioms for what appear manifestly to be different logics in a family of overlapping resemblant logics, rather than the sure sign of a universal logic deep at work in all logical reasoning. Beziau invokes the unity of 'the reality of reasoning', as over against which there can be many different theories, including logics *of* reasoning. Beziau maintains:

> As we have seen in the previous section, it is natural to consider that there are different logical systems even if we consider there is only one reality of reasoning. The variety of logical systems may reflect not only the various ways to understand reasoning at different stages of the history (syllogistic vs first-order logic) and various aspects (modalities, quantifiers, positive logic) of one real reasoning but also the variation of this one reasoning itself through different ages (children/adult) and different cultures. One may focus only on some aspects of the reality of reasoning, and then

> combine, compare, translate, etc., using general techniques developed to deal with logical structures.
>
> On the other hand universal logic leaves open the possibility to consider that reasoning itself has some fundamental variations through history of humanities and also leaves open the logic door to the world: animals, plants, the universe… [11, p. 7].

The problem is that no theorist need dispute that human reasoning is one thing. The question is rather whether one logic serves the explanatory and other formal modeling, expressive and inferential needs of all the logical structures of human reasoning in all its varied exercise. By analogy, water is also one thing, although we need different theories to account for its cause and effect relations when liquid, frozen or vapor.

Family resemblance relations among distinct logical systems is more akin to JC Beall and Greg Restall's *logical pluralism*. Beziau describes their philosophy of logic in these terms, explaining his understanding of Beall and Restall in a succession of remarks, a selection of which appear below:

> …JC Beall and Greg Restall have promoted logical pluralism… Logical pluralism can be simply understood as referring to a view according to which there are different logics: "To be a pluralist about logical consequence, you need only hold that there is more than one true logic." (Beall and Restall 2000 [1], p. 476. See also [2]). But this expression is not innocent; it refers to a position supported by a gang, the logical pluralists, against another gang the logical monists. We are facing the "ism" mania like with communism, physicalism or mysticism [11, p. 9].
>
> Logical pluralism is the claim and the defense that there are various logics, it is not a general theory of logics. The logical pluralist does not make the distinction between reasoning and the theory of reasoning; both are put in the same bag. Beall and Restall use the distinction between "Logic" and "logic" in the following way: "Logic names the discipline, and logic names a logical system." (2000, p. 475). According to this view there is no clear distinction between a logic system and the reasoning it is describing. Furthermore logic, as a discipline does not explicitly appear as a systematic study of logical systems [11, p. 9].
>
> Logical pluralism promotes plurality rather than unity. It is in the line with putting an "s" everywhere; physics, mathematics, religions, logics. The plural for physics or mathematics is indeed quite ambiguous. In English it is rather a syntactic feature [11, p. 9].
>
> From the point of view of intuitionistic logic, classical logic is wrong, the same idea has also been promoted by relevantists and all sort of constructivists, defending other alternative systems. The battle is tough: fuzzy logic has been compared to cocaine, paraconsistent logic to pornography… Some non-classical logicians are not only against classical logic they are also against other logics thinking their logic is the best. Non-classical logic is not necessarily synonymous to pluralism. [11, p. 11].
>
> At the beginning of the XXth century in Europe, people were using logic to develop the unity of science. Logic was identified to the methodology of deductive sciences (cf. Tarski 1936 [15] and the international series of events LMPS launched by Tarski). This project did not stop. Universal logic is the continuation of this project [11, p. 14]. See also [3, 7, 8, 10, 12].

It seems correct to say that family resemblance models of distinct logics is a specific kind of logical pluralism, and not the other way around. There can be pluralism without family resemblance, although it appears unlikely, and the plurality of logics certainly cluster into particular more or less easily identifiable resemblant families. There are classical and intuitionistic systems, relevance logics, paraconsistent and dialethic, and the like, among many others, all of which have spawned families of their own, and are yet related in different ways to other logics in the rhizome. Family resemblance models go beyond what logical pluralism maintains, requiring more structure via overlapping resemblances for inclusion in any number of different families of logics than logical pluralism by itself offers to provide. Should Beziau succeed at blunting objections to universal logic from the standpoint of relative logic, in the essay, he still does not present good enough reasons for preferring universal logic to a family resemblance model of distinct mutually logically irreducible logics. For each of these logics, you can almost take your pick, logical relations may be absolute, each intending a different subdomain of intended objects and singling out these or those objects by virtue of their constitutive properties.

8 Universal Logic Less Stridently Conceived

There is no universal model of deductive inference, and, in stricter diction, we should not speak in universal terms of logical truth, logical form, tautology, contradiction, and the like, without referencing the logical language in which these relations are intended. It is expected that some version of a classical logic will be the default, but an ontically flexible, even neutral referential semantic domain is needed to explain the meaning of many different kinds of propositional expression. If that is what it takes for logic to be relativistic, then a favored archetypal logic, speaking generally, could be both universal *and* relativistic, even in Beziau's sense. The conditions are met if the logic is universal by governing all logical reasoning. If we have a universal logic, it must do this. If we can graft modal logic onto classical first-order predicate logic or functional calculus, then we are in business for a while. Unfortunately, we cannot continue to annex all logical systems in an expanding cluster, if membership is based at least in part on logical compatibility with any chosen core. Some will be logically compatible, and others will certainly not.

The project is theoretical, for the sake of which we can avail ourselves of whatever alternative cluster of logical principles we would prefer to start with before adding a pet theory to what must then be a numerically different logical axiom cluster. If it is the same axiom cluster, then universal logic as Beziau's concept seems to require must be a single all-purpose logic or the backbone or transcendental ground of all single-purpose logics, must harbor explicit contradictions. If this means that universal logic must be paraconsistent or dialethic, then one wants to see Beziau make this argument. If he *has* made the argument, then do so more *loudly*, please, Jean-Yves! Beziau has done highly interesting formally exacting work in paraconsistent logics, and has worked in more traditional areas of classical logic, usually with an interesting off-beat twist of one sort or another. To the limited extent of my knowledge, Beziau has not explicitly drawn a connection between these two strands of his interests, by philosophically advocating paraconsistency as the universal logic of reasoning on which the program for Universal Logic can build. See [4, 5]. The trouble is that if it is anything like that, then loyalty to Tarski's logic which inspired Universal Logic in the first or second place is at risk. One does not lightly set down the axioms for Tarskian inferential semantics next to those for Beziau's paraconsistent or dialethic logics. Or even with those of an ontically neutral Meinongian intensionalist referential semantic domain of existent and nonexistent intended objects. Beziau has acknowledged the need to move beyond Tarski, by rejecting Tarski's axioms for logic. Something of the spirit and expectations of a Tarskian logical language and an algebra of universal inference structures underlying all specialized logics continue to inspire Beziau's project of Universal Logic. This fact may prevent Beziau from identifying the logic of Universal Logic with some development of paraconsistent inference semantics, if that is what should turn out to be the most universal, in the sense of being the most versatile, among distinct logics about the unified facts of reasoning.

We can exploit the fact that absolutism in the philosophy of logic is as much the natural dialectical opponent of relativism as it is of family resemblance-ism. Is family resemblance relative, relativistic, a species of what Beziau considers relative logic? It appears that family resemblances might, but need not be relativistic, if to be relativistic in the present context means for family resemblances to be relative to any particular common cluster within a family. The point is always to take the entire families of logics into view where one would otherwise speak of a universal logic. Family resemblances can be plausibly grouped together in a number of different ways, just because they resemble each other in different ways. It would be trivial to suggest that family resemblances are relative to the resemblances by which they constitute a family. Almost any collection of entities can have its similarities sought out and organized into a variety of different resemblance families for a number of different purposes. If a Wittgensteinian family resemblance relation, along with Wittgenstein's later loosely connected family resemblances, can be arranged in many different ways, including some and excluding other candidates from the entire pool of potential family members, then we need not imagine that in logic research we are striving toward any universal logic. We can organize the same faces in different families for similarities alternatively among any combination of facial features, eyes, ears, eyes and ears, noses, noses and ears, hairline, and so on, and we can include some and exclude others from a given choice of features, on a wide variety of reasonable grounds. In the process of grouping logics into overlapping family resemblance clusters, we may move closer to an ideal favored archetypal logic, but we do not necessarily thereby so much as begin to approach the concept of a *universal* logic. Nor can we take comfort in the thought that there must be a single extended family of logics, each more or less resembling one another, as Wittgenstein remarks, in many overlapping and criss-crossing ways. If family resemblances among distinct logics are irreducible, then there is no prospect for a *universal logic*, while Universal Logic as the quest for an algebraic possibly paraconsistent preferred archetypal logic heroically lives on.

References

1. Beall, J.C., Restall, G.: Logical pluralism. Australas. J. Philos **78**(4), 475–493 (2000)
2. Beall, J.C., Restall, G.: Logical Pluralism. Clarendon, Oxford (2006)
3. Beziau, J.-Y.: Universal logic. In: Childers, T., Majer, O. (eds.) Logica'94—Proceedings of the 8th International Symposium, Prague, pp. 73–93 (1994)
4. Beziau, J.-Y.: What is paraconsistent logic? In: Batens, D., et al. (eds.) Frontiers of Paraconsistent Logic, pp. 95–111. Research Studies Press, Baldock (2000)
5. Beziau, J.-Y.: From paraconsistent to universal logic. Sorites **12**, 5–32 (2001)
6. Beziau, J.-Y. (ed.): Logica Universalis: Towards a General Theory of Logic. Birkhäuser, Basel (2005)
7. Beziau, J.-Y.: 13 questions about universal logic. Bull. Sect. Log. **35**, 133–150 (2006)
8. Beziau, J.-Y.: Logic is not logic. Abstracta **6**, 73–102 (2010)
9. Beziau, J.-Y.: What is a logic? Towards axiomatic emptiness. Log. Invest. **16**, 272–279 (2010)
10. Beziau, J.-Y. (ed.): Universal Logic, an Anthology from Paul Hertz to Dov Gabbay. Birkhäuser, Basel (2012)
11. Beziau, J.-Y.: The relativity and universality of logic. Synthese (2014). doi:10.1007/s11229-014-0419-0 [Note that this online publication of 20 March 2014 has no numbered pages, and all page numbers to this source indicated in square brackets [n] are supplied by the present author]
12. Beziau, J.-Y., Buchsbaum, A.: Let us be antilogical: anti-classical logic as a logic. In: Moktefi, A., Moretti, A., Schang, F. (eds.) Let Us Be Logical. College Publications, London (2013)

13. Birkhoff, G.: Universal algebra. In: Comptes Rendus (ed.) Du Premier Congrès Canadien de Mathématiques, pp. 310–326. University of Toronto Press, Toronto (1946)
14. Etchemendy, J.: Tarski on truth and logical consequence. J. Symb. Log. **58**, 51–79 (1988)
15. Tarski, A.: Logic, Semantics, Meta-Mathematics, 2nd edn. Hackett, Indianapolis (1956). Trans. J.H. Woodger, 2nd ed. John Corcoran
16. Wittgenstein, L.: Tractatus Logico-Philosophicus. Routlege & Kegan Paul, London (1922). Ed. C.K. Ogden
17. Wittgenstein L.: Philosophical Investigations, 3rd edn. Basil Blackwell, Oxford (1968). Ed. and trans. G.E.M. Anscombe

D. Jacquette (✉)
Institut für Philosophie, Abteilung Logik und theoretische Philosophie, Unitobler, Universität Bern, Länggassstraße 49a, Bern 9, 3000, Switzerland
e-mail: dale.jacquette@philo.unibe.ch

Causality and Attribution in an Aristotelian Theory

Srećko Kovač

Abstract Aristotelian causal theories incorporate some philosophically important features of the concept of cause, including necessity and essential character. The proposed formalisation is restricted to one-place predicates and a finite domain of attributes (without individuals). Semantics is based on a labelled tree structure, with truth defined by means of tree paths. A relatively simple causal prefixing mechanism is defined, by means of which causes of propositions and reasoning with causes are made explicit. The distinction of causal and factual explanation are elaborated, and examples of cyclic and convergent causation are given. Soundness and completeness proofs are sketched.

Keywords Predication · Attribute · Subject · Reason · Cause · Proof · Labelled tree

Mathematics Subject Classification (2010) Primary 03A99 · Secondary 01A20

1 Introduction

The role of causality in philosophical and scientific theories, for example, in the last hundred years, ranges from its role as a primitive or fundamental concept (even in logic and set theory) to its dismission at least from "austere" science, and to its confinement to informal, ordinary discourse. What is equally important is that we do not presently have a unique concept of causality to which we may refer as to the standard one.[1]

Our starting background question is whether causality is (or should be) a primitive concept of theories. This leads to the question about causality and logic: since we assume that each theory should include (beside its specific axioms) some sort of logic (i.e. language and a sort of a consequence relation on the sentences of the language), the question arises whether the concept of causality—and in which sense of causality—is connected with logic as such ("essentially"), or originates from it, or is in any other way closely related to it. For example, in a Kantian approach, the concept of causality stems from formal logic (from the "logical function" of hypothetical judgment, in essential connection with the principle of sufficient reason) but in application to representations given in intuition. On the other hand, Gödel reflects on causality as a fundamental philosophical concept from

[1] For instance, the editors of the *Oxford Handbook of Causation* say: "Philosophers have been interested in the nature of causation for as long as there has been philosophy. . . . Despite the attention, there is still very little agreement on the most central question concerning causation: what *is* it?" [7, p. 1].

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_14

which even logic and set theory should be derived [12, p. 432–435]. Be it as it may, the interrelationship of logic, causality and knowledge seems to be one of the fundamental, open questions.[2]

The approach we take here is meant to go back to historical origins of logic and of the theory of causation, and to possibly see whether and in which way these conceptions could give some orientation in current reconsiderations of the concept and the role of causality. We propose a formalisation of Aristotelian theory of causality, with the aim to show that Aristotelian concept of causality reinforces the view on causality as a basic concept of our theories, and moreover, that the concept of causality is of intensional nature and essentially connected with logic. One special interest is to find, along Aristotelian lines, in which way and how much a causal chain leading to some event can be reduced in order to avoid unnecessary complexity and yet to retain the part of conditions that may still be called a cause of the event.

2 Causality and Attribution in Aristotle

There are some general features of the Aristotelian concept of cause:

1. A cause is something *because of what* (*dia ti*), or why, some state of affairs (thing, *pragma*) obtains (see, e.g. *An. Post.* A2 71b 10–11, A6 75a 35, A24 85b 23–24 [1, 3, 5]);
2. The causation of a state of affairs is *necessary*, in the sense that things *cannot* happen otherwise when the cause is present (e.g. *An. Post.* A2 71b 11–12);
3. The cause of a state of affairs is a *necessary* and a *sufficient* condition of the state of affairs (if the affirmation is the cause of ϕ, the negation is the cause of $\neg\phi$; cf. *An. Post.* A13 78b 17–21);
4. Cause *essentially* ("in itself") belongs to the thing it causes, i.e. the cause and the thing should be interconnected by means of *what* they are (their essences), or by means of how they are defined (i.e. by means of their respective concepts) (see *An. Post.* A24 85b 24–25, B8 93a 4).

Feature 1 distinguishes causes from mere (non-causal) reasons explaining the *fact that* something obtains; for example, non-twinkling of planets is a reason that explains *that* planets are near, but not the reason *why* it is so, whereas the reason why planets do not twinkle is that they are near (*An. Post.* A13); feature 2 distinguishes causality from what happens only *accidentally*, even if it happens always (*An. Post.* A6 75a 32–35), or otherwise may and may not happen (cf. *An. Post.* A6 75a 18–21); feature 3 distinguishes a cause from a remote or too general reason, responsible for (or merely explaining) other facts too (see below the analysis of Aristotle's example of why walls are not breathing); feature 4 distinguishes causality from merely equivalent (reciprocal) phenomena and properties (*propria*)—see examples in Sect. 4 below.

[2]These questions are addressed, for example, in a short programmatic paper by J.-Y. Béziau [8], where a very general formalised theory is envisaged that is motivated by da Costa's formalisation of the principle of sufficient reason and Shopenhauer's philosophical views on the principle.

Aristotelian cause of a property is an essential attribute (of what has the property) which is the necessary and sufficient reason why the property obtains. On the other side, for each attribution (*hyparchein*, 'to belong') it could be asked about its "why", looking for its cause in this attribution itself (in which case it is a primitive attribution) or outside the attribution (in which case it has an external cause). Aristotelian formal logic (categorical syllogistic) gives a formal account of attribution interdependences (e.g. if A is universally attributed/belongs to B, and B is universally attributed to C, then A is universally attributed to C). Accordingly, Aristotelian formal logic is nothing but a general formal theory of attribution. Moreover, according to Aristotle's definition of a syllogism, the premises are causes of the conclusion (*An. Pr.* A 1, 24b 18–22 [1, 4, 6]), i.e. forms of reasoning have themselves their formal reasons why (formal causes). Aristotle explicitly states that premises are the material cause from which the conclusion necessarily follows (*Metaph.* Δ2 1013b 20–21 [1], see also [13][3]). Finally, special Aristotelian theories, that is, special sciences, are applied logic—theories of causation in some specific area, with specific primary propositions added to a general formal logic.

3 Preliminary Account of Aristotelian Attribution

As far as can be seen from Aristotle's logic texts, attribution is a primitive logical relationship, and is to be distinguished from set membership in the extensional interpretation of modern Px. It is technically expressed as "A belongs to B" ("B is A", "B can be taken as an A"), where A is an attribute and B the subject of attribution. With respect to a possible expression, attribution is also called predication, attribute is a predicate, and the subject of attribution is now the subject of predication. In [13], we have analysed Aristotle's definition of the meaning of AaB proposition:

A is predicated to all B $(AaB) \Longleftrightarrow$ no B can be taken to which A is not predicated.

Natural reading of this definition seems to imply the existential import of subject B. Namely, non-existence of anything that is B would imply that A (or whatever other predicate) is not predicated to any B, simply because there are no Bs. But this seems to be denied by AaB. So, in this reading, to verify that there is no B to which A is not predicated, there should be a B, and to each B A should not be predicated. It is not modern $\neg\exists x(Bx \wedge \neg Ax)$, but rather something like $A \not< X$ for no X which is chosen as B (with $<$ for 'belongs to'). In this way, we get the following propositions in the square of opposition:

$(a) A \not< X$ for no X which is taken as B,

$(e) A < X$ for no X which is taken as B,

(i)not $(A < X$ for no X which is taken as $B)$,

(o)not $(A \not< X$ for no X which is taken as $B)$,

[3]For the causal meaning of the expression form 'if something holds it is necessary for something (else) to hold', used in Aristotle's definition of the syllogism, see, e.g. *An. Post.* B11.

which we abbreviate in the following way:

(a) $A \not< $ no B,
(e) $A <$ no B,
(i) not ($A <$ no B),
(o) not ($A \not<$ no B).

X that is taken (*ecthesis*) as a B should be meant, according to Aristotle's examples, as a lower species of B, also as a species of the species, till the individuals to which the lowest species (*species infimae*) are attributed. Hence, in the proposed reading "existential import" of subject terms holds for a and i propositions but not for e and o. It can be easily seen that such a reading enables all opposition in the square of propositions (for existential import in general, see [16, 17]). In our formalisation we will use a modified *ecthesis* approach, combined with a sort of (hidden) reflexivity, and will be able to retain in logic only attributes, without individuals.

The alternative Aristotle's definition of the meaning of a proposition does not use *ecthesis* from the subject term but directly relates subject and predicate terms by means of the whole—part relation: AaB means "B is in A as in a whole" (this approach is used in [15]). So B (or that to which B is attributed) seems to be assumed as somehow given in order to be a part of A (or of that to which A is attributed) as a whole.[4] This manner of speaking can completely avoid any mention of individuals (no X should be taken as an example of a lowest species) and enables to consider exclusively attributes in their mutual attribution.

4 Causal Language and Models

In what follows, we propose a formalisation of Aristotle's account of a causal (scientific) theory, possibly with some simplifications, which presents in more detail the general Aristotelian logical structure of causality. Each causal theory (each special science) represents a one-rooted tree structure with a genus at the root and the lowest (unanalysable) species as leaves, interconnected with one another and with the genus by means of a mutual causal attributions.

We define language $\mathcal{L}$ and the semantics for a theory that includes Aristotle's general syllogistic, accompanied with primary propositions that are specific for a given scientific field. The language includes explicit causal prefixes indicating causes in front of propositions.

Vocabulary, $\mathrm{A}_0, \ldots, \mathrm{A}_n$; operators, a, e, i, and o (subscripts and superscripts will be omitted if no ambiguity arises). $\mathcal{P}$ is the set $\mathrm{A}_0, \ldots, \mathrm{A}_n$.

[4]In this sense, Aristotle speaks, for example, of species ("man", "horse") to be a part of a genus ("animal"), Met. Γ, 26. It is the whole–part relation in the distributive sense that each of the many (parts) is one (genus), not in the collective sense of the one that consists of many.

Definition 4.1 (Term)

1. A_i is a term (predicate letter),
2. If Φ is a term, $\overline{\Phi}$ is a term,
3. If Φ and Ψ are terms, $(\Phi\Psi)$ is a term (usually with outer parentheses omitted),
4. If Φ and Ψ are terms, $\Phi \times \Psi$ is a causal (prefix) term.

Definition 4.2 (Sentence) If Φ, Φ', and Γ, are terms, then

1. $\Phi a \Phi'$, $\Phi i \Phi'$ are sentences,
2. $\Gamma : \Phi a \Phi'$, $\Gamma : \Phi i \Phi'$, $\Gamma : \Phi e \Phi'$ and $\Gamma : \Phi o \Phi'$, where Γ can also be a causal term, are sentences. Also
3. if ϕ is a sentence, then $\neg\phi$ is a sentence.

$\Phi e \Phi'$ and $\Phi o \Phi'$ abbreviate $\neg\Phi i \Phi'$ and $\neg\Phi a \Phi'$, respectively.

Definition 4.3 (Theory tree, $\mathcal{T}$) Theory tree $\mathcal{T}$ is a set $\langle W, w_0, < \rangle$, which is a finite ternary tree, i.e. a finite set W that is partially ordered by $<$ with at most three different immediate successors, where for each $w \in W$, $\{w' \mid w' < w\}$ is well-ordered, and with w_0 as a least element.

We call members of W nodes, and $<$ a basic predication relation. Further, we call each maximal totally ordered subset b of W a $\mathcal{T}$*-branch*, and each initial segment p of a branch a $\mathcal{T}$*-path*. The *height* of a w, $h(w)$, is the order-type of $\{w' \mid w' < w\}$. Since p is a sequence we write $\langle w_i, \ldots, w_k \rangle \subseteq p$ for $\{\langle j, w_i \rangle, \ldots, \langle l, w_k \rangle\} \subseteq p$, where j and l are ordinals.

Definition 4.4 (Frame, $\mathcal{F}$) Frame $\mathcal{F}$ is a set $\langle \mathcal{T}, \mathbf{A} \rangle$, where $\mathcal{T}$ is a theory tree, and $\mathbf{A}$ is a finite set of basic attributes such that there is a bijection from $\mathcal{P}$ to $\mathbf{A}$.

Definition 4.5 (Attributive equivalence class, $[A]$) $[A] = \{A\} \cup X \in \wp\mathbf{A}$ such that

1. $A \notin X$,
2. For each A and B with $A \neq B$, $([A] \setminus \{A\}) \cap ([B] \setminus \{B\}) = \varnothing$.

Definition 4.6 (Labelling) $\mathcal{V}$ is a labelling function such that $\mathcal{V}(w \in W) \in \{[A_i]\}_{i \leq n}$ and

1. $h(w) = h(w')$ if $A \in \mathcal{V}(w)$ and $A \in \mathcal{V}(w')$,
2. w and w' do not have the same immediate predecessor if $\mathcal{V}(w) = \mathcal{V}(w')$.

Instead of $\langle w_i, \ldots, w_k \rangle \subseteq p$ we will usually write $\langle [A_i], \ldots, [A_k] \rangle \subseteq p$ if $\mathcal{V}(w_i) = [A_i]$ and $\mathcal{V}(w_k) = [A_k]$.

We can now introduce the concepts of species and genus.

Definition 4.7 (Species) Species is an attribute $A_0 \ldots \ldots A_k$ such that the sequence $\langle [A_0], \ldots, \ldots, [A_k] \rangle$ is by $\mathcal{V}$ in 1–1 correspondence with a $\mathcal{T}$-path p of $\mathcal{F}$.

A *lowest species* (*infima species*, *atomōn eidos*) is a species defined by a 1–1 correspondence with a branch b of $\mathcal{F}$. Each non-lowest species is a *genus*.

As an *example*, the first genus (the general subject) of arithmetic is for Aristotle "number", with "odd" and "even" as the first pair of basic essential attributes. Further, odd numbers as well as even numbers are distinguished with respect to "non-measurable by a number" (prime) and "measurable by a number". The further distinction, applied to both previous attributes, is the distinction with respect to "non-compounded of numbers" (prime in the second sense), and "compounded from numbers" (let us have in mind that for Aristotle "one" is not a number, and that a number is not a measure of itself). For example, the lowest species that we obtain by successively determining numbers by attributes "odd", "non-measurable by numbers", "non-compounded of numbers" is "triad" (see *An. Post.* B13 96a35–96b1).

According to Aristotle, a causal theory should be primarily concerned with essential attributes, but sometimes equivalent peculiar properties (like "non-twinkling" of planets) occur that are dependent on essential attributes (like "near" of planets). Only essential attributes give a scientific, causal proofs (*demonstratio propter quid*, *apodeixis tou dia ti*), whereas peculiar properties could give only factual explanation (*demonstratio quia*, *apodeixis tou hoti*). Another Aristotle's example of essential and peculiar attributes is "spherical", as an essential attribute of the Moon, and "waxing", as a dependent equivalent property: the Moon is waxing in its specific way because it is spherical, not vice-versa (Moon should be here conceived as a lowest species in astronomy, that applies to only one object).

In the definition of a *model*, the actualisation function is included in order to model causal interrelationships. The idea is that the causal interrelationship between immediately connected (or non-connected) attributes in a tree is actualised internally (by means of the essences of the respective attributes themselves), and that otherwise external actualisation (through intermediate attributes) is presupposed. This is a simplification of the prefixing mechanism used in causally interpreted justification logic, where a proposition prefix reproduces the whole causal structure that leads (in a system) to the proposition (see [10, 14]).

Definition 4.8 (Model) Model $\mathfrak{M}$ is a triple $\langle \mathfrak{F}, \mathcal{V}, \mathcal{I}, \mathcal{A} \rangle$ where

1. $\mathfrak{F}$ is a frame,
2. $\mathcal{V}$ is a labelling function,
3. $\mathcal{I}$ is an interpretation function such that $\mathcal{I}(\mathrm{A}_i) = A_i$, $\mathcal{I}(\overline{\Phi}) = \overline{\varPhi}$, $\mathcal{I}(\Phi\Psi) = \varPhi\varPsi$, $\mathcal{I}(\Phi \times \Psi) = \varPhi \times \varPsi$,
4. Let p be a path, and $[A]^k$ a node of height k to which $[A]$ is assigned; $\mathcal{A}$ is an actualisation function from the set of formulas to the set of attributes:
   (a) $B, C \in [A] \implies \mathcal{A}(\mathrm{B}a\mathrm{C}) = A$,
   (b) (i) $\exists p \langle [A]^k, [B]^{k+1} \rangle \subseteq p \;\&\; \neg \exists p' \langle [A']^k, [B]^{k+1} \rangle \subseteq p' \implies \mathcal{A}(\mathrm{A}a\mathrm{B}) = A \times B$, where $A' \neq A$,
   (ii) $\exists p \langle [A]^k, [B]^{k+1} \rangle \subseteq p \implies \mathcal{A}(\mathrm{A}i\mathrm{B}) = A \times B$,
   (iii) $\exists p \langle [A]^k, [B]^{k+1} \rangle \subseteq p \;\&\; \exists p' \langle [A]^k, [C]^{k+1} \rangle \subseteq p' \implies$ $\mathcal{A}(\mathrm{B}e\mathrm{C}) = \mathcal{A}(\mathrm{B}o\mathrm{C}) = B \times C$, $\mathcal{A}(\mathrm{B}o\mathrm{A}) = B \times A$, $\mathcal{A}(\mathrm{C}o\mathrm{A}) = C \times A$,
   (c) $\mathcal{A}(\Phi * \Psi) = \varGamma \iff \mathcal{A}(\Psi \# \Phi) = \varGamma$,
   $*, \#$ are the following possible pairs of operators (according to the possible conversions): a, i; i, i; e, e.
   (d) (i) $\mathcal{A}(\Sigma * \Phi) = \varGamma \;\&\; \mathcal{A}(\Phi \# \Psi) = \varPhi \times \varPsi$
   $\implies \mathcal{A}(\Sigma \S \Psi) = \varPhi$,

(ii) $A(\Sigma * \Phi) = \Gamma \ \& \ \mathcal{A}(\Phi \# \Psi) = \Delta$
$\Longrightarrow \mathcal{A}(\Sigma \S \Psi) = \Delta,$ where $\Delta \neq \Phi \times \Psi$,
$*, \#, \S$ are the following possible sequences of operators (according to the first syllogistic figure): a, a, a; e, a, e; a, i, i; e, i, o.

We defined the actualisation function so as to have one attribute as a value, following Aristotle's view that cause should be a sufficient and necessary condition for its effect (*An. Post.* A 13, 78b 15–21; B 15–16). Other causes are only causes in a non-strict sense (e.g. *propria* of the real cause). However, one and the same effect can sometimes be causally explained, in different approaches, by different sort of causes, e.g. by an efficient and a final cause (light shines through a lantern because of its consisting of small particles as well as in order to save one from stumbling, see *An. Post.* B 11, 94b 27–37). Often it seems that Aristotle takes that one of the explanations is the primary one.

If the causation actualisation of a proposition is internal, i.e. $\mathcal{A}(\Phi * \Psi) = \Phi \times \Psi$, the actualisation is directly due to the related attributes, which in accordance with Aristotle's theory should be recognised by means of direct knowledge of essences if all attributes are essential, or by means of induction and perception if any of the attributes is a proper attribute. In the first case, the knowledge of the causal relationship is according to Aristotle essential, in the second case only factual.

Let us extend the notion of a basic predication path to the notion of *attribution path*, r. If $B \in [A]$ at w and $C \in [B]$ at w', then the attribution path extends from w to w', although w and w' might not be connected by any tree path.

Definition 4.9 Attribution path r satisfies term Φ ($r \models \Phi$) iff

1. $[A] \in r$ for $\Phi = \mathrm{A}$,
2. $r \models \mathrm{A}$ for $\Phi \in [A] \backslash \{A\}$,
3. $r \not\models \Psi$ for $\Phi = \overline{\Psi}$,
4. $r \models \Psi \ \& \ r \models \Psi'$ for $\Phi = \Psi\Psi'$.

In the definition of *truth*, only necessary truth is considered. We have excluded accidental, contingent truth, since it does not pertain to knowledge in Aristotelian sense.

Definition 4.10 (Truth)

1. $\mathfrak{M} \models \Phi a \Phi' \iff \exists r\, (r \models \Phi') \ \& \ \forall r (r \models \Phi' \rightarrow r \models \Phi)$,
2. $\mathfrak{M} \models \Phi i \Phi' \iff \exists r\, (r \models \Phi' \ \& \ r \models \Phi)$,
3. $\mathfrak{M} \models \Psi : \Phi * \Phi' \iff \mathfrak{M} \models \Phi * \Phi' \ \& \ \mathcal{A}(\Phi * \Phi') = \Psi$,
4. $\mathfrak{M} \models \neg\phi \iff \mathfrak{M} \not\models \phi$.

Note that self-predication is also included as essential and internal case of causation. Let us give two examples of how the definition of truth functions in case of compound terms.

1. $\mathfrak{M} \models \mathrm{A}a\mathrm{BC}$ iff $\exists r \langle B, C\rangle \subseteq r \ \& \ \forall r (\langle B, C\rangle \subseteq r \rightarrow A \in r)$,
2. $\mathfrak{M} \models \mathrm{AB}a\mathrm{CD}$ iff $\exists r \langle C, D\rangle \subseteq r \ \& \ \forall r (\langle C, D\rangle \subseteq r \rightarrow (A \in r \wedge B \in r))$.

Remark 4.11 (Paraconsistency) Genus is according to Aristotle, a distributive whole of its species, for example, "number" is a distributive whole of "odd" and "even". Since "odd" and "even" are mutually exclusive, their genus has, in a sense, "paraconsistent" character—it contains a contrariety, without implying triviality. Both of genus' (distributive) parts can be attributed to the genus ("number is odd", "number is even"), and the genus is an element in their essential natures (odd, as well as even, are essentially numbers) (cf. *An. Post.* A4 73b 20–21, A6, A22). This seems to be essentially connected with paraconsistency of Aristotelian logic as described by Gomes and D'Ottaviano [11] on the ground of Aristotle's example with contradictory concepts ("Callias and non-Callias", "man and non-man") occurring under the non-contradictory major term of a syllogism ("animal").

The definition of satisfiability is in its formulation quite usual:

Definition 4.12 (Satisfiability) A set Γ of sentences is satisfiable iff there is model $\mathfrak{M}$ such that for each $\phi \in \Gamma$, $\mathfrak{M} \models \phi$.

5 System

The system of an Aristotelian causal theory contains a finite number of primitive propositions and rules of inference.

(a) Finite number of primitive propositions (*archai*).

We first define *predicative equivalence class* in the same way as an attributive equivalence class by replacing in the definition attributes $A, B, \ldots$ with predicate letters $\mathrm{A}, \mathrm{B}, \ldots$, respectively. The system may be called hybrid because we use the same tree structure (theory tree) as in semantic frame, and a labelling function $\mathcal{U}$, which is isomorphic to $\mathcal{V}$, the only difference being that each attribute A_i in an equivalence class that is associated to a node is replaced by a predicate letter A_i.

Now we build a finite ternary tree by means of immediate relation $[\mathrm{A}_i] < [\mathrm{A}_j]$ (where A_i has height h, and A_j has height $h+1$) with predicative equivalence classes as nodes. We describe the tree by primitive propositions (*archai*):

1. For each pair $[\mathrm{A}_i] < [\mathrm{A}_j]$ where A_i has height h, and A_j has height $h+1$, there is primitive proposition $\mathrm{A}_i * \mathrm{A}_j$, where $* = a$ if A_j does not occur more than once at $h+1$ (and in the tree at all), and otherwise $* = i$.
2. If there is A_i such that $\mathrm{A}_i * \mathrm{A}_j$ and $\mathrm{A}_i * \mathrm{A}_k$ are primitive propositions, then $\mathrm{A}_j e \mathrm{A}_k$ is a primitive proposition.
3. If $\mathrm{A}_j, \mathrm{A}_k \in [\mathrm{A}_i]$, then $\mathrm{A}_j a \mathrm{A}_k$ and $\mathrm{A}_k a \mathrm{A}_j$ are primitive propositions.

(b) Conversion rules:

$$\Phi a \Psi / \Psi i \Phi; \qquad \Phi i \Psi / \Psi i \Phi; \qquad \Phi e \Psi / \Psi e \Phi.$$

(c) Indirect inference rule (S: a set of propositions, Contrd: negation of a member of S):

$$S, \neg\chi/\text{Contrd} \quad \Longrightarrow \quad S/\chi.$$

(d) Categorical syllogism rules:

$$\Sigma a\Phi,\, \Phi a\Psi/\Sigma a\Psi; \qquad \Sigma e\Phi,\, \Phi a\Psi/\Sigma e\Psi;$$
$$\Sigma a\Phi,\, \Phi i\Psi/\Sigma i\Psi; \qquad \Sigma e\Phi,\, \Phi i\Psi/\Sigma o\Psi.$$

(e) Causal rules:

$\mathrm{A} * \mathrm{B}/\mathrm{A} \times \mathrm{B} : \mathrm{A} * \mathrm{B}$ if $\mathrm{A} * \mathrm{B}$ is a primitive proposition on the ground of $<$;
$\mathrm{A} * \mathrm{B}/\mathrm{C} : \mathrm{A} * \mathrm{B}$ if $\mathrm{A} * \mathrm{B}$ is a primitive proposition on the ground of the membership in $[C]$;
$\Gamma : \Phi * \Psi/\,\Gamma : \Psi\#\Phi$, for $*$ and # as in the conversion rules above;
$S, \Gamma : \neg\chi/\mathrm{Contrd} \Longrightarrow S/\,\Gamma : \chi$;
$\Gamma : \Sigma * \Phi,\ \Phi \times \Psi : \Phi\#\Psi/\Phi : \Sigma\S\Psi$,
$\Gamma : \Sigma * \Phi,\ \Delta : \Phi\#\Psi/\Delta : \Sigma\S\Psi$, for $\Delta \neq \Phi \times \Psi$,
$\Gamma : \phi/\phi$.

If we apply the usual rules for the reduction to the first syllogistic figure, we obtain the following causal syllogisms for the second figure:

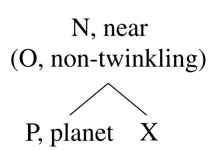
Γ: $\Phi * \Sigma,\ \Delta : \Phi\#\Psi/\Delta : \Sigma\S\Psi$ ($\Gamma : \Sigma\S\Psi$ in *Camestres*),
for $\Delta \neq \Phi \times \Psi$ (for $\Gamma \neq \Phi \times \Sigma$ in *Camestres*),
Γ: $\Phi * \Sigma,\ \Delta : \Phi\#\Psi/\Phi : \Sigma\S\Psi$,
for $\Delta = \Phi \times \Psi$ (for $\Gamma = \Phi \times \Sigma$ in *Camestres*),

and the following causal syllogisms for the third figure:

Γ: $\Sigma * \Phi,\ \Delta : \Psi\#\Phi/\Delta : \Sigma\S\Psi$ ($\Gamma : \Sigma\S\Psi$ in *Disamis*),
for $\Delta \neq \Phi \times \Psi$ (for $\Gamma \neq \Phi \times \Sigma$ in *Disamis*),
Γ: $\Sigma * \Phi,\ \Delta : \Psi\#\Phi\,/\,\Phi : \Sigma\S\Psi$,
for $\Delta = \Phi \times \Psi$ (for $\Gamma = \Phi \times \Sigma$ in *Disamis*),

Let us analyse some characteristic situations that can appear within an Aristotelian causal theory T, containing all primitive propositions and closed under consequences.

Example (Cause and fact, see *An. Post.* A 13, 78a30–78b4) We show how "non-twinkling" may be used as a middle term in a syllogism to prove the nearness of planets, although it is not the real cause of the nearness of planets, "non-twinkling" being only a peculiar property corresponding to the nearness of planets. Such non-causal middle terms serve for Aristotle only to demonstrate a fact (*demonstratio quia*), not the reason why (*demonstratio propter quid*). The proposition that planets are near is taken to be a primitive (astronomical) proposition. In the right side proof below, we start from causal prefixes according to the causal theory of the left-side proof—middle term O of the right-side proof does not have a causal role.



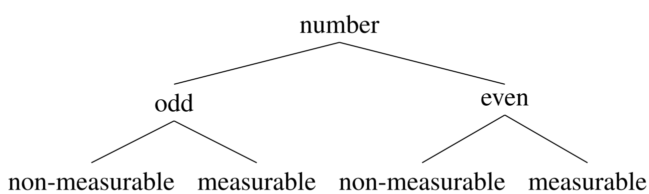
1	N : O*a*N	prim. prop.	1	N : N*a*O	prim. prop.
2	N × P : N*a*P	prim. prop.	2	N : O*a*P	left. syll, line 3
3	N : O*a*P	from 1 and 2, caus.	3	N : N*a*P	from 1 and 2, caus.

Example (A too remote "cause", *An. Post.* A 13, 78b13–31) It seems that in the second syllogistic figure a proof may be given through a non-adequate, too remote cause. In Aristotle's example, "not being animal" is not an adequate reason for why wall does not breathe, since wall is also not an animal that does *not* breathe (there are non-breathing animals)—although it is a reason of the sole fact *that* wall does not breath. However, the syllogism seems to suggest "not being an animal" as the reason due to the minor, negative, premise. The real cause for why wall does not breathe seems to be, in the analysis below, "being an animal" (that what breathes *is* an animal). Let 'A' stand for 'animal', 'B' for 'breathing', and 'W' for 'wall'. Second-figure proof (*Camestres*) is to the left, and its reduction to the first figure (with an improvement of causal explanation) is to the right.

1	A × B : A*a*B	primitive prop.
2	A × W : A*e*W	primitive prop.
3	$\overline{\mathrm{A}}$: B*e*W	∗from 1 and 2

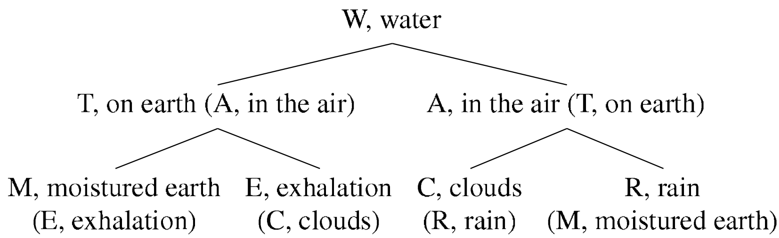
1	A × W : W*e*A	caus. conv.
2	A × B : A*a*B	primitive prop.
3	A : W*e*B	from 1 and 2, caus.
4	A : B*e*W	from 3, caus. conv.

Let us note that because of possible re-occurring of attributes in a tree, a *convergence* can appear in a labelled tree. For example, in the mentioned example of Aristotle's arithmetic, an odd number (triad), as well as an even number (dyad), can be prime in the sense of being "non-measurable by numbers":



However, there are no two different essences of prime numbers (odd, even) but of different *species* of prime numbers.

Example (Causal cycle, see *An. Post.* B 12, 95b38–96a7) Sometimes, a specific causal substructure has to be added in a model, more precisely, into an equivalence class $[A]$ at a tree node. This should be done, for instance, in case of water cycle (see *An. Post.*B12 95b38–96a7). We describe such causal cycles by transitivity, assigning to properties in $[A]$ a double relative role of cause attributes as well as of peculiar properties (*propria*). A peculiar property corresponding to an attribute and relatively caused by the attribute is given in our scheme in parentheses. For example, there is moistured earth iff there is exhalation (exhalation belongs to moistured earth as its peculiar property)—at least on some regular but essential basis. However, moistured earth causes exhalation (under some general conditions, like the sun), not vice versa.



Here are the first syllogistic proof of the syllogistic chain (left side) and the whole circular chain of the causal inference in a form of sorites (right side). We have put causal prefixes in parentheses to indicate that they have only relative causal role (water being a real cause):

1	(M) : MaE	prim. prop.
2	(E) : EaC	prim. prop.
3	(E) : MaC	from 1 and 2, caus.

1	(M) : MaE	prim. prop.
2	(E) : EaC	prim. prop.
3	(C) : CaR	prim. prop.
4	(R) : RaM	prim. prop.
5	(R) : MaM	from 1 to 5, caus.

6 Soundness and Completeness

We outline main features of the soundness and completeness proofs.

6.1 Soundness

According to the construction of the set of primitive propositions of a system, it is immediate that due to an interpretation function (in a model) there is a set of attributive equivalence classes that corresponds to a chosen set of predicative equivalence classes. It is also obvious that to each tree of primitive propositions there corresponds some theory tree that is a part of a frame. Finally, it is straightforward to prove that each inference rule of the system preserves truth. For example, for conversion from $\Phi a \Phi'$ to $\Phi' i \Phi$, if $\Phi a \Phi'$ is true in a model, it is obvious from Definition 4.10 that there is a an attribution path r in the theory tree of the model such that r satisfies both Φ and Φ', i.e. makes $\Phi' i \Phi$ true. As a further example, categorical syllogism *Barbara* is obviously semantically confirmed by means of the transitivity of $<$ in a theory tree, i.e. if each attribution path r that satisfies Ψ also satisfies Φ, and each r that satisfies Φ also satisfies Σ, then each r that satisfies Ψ also satisfies Σ. In addition, explicit causal conditions in a syllogism strictly correspond to the definition of the actualisation function within Definition 4.8 of a model.

6.2 Completeness

We sketch a proof that for each consistent Aristotelian causal theory there is a corresponding model confirming precisely the sentences that are members of the theory (completeness).

We say that a set S of sentences is *inconsistent* iff it contains contradictories (ϕ and $\neg\phi$) as members, or contradictories are syllogistically deducible from S. Let a theory T contain all sentences deducible from it. Then T is inconsistent iff T contains contradictories as members.

Let us start from a consistent set S of sentences of Aristotelian causal theory T. We extend S to a **maximal consistent** set U adding each ϕ (of T) that can be added without a contradiction. It can be easily seen that U obeys the square of opposition conditions for SP (subject-predicate) sentences: U contains one and only one sentence of each contradictory pair of SP sentences, at least one of subcontraries and at most one of contraries (cf. [9] in a different Aristotelian formal system).

Proposition 6.1

1. $\Phi a \Psi \in U$ *iff for each* X *such that* $\Psi a X \in U$, $\Phi a X \in U$,
2. $\Phi i \Psi \in U$ *iff there is* X *such that* $\Psi a X \in U$, *and* $\Phi a X \in U$,
3. $\Phi * \Psi \in U$ *iff for some* Γ, $\Gamma : \Phi * \Psi \in U$,
4. $\phi \in U$ *iff* $\neg\phi \notin U$.

Proof

1. Suppose $\Phi a \Psi \in U$ as well as $\Psi a X \in U$ but $\Phi a X \notin U$. Then $\Phi o \Psi \in U$. Contradiction.
2. Similarly for $\Phi i \Psi$.
3. Suppose that $\Phi * \Psi \in U$. The proof is based on the causal rules of the system. If (a) $\Phi * \Psi \in U$ is a primitive proposition, then for some Γ, $\Phi * \Psi / \Gamma : \Phi * \Psi$, and thus $\Gamma : \Phi * \Psi \in U$. If (b) $\Phi * \Psi$ is derived by conversion and Γ is the causal prefix of the starting proposition, then Γ is the causal prefix of $\Phi * \Psi$. If (c) $\Phi * \Psi$ is derived by indirect proof, then it gets the same causal prefix under which the negation of $\Phi * \Psi$ was supposed. If (d) $\Phi * \Psi$ is obtained by a syllogism, after establishing the causal prefixes of the premises, we can derive the causal prefix of the conclusion (according to the causal inference rules).
4. Obvious from the definition of U. □

Canonical model $\mathfrak{M}_U$ is a model where, instead of labelling attributes, corresponding labelling predicates are associated with the nodes of a theory tree. Hence, in $\mathfrak{M}_U$ attribution path r **satisfies** Φ, in the basis case where $\Phi = \mathrm{A}$, iff $[\mathrm{A}] \in r$. Hence, **truth** in a canonical model is based on predicates and on their association with nodes of a theory tree (attribution collapses to predication).

Proposition 6.2 $\phi \in U$ *iff* $\mathfrak{M}_U \models \phi$.

Proof

- Basis ($\phi = \mathrm{A}_i * \mathrm{A}_j$). The proposition is immediate on the ground of the identity of the canonical attribution tree with the tree of primitive propositions for theory $\mathcal{T}$.
- We take $\Phi a \Psi$ as an example of $\phi = \Phi * \Psi$. $\Phi a \Psi \in U$ means that for each X such that $\Psi a X \in U$, also $\Phi a X \in U$ ($X = \Psi$ for Ψ being a primitive predicate with the greatest height in the theory tree). According to the hypothesis, $\mathfrak{M}_U \models \Phi a X$ for every X such that $\mathfrak{M}_U \models \Psi a X$. This means that each path satisfying Ψ also satisfies Φ, that is, $\mathfrak{M}_U \models \Phi a \Psi$.

- Let $\phi = \Gamma : \Phi * \Psi$. This means that $\Phi * \Psi \in U$. According to the hypothesis, $\mathfrak{M}_U \models \Phi * \Psi$. But in a model, actualisation function $\mathcal{A}$ assigns to each formula some causal prefix Γ (Definition 4.8). It can be checked that this prefixing mechanism corresponds to the prefixing mechanism of the system. Accordingly, $\mathfrak{M}_U \models \Gamma : \Phi * \Psi$.
- Let $\phi = \neg\psi$. Sentence $\neg\phi \in U$ iff $\phi \notin U$. That is, in accordance with the hypothesis, $\mathfrak{M}_U \not\models \psi$, and equivalently, $\mathfrak{M}_U \models \neg\psi$. □

Lemma 6.3 *Each consistent set of $\mathcal{T}$ is satisfiable.*

Proof Follows from Proposition 6.2. □

7 To Sum Up

Aristotle's account of causality is deeply interconnected with his accounts of logic and attribution. Although an Aristotelian causal theory, as formalised in this paper, is restricted to one-place predicates and a finite domain (of attributes) it can, as an example, offer some hints on how to unify some philosophical features of the concept of causality (including necessity and essential character). At the same time, and in comparison with causally interpreted justification logic, the proposed formalisation indicates a possible way how to simplify (although not without a loss in expressivity) a causal prefixing mechanism.

References

1. Aristotle: Analytica Priora et Posteriora. Oxford University Press, Oxford (1964). D. Ross and L. Minio-Paluello (eds.)
2. Aristotle: Metaphysica. Oxford University Press, Oxford (1973). W. Jaeger (ed.)
3. Aristotle: Posterior Analytics. Transl. by G.R.G. Mure. The Internet Classics Archive (1994–2000). http://classics.mit.edu//Aristotle/posterior.html
4. Aristotle: Prior Analytics. Transl. by A.J. Jenkins. The Internet Classics Archive (1994–2000). http://classics.mit.edu//Aristotle/prior.html
5. Aristotle: Posterior Analytics, 2nd edn. Oxford University Press, Oxford (2002). Transl. by J. Barnes
6. Aristotle: Prior Analytics, Book 1. Oxford University Press, Oxford (2010). Transl. by G. Striker
7. Beebee, H., Hitchcock, C., Menzies, P.: The Oxford Handbook of Causation. Oxford University Press, Oxford (2012)
8. Béziau, J.-Y.: On the formalisation of the principium rationis sufficientis. Bull. Sect. Log. **22**, 2–3 (1993)
9. Corcoran, J.: Completeness of an ancient logic. J. Symb. Log. **37**, 696–702 (1972)
10. Fitting, M.: Possible world semantics for first order LP. Ann. Pure Appl. Log. **165**, 225–240 (2014)
11. Gomes, E.L., D'Ottaviano, I.M.L.: Aristotle's theory of deduction and paraconsistency. Principia **14**, 71–97 (2010)
12. Gödel, K.: Texts relating to the ontological proof. In: Feferman, S., et al. (eds.) Collected Writings, vol. 3, pp. 429–437. Oxford University Press, Oxford (1995)
13. Kovač, S.: Causation and intensionality in Aristotelian logic. Stud. Philos. Christ. **49**, 117–136 (2013)
14. Kovač, S.: Modal collapse in Gödel's ontological proof. In: Szatkowski, M. (ed.) Ontological Proofs Today. Ontos [de Gruyter], Frankfurt (2012)
15. Malink, M.: Aristotle's Modal Syllogistic. Harvard University Press, Cambridge (2013)
16. Parsons, T.: Things that are right with the traditional square of opposition. Log. Univers. **2**, 3–11 (2008)

17. Read, S.: Aristotle and Łukasiewicz on existential import (2013). http://www.st-andrews.ac.uk/~slr/Existential_import.pdf

S. Kovač (✉)
Institute of Philosophy, Ul. grada Vukovara 54, 10000 Zagreb, Croatia
e-mail: skovac@ifzg.hr

Using Each Other's Words

Marcus Kracht

Abstract Natural languages are polyphonic: typically, no two speakers associate identical meanings with all the words they are using. Also, the meanings of words may change over time. Yet, we are still missing a formal framework with which to handle this variety. This paper is making a first step by introducing the logic of so-called *deflectors*. These are devices to borrow someone else's language.

Keywords Deflectors · Modal logic · Products of modal logics

Mathematics Subject Classification (2000) Primary 03B45 · Secondary 03B65

1 Varieties of Meaning

There are essentially two ways in which a sentence can cease be true. One is that it describes a state of affairs that no longer obtains. The other is that the state of affairs is no longer described by that same sentence. The first one is the usual scenario. The sentence "The door is open." is true on one occasion because the door actually is open; and it ceases to be true because the door gets closed. So far so good. Now consider this: nothing changes, and yet the sentence ceases to be true. A spectacular case of this sort is the sentence

`There are nine planets.` (1)

(Here, as further down, we are talking about planets of the Sun rather than what is also referred to as "moons", like Ceres. Also, in informal discussion I will continue to use the English word "planet" as if it still denotes all the nine customary planets.) It was true in 1980, but it no longer is true. What has happened? Astronomers have decided in 2006 on a new definition of the word "planet", and that new definition excludes Pluto. Hence what is true in astronomical sciences now, in 2014, is rather

`There are eight planets.` (2)

So, (1) ceased to be true not because the world has changed, but because (1) no longer describes a state of affairs that obtains.

This situation is pervasive, actually. As Kracht and Klein [2] have argued, there is no guarantee that the meaning of words is the same for two different speakers. Any learning algorithm will lead to slight differences between individuals. Thus, whether or not a certain object will be called "green" or "yellow" will differ when the colour is in the middle

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_15

between the clear cut cases. Or, with respect to the above case, it may very well be that plenty of people have not heard about the scientific reform and so continue to use the word "planet" in the old sense. Hence for them (1) is still true. To say that they are wrong is to maintain that there can be only one meaning to a word and that it is in this case up to the astronomers to say what is the correct use of "planet". But if there is only one correct meaning, how is one to interpret the change that happened in 2006? Can we say that the astronomers talking in the 1980s have been wrong in saying there are nine planets? Is "planet" one of these infamous non-natural kinds that Goodman defined? I think, this is clearly not the case. And we can diagnose that by looking at ways in which we refer to the alternative way of talking. Not even astronomers would say, for example,

`I was wrong in 1986 in calling Pluto a planet.` (3)

Obviously, they think, as we do, that they had every right in 1986 to call Pluto a planet. And that is because the meaning of that term at the time of utterance was such that it included Pluto.

To describe the situation, people would say, for example,

`Pluto is no longer called a planet.` (4)

Contrast this with

`Pluto no longer is a planet.` (5)

Consider an event in 2000 (before the change in definition) whereby Pluto collides with a big asteroid and is being kicked out of its orbit so that it leaves the solar system. As a consequence, astronomers have decided to stick with the original definition of "planet". Then, while (5) would be appropriate today in 2014, (4) definitely would not. The reason is that (4) asserts that a change in the way we call things has happened. Statement (5), however, only states that the sentence "Pluto is a planet." has become false.

This much has always been known, of course. A sentence is true because of the way the world is and what the words mean together with the composition algorithm for meanings. Change one of the ingredients and your sentence may change truth value. What has been missing so far is a logical treatment of this phenomenon. Or, to say it more modestly, I am not aware of any such treatment. I wish to provide one here.[1]

2 Deflectors

If the various meanings would simply coexist side by side, there would be no point in developing a theory. However, language also provides means to quote the meaning of a word from different sources. In addition to talking our own talk, we can also talk other

[1]There is a resemblance with what is known as two dimensional semantics, see, for example, Stalnaker [5]. However, the approach taken here is different in that I use two distinct sets of worlds as coordinates and that the content attributed to a speaker is not obtained by diagonalisation (obviously, if the coordinate sets are different, this cannot be done). Moreover, the quotation operators used here (called deflectors) are absent in [5]. I thank Daniel Milne for urging me to consider that issue.

people's talk. We can say, for example, "in your words" or "in Finnish" and the like and thereby change the way in which our words are to be interpreted. I call the devices to achieve this *deflectors*. We shall provide a formal theory of such deflectors.

We assume a fixed language, L, that has constants and a fixed mode of interpreting complex expressions. However, the so-called constants will lack a uniform interpretation. Instead, their interpretation may vary according to some *index*. This index can be a time point, a speaker, a dialect, a combination of the two, etc. Given an index i and a sentence φ, the expression "$\langle\!\langle i \rangle\!\rangle\varphi$" may be translated as "in the words of i, φ" when i is a person, or as "speaking i, φ" or "in the terminology of i, φ", when it is a mode of expression.

In addition to indices, we also have worlds and a modality $\Diamond$ that ranges over possible worlds. So, "$\Diamond\varphi$" says that φ is possible. Although we can think of several more modal operators, one will suffice to demonstrate the interaction between modality and deflection.

The basic symbols of the language are the following:

1. A set C of constants.
2. A set I of indices.
3. $\langle\!\langle \cdots \rangle\!\rangle$, $\Diamond$, $\neg$, $\wedge$.

There are no variables. A *proposition* is formed as follows:

- Any constant is an proposition.
- If φ, χ are propositions, so are $\neg\varphi$, $\Diamond\varphi$ and $\varphi \wedge \chi$.
- If φ is a proposition and i an index, $\langle\!\langle i \rangle\!\rangle\varphi$ is a proposition.

The brackets "$\langle\!\langle \cdots \rangle\!\rangle$" are called *deflectors*. They allow quoting a proposition in the words of someone else. A *preframe* is a triple $\mathcal{D} = \langle W, R, \{J_i : i \in I\}\rangle$, where $R \subseteq W^2$ is a binary relation and for every index i, $J_i : C \to 2^W$ is an interpretation of the constants.

$$
\begin{aligned}
\big\langle \mathcal{D}, (w, i)\big\rangle \vDash c \quad &:\Leftrightarrow \quad w \in J_i(c),\\
\big\langle \mathcal{D}, (w, i)\big\rangle \vDash \neg\varphi \quad &:\Leftrightarrow \quad \big\langle \mathcal{D}, (w, i)\big\rangle \nvDash \varphi,\\
\big\langle \mathcal{D}, (w, i)\big\rangle \vDash \varphi \wedge \chi \quad &:\Leftrightarrow \quad \big\langle \mathcal{D}, (w, i)\big\rangle \vDash \varphi; \chi,\\
\big\langle \mathcal{D}, (w, i)\big\rangle \vDash \Diamond\varphi \quad &:\Leftrightarrow \quad \text{for some } w' \text{ such that } wRw' : \big\langle \mathcal{D}, \big(w', i\big)\big\rangle \vDash \varphi,\\
\big\langle \mathcal{D}, (w, i)\big\rangle \vDash \langle\!\langle j \rangle\!\rangle\varphi \quad &:\Leftrightarrow \quad \big\langle \mathcal{D}, (w, j)\big\rangle \vDash \varphi.
\end{aligned}
\tag{6}
$$

Formulas are evaluated at pairs consisting of a world and an index.

The effect of $\langle\!\langle i \rangle\!\rangle$ is to change the index of interpretation. This calls for a polymodal reformulation. The preframe $\mathcal{D}$ will be changed to a *frame*

$$
\mathcal{D}^\sharp := \big\langle W \times I, R^\sharp, \{f_i : i \in I\}, J\big\rangle,
\tag{7}
$$

where

1. $R^\sharp := \{((w, i), (w', i)) : wRw', i \in I\}$.
2. $f_j((w, i)) := (w, j)$.
3. $J(c) := \{(w, i) : w \in J_i(c)\}$.

(The f_j are functions. However, note that as functions they are also the relations $\{((w, i), (w, j)) : (w, i) \in W \times I\}$. So we have a standard polymodal Kripke-frame, which also interprets the propositional constants.) The clauses of (6) will be transferred in the natural way. For example, here are the first and the fourth clause:

$$\begin{aligned} \langle \mathcal{D}^\sharp, (w,i) \rangle \vDash c \quad &:\Leftrightarrow \quad (w,i) \in J(c) \\ \langle \mathcal{D}^\sharp, (w,i) \rangle \vDash \Diamond\varphi \quad &:\Leftrightarrow \quad \text{for some } (w',i') \text{ such that} \\ &\qquad (w,i) R^\sharp (w',i') : \langle \mathcal{D}^\sharp, (w',i') \rangle \vDash \varphi. \end{aligned} \tag{8}$$

But note that by definition of $R^\sharp$, $i = i'$, so the clause for $\Diamond$ in (8) is equivalent to the one in (6). A *deflector frame* is a frame of the form $\mathcal{D}^\sharp$.

3 An Example

Consider the example of the introduction. We have two constants c_9 and c_8, which are the sentences /`There are nine planets`/ and /`There are eight planets`/, respectively. There are two indices, a and p. The index a corresponds to time points to interpret language before 2006, p to time points from 2006 on. And there are three worlds. In w_0, we have the usual nine planets. In w_1, of the nine Pluto is missing; in w_2, Venus is missing instead of Pluto. (Hence, while w_0 has nine planets in the standard sense, w_1 and w_2 each have eight.) We have $W := \{w_0, w_1, w_2\}$, $R = W^2$ and $J_a(c_9) = \{w_0\}$, $J_a(c_8) = \{w_1, w_2\}$; $J_p(c_9) = \varnothing$, $J_p(c_8) = \{w_0, w_1\}$. This defines the preframe $\mathcal{P}$. We turn this into the frame

$$\mathcal{P}^\sharp = \langle W \times \{a, p\}, R^\sharp, \{f_a, f_b\}, J \rangle, \tag{9}$$

with $f_a((w,i)) := (w,a)$, $f_b((w,i)) := (w,p)$. Consider an astronomer in 1980. If he utters (1), his world is w_0 and his index is a. We find that (1) is true since $J_a(c_9)$ contains w_0. If he utters (1) in 2014, it is false since $J_p(c_9)$ does not contain w_0. However, $\langle\langle a \rangle\rangle c_9$ is true at (w_0, p), as is easily checked. It is the formal rendering of

`In the terminology of 1980, there are nine planets.` (10)

The deflector $\langle\langle a \rangle\rangle$ allows to decontextualise the interpretation.

The deflector frames dissociate the worlds of the preframes into several counterparts. Thus, while the worlds of the preframe $\mathcal{D}$ constitute different states-of-affairs, this is no longer true of the worlds of the deflector frame $\mathcal{D}^\sharp$. How is this to be explained? The idea I am pushing here is that there are two ways of looking at the notion of "state-of-affairs". In the customary meaning, it denotes a way of being of the world without taking into account the language; in the other, new meaning, the language itself also is a way of being of the world. The fact that Pluto is called a planet in a but not in p is in the second sense a way of being of the worlds a and p. Thus, the worlds comprise in addition to the languageless facts also facts of the language.

This can only work nontrivially if there is a separation between a core that remains constant, and a remainder whose interpretation is free. Here the core consists in the boolean connectives, the deflectors and the modal operator $\Diamond$ whose interpretation remains fixed, while the interpretation of the constants is freely assignable.[2]

[2]As natural language allows phrasing deflectors using plain words, such as "in the parlance of the prime minister", the translation of deflectors of the language into those of the logical language is, of course, nontrivial. I shall ignore this problem here.

To encompass this difference, we need a new terminology. We will say that the pair (w_0, p) is the same *world* as (w_0, a), the only difference being how we encode the state of affairs. This calls for a definition. A *polyphonic proposition* in a deflector frame $\mathcal{D}^\sharp$ is an arbitrary subset of $W \times I$. The proposition is *monophonic* if it is of the form $V \times I$ for some $V \subseteq W$. Standardly, a proposition is considered to be monophonic. For we wish to say that whatever a proposition is, as a set of worlds it should not depend on the code that we are using. However, this would mean that the constants "c_9" and "c_8" do not express propositions since their values are not of the required form. Indeed, in the present circumstances we would be inclined to say exactly that. Consequently, we distinguish two types of formulae: those that express a (monophonic) proposition and those that do not. An *utterance* is meant to express a proposition. Hence, when φ is uttered at index i, the proposition that is being expressed is not φ (because that may not denote a proposition) but rather $\langle\!\langle i \rangle\!\rangle\varphi$. In this way, the utterance of the same expression can denote a different proposition in *the same world*. This is a general fact worthy of note.

Proposition 1 *A polyphonic proposition c is monophonic if and only if $c = \langle\!\langle i \rangle\!\rangle c$ for some* (*and hence for all*) i.

Now, when we use variables, it seems that we are forced to say that they denote monophonic propositions rather than polyphonic ones. This will then introduce a dichotomy between constants and propositions, since the latter are constrained in the way the former are not.

However, given the observation above, we can always explicitly reduce a proposition to become monophonic by prefixing it with a deflector.

4 The Logic of Deflector Frames

We will now proceed to an axiomatisation of deflector frames. The resulting logic will not axiomatise all and only the deflector frames. Hence, we shall generalise the terminology to include those frames that satisfy the logic of deflector frames. *Quasi-deflector frames* are frames of the form $\mathcal{F} = \langle W, R, f \rangle$, where W is a set of worlds, $R \subseteq W^2$ and f a function assigning to each $i \in I$ a function f_i such that the following holds.

① For all $w \in W$, $i, j \in I$: $f_j(f_i(w)) = f_j(w)$.
② For all $w \in W$ there is a $i \in I$ such that $w = f_i(w)$.
③ For all $w, w' \in W$, if wRw' and $i \in I$, then $f_i(w) R f_i(w')$.
④ For all $w, w' \in W$ and $i \in I$, if $f_i(w) R w'$ then there is a $w'' \in W$ such that wRw'' and $f_i(w'') = w'$.

The first condition says that the quoted meanings are absolute. The second says that every meaning has some index. The last two conditions state that the maps $w \mapsto f_i(w)$ are p-morphisms. To see why ① is the case, let us take a preframe $\mathcal{D}$. Then in $\mathcal{D}^\sharp$, $f_j(f_i((w, k))) = f_j((w, i)) = (w, j) = f_j((w, k))$. Using (6), this translates into $\langle\!\langle i \rangle\!\rangle \langle\!\langle j \rangle\!\rangle p \leftrightarrow \langle\!\langle j \rangle\!\rangle p$. To see this, observe that

$$\begin{aligned} &\left\langle \mathcal{D}, (w, k) \right\rangle \vDash \langle\!\langle i \rangle\!\rangle \langle\!\langle j \rangle\!\rangle p \\ \Leftrightarrow \quad &\left\langle \mathcal{D}, (w, i) \right\rangle \vDash \langle\!\langle j \rangle\!\rangle p \end{aligned}$$

$$\Leftrightarrow \quad \big\langle \mathcal{D}, (w, j)\big\rangle \vDash p$$
$$\Leftrightarrow \quad \big\langle \mathcal{D}, (w, k)\big\rangle \vDash \langle\!\langle j\rangle\!\rangle p. \tag{11}$$

The other conditions are also easily verified. The satisfaction clauses for deflector frames are standard. Observe that since the f_i are functions we have

$$\langle \mathcal{F}, \beta, w\rangle \vDash \langle\!\langle i\rangle\!\rangle\varphi \quad \Leftrightarrow \quad \big\langle \mathcal{F}, \beta, f_i(w)\big\rangle \vDash \varphi. \tag{12}$$

A *general quasi-deflector frame* is a structure $\langle W, R, f, U\rangle$, where $\langle W, R, f\rangle$ is a quasi-deflector frame, and $U \subseteq \wp(W)$ is a collection of sets closed under complement, intersection; which for every i is closed under the operator

$$\langle\!\langle i\rangle\!\rangle a := \big\{w : f_i(w) \in a\big\} \tag{13}$$

and which is closed under the operator

$$\Diamond a := \big\{w : \exists w' : wRw' \in a\big\}. \tag{14}$$

From this we can extract a logic of quasi-deflector frames. It is characterised by the following axioms. We use $[\![i]\!]\varphi$ to abbreviate $\neg\langle\!\langle i\rangle\!\rangle\neg\varphi$.

$$\begin{array}{ll}
\text{(Ax 1)} & \langle\!\langle i\rangle\!\rangle p \leftrightarrow [\![i]\!] p.\\
\text{(Ax 2)} & \langle\!\langle i\rangle\!\rangle\langle\!\langle j\rangle\!\rangle p \leftrightarrow \langle\!\langle j\rangle\!\rangle p.\\
\text{(Ax 3)} & \Diamond\langle\!\langle i\rangle\!\rangle p \leftrightarrow \langle\!\langle i\rangle\!\rangle\Diamond p.\\
\text{(Ax 4)} & \langle\!\langle i\rangle\!\rangle\Box p \to \Box\langle\!\langle i\rangle\!\rangle p.\\
\text{(Ax 5)} & \Diamond[\![i]\!] p \to [\![i]\!]\Diamond p.\\
\text{(Ax 6)} & \varphi \leftrightarrow \bigvee_{i\in I}\langle\!\langle i\rangle\!\rangle\varphi.
\end{array} \tag{15}$$

(The last axiom requires that I is finite.) (Ax 4) and (Ax 5) are derivable. (Ax 4) is obviously equivalent to $\Diamond[\![i]\!]p \to [\![i]\!]\Diamond p$ (replace p by $\neg p$, and then do contraposition). Using the fact that $[\![i]\!]p \leftrightarrow \langle\!\langle i\rangle\!\rangle p$ is derivable, we get an equivalence with $\Diamond\langle\!\langle i\rangle\!\rangle p \to \langle\!\langle i\rangle\!\rangle\Diamond p$, which is one half of (Ax 3). Similarly, (Ax 5) follows from (Ax 3) by replacing $\langle\!\langle i\rangle\!\rangle$ by $[\![i]\!]$.

Notice that deflector frames are products of frames, one component dealing with the deflectors $\langle\!\langle i\rangle\!\rangle$ and the other with the modality $\Box$. The first logic is called Def_I the second L. We do allow to add any set of postulates for the modality $\Box$. If L is the logic of $\Box$, we denote by $\mathsf{Def}_I(L)$ the logic of adding the postulates of L to Def_I. As we shall show below, this logic is identical to the product of the logics Def_I and L (see [3] for an overview of products). This means that the logic of some class of quasi-deflector frames is the logic of some class of deflector frames, no matter what L is. This is a rather strong result, akin to a result by Gabbay and Shehtman [1] on products of logics with functional operators. For definition, if L is a modal logic based on the operators taken from O and L' is a modal logic based on the modal operators taken from O' where O' is disjoint from O, write $[L, L']$ for the logic axiomatised by

(a) L and L';
(b) $\langle m\rangle\langle m'\rangle p \leftrightarrow \langle m'\rangle\langle m\rangle p$, where $m \in O$ and $m' \in O'$;

(c) $\langle m\rangle[m']p \to [m']\langle m\rangle p$, $m \in O$ and $m' \in O'$;
(d) $\langle m'\rangle[m]p \to [m]\langle m'\rangle p$, $m \in O$ and $m' \in O'$.

This logic is called the *commutator* [3]. The postulates under (b) encode the commutation, the postulates under (c) and (d) encode the Church–Rosser property. (They are dual to each other, so either of (c) and (d) is sufficient.) They state that for each pair of modalities $m \in O$ and $m' \in O'$, if $wR(m)w'$ and $wR(m')w''$ there exists w''' such that $w'R(m')w'''$ and $w''R(m)w'''$. Call a frame $\langle W, R, D\rangle$, with $R : O \cup O' \to W \times W$, $D \subseteq \wp(W)$ a *product frame* if (i) $W = W_0 \times W_1$, and $(u_0, u_1)R(m)(v_0, v_1)$ iff either (iia) $m \in O$, $u_0 R(m) v_0$ and $u_1 = v_1$, or (iib) $m \in O'$, $u_0 = v_0$ and $u_1 R(m) v_1$. Finally, D must be a field of sets closed under the modal operators. A complete logic over $O \cup O'$ is the *product logic* $L \times L'$ if it is the logic of products of frames for L and frames for L'. The question is whether the above axioms suffice to axiomatise $L \times L'$.

Let us now return to the logic of deflector frames. Notice first that the postulates for $\mathsf{Def}_I(L)$ include the postulates for the commutator $[\mathsf{Def}_I, L]$. The remainder of this paper is devoted to showing that this axiomatises the product, whence that it *is* the logic of deflector frames. The following is easily proved.

Proposition 2 *The deflectors $\langle\!\langle i\rangle\!\rangle$ commute with $\neg$, $\wedge$, $\Diamond$, $\Box$. Moreover, every formula is equivalent in Def_I to a formula built from atomic formulae or formulae of the form $\langle\!\langle i\rangle\!\rangle p$, $\langle\!\langle i\rangle\!\rangle c$ $(i \in I)$ using only $\neg$, $\wedge$, and $\Diamond$.*

Observe namely that $\langle\!\langle i\rangle\!\rangle p \equiv [\![i]\!]p$, so that $\neg\langle\!\langle i\rangle\!\rangle p \equiv [\![i]\!]\neg p \equiv \langle\!\langle i\rangle\!\rangle\neg p$. Commutativity over $\wedge$ is clear. Moreover, we have commutativity with $\Diamond$ as an axiom. Hence, we can always push the deflectors inside. Now observe that any sequence of deflectors can be reduced to the innermost deflector.

The axioms (Ax 1–6) above correspond to first-order conditions on the frames saying that for each i, the relation associated with $\langle\!\langle i\rangle\!\rangle$ is a function (Ax 1) such that $f_j(f_i(w)) = f_j(w)$ (Ax 2). (Ax 3) corresponds to the properties (16a) and (16b), respectively:

$$\big(\forall w w' w''\big)\big(wRw' \wedge w'' = f_i\big(w'\big) \to \big(\exists w'''\big)\big(w''' = f_i(w) \wedge w''' R w''\big)\big), \tag{16a}$$

$$\big(\forall w w' w''\big)(w' = f_i(w) \wedge w' R w'' \to \big(\exists w'''\big)\big(wRw''' \wedge w'' = f_i\big(w'''\big)\big). \tag{16b}$$

Equations (16a) and (16b) can be simplified to (17a) and (17b):

$$(\forall w)\big(\forall w'\big)\big(wRw' \to f_i(w) R f_i\big(w'\big)\big), \tag{17a}$$

$$(\forall w)\big(\forall w'\big)\big(f_i(w) R w' \to \big(\exists w''\big)\big(wRw'' \wedge f_i\big(w''\big) = w'\big)\big). \tag{17b}$$

(Ax 4) and (Ax 5) are derivable. (It can also be checked that the Church–Rosser property must hold. Suppose namely that wRw' and $w'' = f_i(w)$ then with $w''' := f_i(w')$ we have $w'' R w'''$.) Thus we have managed to reproduce conditions ①–④. Finally, (Ax 6) corresponds to the first-order condition

$$(\forall w)\Big(\bigvee_{i \in I} f_i(w) = w\Big). \tag{18}$$

This finishes the axiomatisation. Now we shall proceed to show that quasi-deflector frames for a logic $\mathsf{Def}_I(L)$ can be replaced by deflector frames for that same logic.

Let us be given a quasi-deflector frame $\mathcal{F}$ and a world $w_0 \in W$. Define $C_{\mathcal{F}}(w_0) := \{f_i(w_0) : i \in I\}$ and call this the *cycle* of w_0. Likewise, put $S_{\mathcal{F}}(w_0) := \{w' : w_0 R(\Box)^* w'\}$, and call it the *sheaf* of w_0. (As is customary, $R(\Box)^*$ denotes the reflexive transitive closure of $R(\Box)$.) It becomes a frame $\mathcal{S}_{\mathcal{F}}(w_0)$ with the relation R^S, where $w R^S w'$ iff $w R w'$. Furthermore, let $\mathcal{C}_I$ be the following frame: $\langle I, \{g_i : i \in I\}, \wp(I)\rangle$, $g_i(i') := i$ for all $i' \in I$. We shall show that given a model for φ based on $\mathcal{F}$ at w_0 we can define a product frame over $S_{\mathcal{F}}(w_0) \times \mathcal{C}_I$ and a model for φ on this product frame. Notice that the construction even works for general frames. Thus we shall lift the restriction implicit in the definition of products that the component logics be complete.

We start with a model $\langle \mathcal{F}, \beta, w_0\rangle \vDash \varphi$. The frame $\mathcal{F}^\circ$ is defined to be the product of the frames $S_{\mathcal{F}}(w_0)$ and $\mathcal{C}_I$, with the following internal sets. For an internal set a of $\mathcal{F}$, put $a^\circ := \{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \in a\}$. By definition, $a^\circ = \{(w, i) : w \in S_{\mathcal{F}}(w_0), w \in \langle\!\langle i\rangle\!\rangle a\}$,

$$\begin{aligned}(-a)^\circ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \in -a\big\}\\ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \notin a\big\}\\ &= -\big\{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \in a\big\}\\ &= -a^\circ, \end{aligned} \tag{19}$$

$$\begin{aligned}(a \cap b)^\circ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \in a \cap b\big\}\\ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \in a\big\}\\ &\quad \cap \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \in b\big\}\\ &= a^\circ \cap b^\circ, \end{aligned} \tag{20}$$

$$\begin{aligned}\big(\langle\!\langle j\rangle\!\rangle a\big)^\circ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \in \langle\!\langle j\rangle\!\rangle a\big\}\\ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), w \in \langle\!\langle i\rangle\!\rangle\langle\!\langle j\rangle\!\rangle a\big\}\\ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), w \in \langle\!\langle j\rangle\!\rangle a\big\}\\ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), (w, j) \in a^\circ\big\}\\ &= \langle\!\langle j\rangle\!\rangle a^\circ. \end{aligned} \tag{21}$$

For note that by (13), $\langle\!\langle j\rangle\!\rangle a^\circ = \{(w, i) : w \in S_{\mathcal{F}}(w_0), f_j((w, i)) \in a^\circ\} = \{(w, i) : w \in S_{\mathcal{F}}(w_0), (w, j) \in a^\circ\}$. And, finally,

$$\begin{aligned}(\Diamond a)^\circ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), f_i(w) \in \Diamond a\big\}\\ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), \exists w' : f_i(w) R w' \in a\big\}\\ &= \big\{(w, i) : w \in S_{\mathcal{F}}(w_0), \exists w' : w R w' \wedge f_i\big(w'\big) \in a\big\}\\ &= \Diamond\big\{\big(w', i\big) : w' \in S_{\mathcal{F}}(w_0), f_i\big(w'\big) \in a\big\}\\ &= \Diamond a^\circ. \end{aligned} \tag{22}$$

The sets of the form a° are therefore closed under complement, intersection and the modal operators. Hence we have a general frame. By construction, it is a deflector frame.

Now put $\beta^\circ(p) := \beta(p)^\circ$. From the previous considerations, we get that $\beta^\circ(\varphi) = (\beta(\varphi))^\circ$. It follows that

$$\left\langle \mathcal{F}^\circ, \beta^\circ, (w, j) \right\rangle \vDash \varphi. \tag{23}$$

Let us now turn to the general logic of deflector frames. Let L be the logic of $\Diamond$. Assume as above that there are no further axioms concerning deflectors. Thus, additional axioms over Def_I only contain $\Diamond$ as a modal. Denote the resulting logic by $\mathsf{Def}_I(L)$. Denote the logic of pure deflectors by Def_I. Then we have

Theorem 3 *If I is finite*, $\mathsf{Def}_I(L) = \mathsf{Def}_I \times L = [\mathsf{Def}_I, L]$. *If L is complete, so is* $\mathsf{Def}_I(L)$.

Proof We have just shown that $\mathcal{F} \nvDash \varphi$ implies that $\mathcal{F}^\circ \nvDash \varphi$. If $\mathcal{F} \vDash \mathsf{Def}_I(L)$, we also have $\mathcal{F}^\circ \vDash \mathsf{Def}_I(L)$. This is seen as follows. By construction, $\mathcal{F}^\circ$ is a deflector frame, so it satisfies the postulates of Def_I. Moreover, it satisfies the postulates of L since every sheaf is isomorphic to $\mathcal{S}_{\mathcal{F}}(w_0)$, which by assumption on $\mathcal{F}$ is a frame for L. □

5 Conclusion

People do not talk alike, the meaning of words or constructions change from people to people and over time. This does not mean, however, that no analytic tools can be used. In this essay, I have shown how we can borrow each other's language. The logic of deflectors is rather well behaved. Quoting your language is easy. The hard part, of course, is knowing what we are getting into when we do that. I have used a rather well documented case of language reform to demonstrate how we can handle different ways of talking in a fully rational way. It requires knowing when the meaning changed and how. Informal language, however, never is like that. We are often not even aware how subtle the differences are. Yet, as [4] has reminded us, we often *do* borrow meanings from each other when, for example, we tacitly rely on expert opinion. There is then a place for a thorough investigation of the logic and pragmatics of deflection.

Acknowledgements I wish to thank Daniel Milne and Udo Klein for useful discussions.

References

1. Gabbay, D., Shehtman, V.: Products of modal logics. Part I. Log. J. IGPL **6**, 73–146 (1998)
2. Kracht, M., Klein, U.: Notes on disagreement. In: Gutzman, D., Köpping, J., Meier, C. (eds.) Evaluations—Denotations—Entities. Studies on Context, Content and the Foundation of Semantics. Brill, Leiden (2014)
3. Kurucz, A.: Combining modal logics. In: van Benthem, J., Blackburn, P., Wolter, F. (eds.) Handbook of Modal Logic, pp. 869–924. Elsevier, Amsterdam (2007)
4. Putnam, H.: Representations and Reality. MIT Press, Cambridge (1988)
5. Stalnaker, R.: Assertion. In: Cole, P. (ed.) Syntax and Semantics, vol. 9, pp. 315–332. Academic Press, New York (1978)

M. Kracht (✉)
Fakultät Linguistik und Literaturwissenschaft, Universität Bielefeld, Postfach 10 01 31,
33501 Bielefeld, Germany
e-mail: marcus.kracht@uni-bielefeld.de

On Universality and Formality in 19th Century Symbolic Logic: The Case of Schröder's "Absolute Algebra"

Javier Legris

Abstract This paper deals with conceptions of formality underlying 19th century symbolic logic, where notations and manipulation of signs played an important role. It is devoted specifically to the case of Ernst Schröder's "formal algebra", which extended with the algebra of relatives (as developed by C.S. Peirce) constituted the basis for a *Pasigraphy* as a universal notation system. The discussion will begin with the well-known distinction devised by Gottlob Frege between two sorts of formal theories. In the paper, both conceptions of formality will be connected with the corresponding attempts of constructing universal scientific notations (Schröder's *Pasigraphy* and Frege's *Begriffsschrift*). It will be shown that the Pasigraphy was an interpretation of that formal algebra. As a further conclusion, it will be suggested that each of the two conceptions of formality places logic in different levels and determines different conceptions of universality.

Keywords History of logic · Algebra of logic · Formality · Ernst Schröder

Mathematics Subject Classification 01A55 · 0303 · 03A05

1 Introduction

During the 19th century different ideas on formality coexisted in symbolic logic and foundations of mathematics. These ideas concerned the nature of formal objects and formal theories. They were essential in the determination of the scope of logic and its relation to mathematics, and they constitute the background for the discussion about *logical form* in the first decades of the 20th century. Examples of them can be found, first of all, in English symbolic algebra, but also in the theory of forms (*Formenlehre*) due to Hermann and Robert Grassmann, and even in Georg Cantor's theory of sets. The problem was related to the generalization of the notion of number that became independent of the notion of quantity. In many cases, numbers were conceived as formal objects: 'thought objects' or 'creations of the mind'. So, it was not an easy task to make a clear-cut distinction between numbers and logical concepts, and in some cases logic and mathematics were considered to be the same.

It is widely known that in Kant there was already a conception of formality related to logic. In fact, it is regarded as the first explicit characterization of deductive logic as *formal* logic. Kant formulated it in the *Critique of Pure Reason*, when he distinguished it from his transcendental logic [9]:

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_16

> "Now, logic in its turn may be considered as twofold—namely, as logic of the general, or of the particular use of the understanding. The first contains the absolutely necessary laws of thought, without which no use whatsoever of the understanding is possible, and gives laws therefore to the understanding, without regard to the difference of objects on which it may be employed." (*KrV.* A51 B76)

Formal logic had two basic features:

1. As general logic, it makes abstraction of all content of the knowledge of the understanding, and of the difference of objects, and has to do with nothing but the mere form of thought.
2. As pure logic, it has no empirical principles. [. . .] It is a demonstrated doctrine, and everything in it must be certain completely a priori. (Kant *Krv.* A54 B78)

Through this characterization, a notion of formality is also achieved. Formality is related to "object in general", so that a particular domain is not presupposed and logic turns to be universally (or, better, *generally*) applicable. This conception of formality was very influential in 19th century logic, as it has been observed in many opportunities (see, e.g. [12]). A historical survey of the notion of formal logic in 19th century in the United Kingdom can be found in Hodges [8].

In the sequel, a particular case of formality underlying the origins of symbolic logic in the second half of 19th century will be examined. This idea is exemplified by the project of a formal algebra formulated by Ernst Schröder (1841–1902). In its last phase, this project turned out to be *a Pasigraphy* (an universal notation). As a *Leitfaden* the well-known Fregean distinction between two different types of formal theory will be discussed. In 1885, Gottlob Frege (1848–1925) argued for a particular conception of formal theories of Arithmetic. According to it, a theory for Arithmetic is formal if every law of the theory is logically derived only from logical notions via definitions. In vol. II of *Grundgesetze*, Frege characterized his conception of formal theories as contentual. For it is the meaning of the symbols that is the real object of the theory, the symbols being only a medium to express this meaning. Frege opposed this conception of formal theory to the idea of symbolic systems whose elements had no meaning at all. Now, Ernst Schröder with his program of a formal algebra had just advocated for such an idea of formal theory: Formal algebra could be applied to different domains, and the operations received diverse interpretations depending on the domain considered. *This general applicability of formal algebra suggested an alternative idea of universality*.

As a conclusion, I shall suggest that *to different conceptions of formality correspond different conceptions of universality*. Moreover, it will be argued that these different conceptions of universality can be understood properly as alternative ideas of formality.

2 Frege's Ideas on Formal Theories

In a paper from 1885 delivered at the Society of Medicine and Natural Science in Jena, entitled "On Formal Theories of Arithmetic" [*Über formale Theorien der Arithmetik*], Frege maintained a particular conception of formal theories in Arithmetic. According to it, a theory for Arithmetic is formal if every law of the theory is logically derived only from logical notions via definitions.

> "I here want to consider two views, both of which bear the name 'formal theory'. I shall agree with the first; the second I shall attempt to confute. The first has it that all arithmetic propositions

> can be derived from definitions alone using purely logical means, and consequently that they also must be derived in this way" [5, p. 94].

Frege connects his conception of formal theory to the universality of arithmetic: the notion of number is universally applicable (at this point Frege refers to the "general applicability" or comprehensive applicability [*umfassende Anwendbarkeit*] of arithmetic). The notion of number, Frege says, can be applied to ideal or real entities, to the temporal and spatial, and to objects in general or to events, methods, etc.

The other conception taken into account by Frege considers that the signs included in the theory are void of meaning, they have no content at all. A formal theory is just a symbolic system without meaning [5, p. 105]. In the specific case of arithmetic, signs like $1/2$ or π do not necessarily have the current and usual ("pre-formal") meanings.

Some years later, close to the end of vol. II of *Grundgesetze*, from 1903, when he deals with the real numbers, Frege again makes the distinction. In § 156, he refers to a formal trend and a *contentual* trend (*inhaltlich*) in the characterization of arithmetic. The "formal current of thought"

> "calls numbers certain figures produced by writing, which are manipulated according to arbitrary rules" [6, p. 154].

In other contexts, these figures can have other meanings, but the "formal arithmetician", as he says, does not take into account these other meanings. Frege was thinking of his colleague in Jena Carl Johannes Thomae (1840–1921) as the representative of this "formal arithmetician". He had proposed such a formal conception of mathematics in his book *Elementare Theorie der analytischen Funktionen einer komplexen Veränderlichen* (1880).[1]

For the other trend, that Frege calls *contentual*, "arithmetic are signs of their proper objects, numerical signs, external auxiliary means" (*loc. cit.*). The rules follow from the meaning of the signs, and these meanings are the real objects of arithmetic: the arithmetical entities [6, p. 156]. This point of view can be also denominated 'formal', but in another sense: in the sense that the nature of arithmetic is "purely logical" (*loc. cit.*). 'Formal' here means 'logical'.

Frege criticized the formal trend, trying to show its unfeasibility. He argued for the "contentual trend", which clearly reflects the logicist project initiated by him some years before with his *Begriffsschrift*. He repeats some arguments of those papers written after *Begriffsschrift* in order to explain the structure and function of his system of the conceptual writing. In this case, he adds its *universal applicability* as an argument for his logicism in the following sense: The definitions he provides for the signs for arithmetic make possible their correct application in everyday language like in "There are exactly two satellites rotating around Mars."[2]

Frege does not explicitly mention defenders of this position. In the quotation from vol. II of *Grundgesetze*, Frege does not refer to Hilbert's formal axiomatics as developed in his *Foundations of Geometry* from 1899.[3] Some years later, Frege criticized

---

[1]Thommae, together with Heinrich Eduard Heine and Hermann Hankel among others, would represent the school of "old formalism", different from the later formalism of Hilbert's school.

[2]Ignacio Angelelli called this view "philosophical logicism": Arithmetic would deal with the more general features of reality, that is, these features that are in every domain. In this sense it could be compared with ontology or logic (see [1, p. 244]).

[3]In fact, Frege mainly refers to Alwin R. Korselt and his paper "Über die Grundlagen der Geometrie" (see [7, p. 281]). Korselt had worked on Schröder's algebra of logic.

Hilbert's achievements (see [7]). According to this criticism, the basic formulas of an axiomatic system were "pseudo-axioms" and not real statements or sentences. Consequently, Hilbert's system provided also "pseudo-proofs" [*Scheinbeweise*] lacking of an authentic *epistemic* content. In Frege's own words:

> "From the fact that the pseudo-axioms [*Pseudoaxiome*] do not express thoughts, it further follows that they cannot be premises of an inference-chain [*Schlußkette*]. Of course, one really cannot call them propositions [*Sätze*]—groups of audible or visible signs—premises anyway, but only the thoughts expressed by them. Now, in the case of the pseudo-axioms, there are no thoughts at all, and consequently no premises. Therefore when it appears that Mr. Hilbert nevertheless does use his axioms as the premises of inferences and apparently bases proofs on them, these can be inferences and proofs in appearance [*Scheinschlüsse*, *Scheinbeweise*] only" [7, pp. 306 f].

With this criticism, Frege tried to show that the results obtained in a purely formal system, that is, through a formal proof, cannot bring authentic knowledge.[4]

The motivations underlying the *Begriffsschrift* are very well known. Frege's foundational aims led him to the explicit formulation of a logic system in a symbolic language. In this language, there are, firstly, some basic symbols. Every further symbol can be defined from this basis. Secondly, the system also contained axioms and rules of inference, so that the proofs were carried out "under the form of a calculus". Arithmetic knowledge has then a 'logical source'. It can be obtained by means of proofs constructed on the basis of the application of rules of inference. More specifically, the analytic character of mathematical truths follows from logical truths (that are analytic per se) by means of proofs and definitions. Proofs and definitions are the warranting procedures for the truth of a mathematical proposition. For this reason, Frege developed a theory of proofs and a theory of definitions.

So, what was for Frege a formal theory? A formal theory should be constructed from logical axioms, and their theorems should be obtained from definitions and proofs by means only of logical proofs. *Formal is identified with logical*. (In this sense, it was not possible to construct a formal theory for geometry. Geometry, for instance, had for Frege non logical sources; instead it should be based on the intuition of space.)

3 Schröder's Formal Algebra

In the context of his characterization of a formal theory as a symbolic system devoid of meaning, Frege observed that in this case numbers were identified with symbols. A similar idea had been formulated by the mathematician Ernst Schröder some years before. In the introduction of his *Handbook of algebra and arithmetic*, Schröder argued that numbers are signs (*Zeichen*) produced by us in some arbitrary way (see [14, p. 2]). This statement must be understood in its historical context. At that time this point of view meant a deviation from the traditional idea of number that connected number with quantity and measurement in order to obtain a more abstract notion of arithmetic. The domain (*Zahlengebiet*) in algebraic systems can have as elements entities of arbitrary nature and its objects are the regularities characterizing the operation, the "pure" laws governing them. In a text published one year later, *Über die formalen Elemente der absoluten Algebra*, Schröder

[4] This is essentially what Detlefsen called "Frege's Problem" (see [3, pp. 9 ff]).

took into account operations (sum, for example) between concepts, judgements, numbers and points on the plane. They were "symbolic operations". Similar considerations can be found in his review of Frege's *Begriffsschrift*: the abstract calculus can be applied to an arbitrary aggregate (*Mannigfaltigkeit*) constituted by arbitrary domains (*Gebiete*), such as parts of a surface, without taking into account their measure (see [15, p. 85]).

According to this, Schröder defined formal algebra (*formale Algebra*) as the study of the laws on algebraic operations dealing with 'general numbers' in a domain (*Zahlgebiet*). In doing this, no assumption about the nature of their elements is made (see [14, p. 233]). The domain of numbers should not be restricted to number [14, p. 2]. From his point of view, algebra consisted *in manipulation of signs*.

Schröder conceived his formal algebra also as a reconstruction program for the whole mathematics, in which he distinguished between the symbolic system and their applications. (see [13, p. 567]). Schröder called the formal algebra with its applications *absolute algebra*, and he considered it as a general theory about connexions. Logic is just one of the applications in this absolute algebra.

Later, Schröder was able to extend his program, when he adopted the algebra of relatives of Charles S. Peirce. He then considered his symbolic algebra as a universal language, a 'pasigraphy' as he called it. Symbolic algebra of relatives was regarded by Schröder as a universal language and, at the same time, as a foundational theory for *prima facie* every scientific domain [16]. According to Schröder, pasigraphy is a 'new discipline', the aim and purpose of which is to lay down a universal scientific language (*wissenschaftliche Universal-Sprache*). In this language, the basic or fundamental notions (*Grundbegriffe*) of 'pure mathematics' (logic, arithmetic and geometry) are introduced.

> "The problem to be solved for any given branch of science amounts to: expressing all the notions which it comprises, adequately and in the concisest possible way, through a minimum of primitive notions, say "categories", by means of purely logical operations of general applicability, thus remaining the same for every branch of science and being subject to the laws of ordinary Logic, but which latter will present themselves in the shape of a "calculus ratiocinator". For the categories and operations of this "lingua characteristica" or "scriptura universalis" easy signs and simple symbols, such as letters, are to be employed, and—unlike the "words" of common language—they are to be used with absolute consistency" [16, 46].[5]

Summarizing, what is for Schröder a *formal theory*? It is a set of basic symbols with certain combination rules. For these operations and relations certain postulates hold, from which theorems are derived using this 'general logic', that is at a higher level. Eventually, the formal theory can received different interpretations.

4 Discussion: Formality in Schröder

In a quite recent paper, Catarina Dutilh Novaes distinguish two basic senses of formality. The first sense relates formal to schemes, to indifference of properties in objects and

[5] In the Pasigraphy-paper, the algebra of relatives has a privileged interpretation in set theory and Schröder proposed a set-theoretic reconstruction of arithmetic. Therefore, it can be suggested that the formal algebra has an intended interpretation, being then a language in a strict sense. However, the distinction between formal system and possible interpretations remains. For a further discussion of this problem, see [13, pp. 597 ff].

to abstraction from content. The second sense formality is predicated of (the strict application of) rules (see [4]). In the first sense, both objects and signs can be said to be formal. It seems to correspond to the 'schematic notion of the formal' (see [12]). This is the idea of replacing terms with different terms in arguments, the idea of variability of *terms* preserving the validity of arguments.

In a general context, it is useful to understand this idea by saying that a formal theory is merely *schematic*. In the more abstract and general sense, a formal theory consists merely of *schemes*. John Corcoran recently investigated the nature of these symbolic structures beyond their metalinguistic properties. He characterized a schema in the following words:

> "In some of several closely related senses, a *schema* ... is a complex system having several components one of which is *a template-text* or scheme-template, a syntactic string composed of one or more 'blanks' and usually also words and/or symbols to which meaning is assigned" [2, p. 219].[6]

In this framework, Schröder's perspective should include the extreme case of schemata consisting exclusively of blanks. Different interpretations of a formal system can lead to formal systems with more meaningful symbols and less blanks. When there are no blanks (all signs are interpreted), it would be possible to speak of a language. This fact can be visualized through the following diagram.

[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]   only blanks

↓ formal interpretation

⊗ [ ] [ ] [ ] ⊗ [ ][ ][ ] ⊗   fewer blanks

↓ formal interpretation

⊗ [ ] ⊕ [ ] ⊗ [ ] ⊕ [ ] ⊗   fewer blanks

↓ interpretation in a domain

⊗ • ⊕ • ⊗ • ⊕ • ⊗   no blanks

↑ Language ↑

When there are no blanks, every symbol is interpreted. We have then a full interpreted system or theory.

Now, taking into account the whole program of the absolute algebra, (extended with the algebra of relatives) as a tool, the result of the interpretation would be a "language" in a sense close to Frege's "conceptual notation" (*Begriffsschrift*). In the language, some signs should be used now as *variables* for the entities of the domain, and laws or lawlike sentences could be expressed. Formal theories in Frege's sense could represent the other extreme or limit case of formal theories without blanks. But this is erroneous because the very idea of schema does not play a role in Frege's conception of axioms in an axiomatic system (although inference rules are obviously schematic).[7]

Frege and Schröder both insisted (together with Peano among others) on the importance and significance of symbolism in foundational research. The aim of this research

[6]Thommae also referred to "schemes" in his aforementioned book: real numbers should be understood as "pure schemes without content".

[7]Frege made use of the notion of schema in his criticism on Korselt's idea of formal system (see [7, p. 304]).

was to clarify and elucidate the notions considered basic or fundamental for the construction of mathematics and to show clearly and precisely the relations between these basic notions. As a result and from a general point of view, certain *basic structures* were stand out. Symbolic languages seemed to be specially suitable for showing these structures and for the presentation of the laws appropriate for different mathematical domains according to a 'calculus' based on rules of symbolic combination and manipulation. In both cases, we find a handful of basic symbols, from which the whole language is developed.

However, the programs of Frege and Schöder are quite different. Historians of logic have repeatedly noted these differences. For Frege the conceptual writing consisted of basic meaningful symbol (they have a *logical* meaning). All other symbols must be defined from them. On the contrary, for Schröder, the algebraic language consists of basic symbols *without meaning*, but they are subject to certain laws of combination, so that they have a certain *structure*. In a second stage, different interpretations (or applications) can be assigned to the symbols. In this case, the algebraic language can be characterized as a *general language*: it can receive different interpretations.

Also Frege refers to the 'general applicability' of his own system, in the following sense: logical categories (and then arithmetical, too) hold for the world in its totality; they are general in an *ontological* sense. Frege's symbolism was descriptive (contentual), serving to represent the laws *that were universally valid*. At the time of his 1885 lecture on formal theories, Frege did not have a clear notion of orders or levels. Nevertheless, it is possible to understand this 'general applicability' relative to first order, that is, to the level of objects. Logical laws are second order and valid for every object of first order. Thus, 'formal' corresponds to 'higher order'.

Now, some questions arise with regard to the Pasigraphy. According to Schröder, The Pasigraphy is a 'new discipline', the aim and purpose of which is to lay down a universal scientific language (*wissenschaftliche Universal-Sprache*). In this language, the basic or fundamental notions (*Grundbegriffe*) of 'pure mathematics' (logic, arithmetic and geometry) are introduced.

> "The problem to be solved for any given branch of science amounts to: expressing all the notions which it comprises, adequately and in the concisest possible way, through a minimum of primitive notions, say "categories", by means of purely logical operations of general applicability, thus remaining the same for every branch of science and being subject to the laws of ordinary Logic, but which latter will present themselves in the shape of a "calculus ratiocinator". For the categories and operations of this "lingua characteristica" or "scriptura universalis" easy signs and simple symbols, such as letters, are to be employed, and—unlike the "words" of common language—they are to be used with absolute consistency" [16, p. 46].

The primitive signs of the theory are the following [16, p. 47]:

=	·	¯	˘	;
1'	Π			

corresponding to relations and operations in the algebra of relatives. They indicate only *categories* in Schröder's sense: from them, the remaining algebraic signs can be defined. If this is the case, the Pasigraphy has an intended (logical) meaning. Thus, *it is an application of the program of a formal algebra extended with the algebra of relatives* (for further details see [10, 11]).

5 Concluding Remarks: Formal Theories and Universality

Regarding Kant's notion of formality and his distinction between formal and transcendental logic, Frege's logic is as a contentual logic as Kant's transcendental logic was. Frege tried to determinate the nature of numbers (as entities), whereas Schröder merely aimed at knowing their structural properties, independently from their ontological nature.

The distinction between two conceptions of formal theories contributes to a clarification of the other well-known distinction between 'logic as language' and 'logic as calculus' devised by Jean van Heijenoort as a way to identify two traditions that coexisted in the origins of symbolic logic (see [17]). van Heijenoort distinguished between two fundamental lines of thought in the history of modern logic: *logic as language* (represented paradigmatically by Frege's conceptual notation) and *logic as calculus* (represented mainly by the algebra of logic). He took this distinction from Frege's own opposition between *lingua characterica* and *calculus ratiocionator*, formulated in a paper published posthumously. With the expression 'lingua characterica', Frege meant a language with a fixed interpretation (a mathematical domain), serving "to express a content", as Frege himself wrote. On the contrary, a calculus was conceived only as a symbolic system with a manifold of possible interpretations, and it was intended to be a formal representation of logic for solving logical problems.

Now, to each conception of formal theory corresponds a definite idea of universality—and therefore an idea of *universal language*. For Frege (at least at the time of *Begriffsschrift*) a universal language should have the logic contentual basis and any further concept should be introduced by definition. Its universality rests on the fact that *every* concept should be defined in terms of the basic concepts. In Schröder's program, the universality of absolute algebra rests on the fact that it can be applied to *every* domain. At the same time, Schröder realized the algebraic symbols should have a meaning or content, in order to become a *language*. Therefore, in both cases we have pretensions of universality and ideas of universal languages. The differences between both notions of language lay in the semantics: in the case of Frege we have an 'internal' semantics, but it is 'external' in the case of Schröder.

If we take into account the evolution of symbolic logic at the beginning of the 20th century, the continuity of both conceptions of formality and universality can be clearly seen. Frege's and Schröder's programs contain *in nuce* different formal methodologies. On the one hand, both conceptions had an influence on the drawing up of *Principia Mathematica* by Whitehead and Russell. On the other, Schröder's perspective is known to be implicit in the development of model theory, with his typical distinction between language and metalanguage, between the formal systems and their interpretations.

Acknowledgements It is a truism that the idea of universality is essential to Jean-Yves Beziau's work. In some respect, Jean-Yves devotes his life to this idea. The following historical comments can be interesting in connection with it. This paper is a result of a research project supported by the *Consejo Nacional de Investigaciones Científicas y Técnicas* from Argentina (PIP 11220080101334, CONICET). It also benefited from the additional financial support of the Alexander von Humboldt Foundation. I wish to thank Volker Peckhaus for helpful and valuable comments on an earlier draft of this paper.

References

1. Angelelli, I.: Studies on Gottlob Frege and Traditional Philosophy. Dordrecht, Reidel (1967)

2. Corcoran, J.: Schemata: the concept of schema in the history of logic. Bull. Symb. Log. **12**, 219–240 (2006)
3. Detlefsen, M.: Hilbert's Program. An Essay on Mathematical Instrumentalism. Dordrecht, Reidel (1986)
4. Dutilh Novaes, C.: The different ways in which logic is (said to be) formal. Hist. Philos. Logic **32**, 303–332 (2011)
5. Frege, G.: Über formale Theorien der Arithmetik. Sitz.ber. Jeaneischen Ges. Med. Nat.wiss. **12**, 94–104 (1885/1886). Reprinted in Frege G.: Kleine Schriften, 2nd edn. (1990) Edited by Ignacio Angelelli. Olms, Hildesheim–Zürich–New York, Olms, pp. 103–111
6. Frege, G.: Grundgesetze der Arithmetik, vol II. Jena (1903). Reprinted Hildesheim, Olms, 1966
7. Frege, G.: Über die Grundlagen der Geometrie. Jahresber. Dtsch. Math.-Ver. **15**, 293–309 (1906). Reprinted in Frege 1990, pp. 281–323
8. Hodges, W.: A formality. CD Festschrift for the 50th birthday of Johan van Benthem, Vossiuspers AUP (1999). www.aup.nl
9. Kant, I.: KrV. Kritik der reinen Vernunft, 1st edn. 1781, 2nd edn. 1787. In: Weischedel, W. (ed.) Werke in zwölf Bänden, vol. III. Suhrkamp, Frankfurt (1968)
10. Legris, J.: Deux approches des relations logique-mathématiques Frege et Schröder. In: En Justifier en Mathématiques, comp. por Dominique Flament y Philippe Nabonnand, pp. 215–254. Editions de la Maison des sciences de l´homme, Paris (2011)
11. Legris, J.: Universale Sprache und Grundlagen der Mathematik bei Ernst Schröder. In: Mathematik—Logik—Philosophie. Ideen und ihre historischen Wechselwirkungen, comp. por Günther Löffladt. pp. 255–269. Harri Deutsch, Frankfurt am Main (2012)
12. MacFarlane, J.: What does it mean to say that logic is formal? PhD Dissertation, University of Pittsburgh (2000)
13. Peckhaus, V.: Schröder's logic. In: Gabbay, D.M., Woods, J. (eds.) Handbook of the History of Logic. Vol. 3. The Rise of Modern Logic: From Leibniz to Frege, pp. 557–609. Elsevier-North Holland, Amsterdam (2004)
14. Schröder, E.: Lehrbuch der Arithmetik und Algebra. Teubner, Lepizig (1873)
15. Schröder, E.: Review of Begriffsschrift by Gottlob Frege. Z. Math. Phys. **25**, 81–94 (1880)
16. Schröder, E.: Über Pasigraphie, ihren gegenwärtige Stand und die pasigraphische Bewegung in Italien. In: Rudio, F. (ed.) Verhandlungen der Ersten Internationales Mathematiker-Kongresses in Zürich vom 9. bis 11. August 1897. The Monist, vol. 9, pp. 147–162. Teubner, Leipzig (1898). Reprint Nendeln, Kraus, 1967, English translation: on Pasigraphy. Its Present State and The Pasigraphic Movement in Italy en *The monist* 9, 1898, pp. 44–62 (corrigenda, p. 320)
17. van Heijenoort: Logic as calculus and logic as language. Synthese **24**, 324–330 (1967)

J. Legris (✉)
IIEP-BAIRES, Facultad de Ciencias Económicas, Universidad de Buenos Aires,
Av. Córdoba 2122-2º (C1120 AAQ), Ciudad Autónoma de Buenos Aires, Argentina
e-mail: jlegris@retina.ar

Caramuel and the "Quantification of the Predicate"

Wolfgang Lenzen

Abstract The theory of the "Quantification of the Predicate" attempts to transform the traditional logic of the four categorical forms (Every S is P; No S is P; Some S is P; Some S isn't P) into a system of eight or even twelve propositions in which the simple predicate P is replaced by a quantified predicate like 'some P', 'every P' and perhaps even 'no P'. According to the standard historiography of logic, such a theory was invented in the 19th century by W. Hamilton and Augustus De Morgan. However, already in the 17th century, the Spanish logician Juan Caramuel y Lobkowitz published a book "Theologia rationalis" in which propositions with quantified predicates are systematically investigated. By way of a remarkable extension of the traditional theory of conversion, Caramuel arrives at a system of logical inferences which might be considered as a forerunner of Hamilton's theory. However, Caramuel's "method" basically consists only in listing various *examples* of true and false propositions. Therefore, his theory fails to provide a *general semantics* for propositions with a quantified predicate. One variant of such a semantics was developed in the 18th century by Gottfried Ploucquet. Another completely different one had been sketched already in the 17th century by Gottfried Wilhelm Leibniz.

Keywords History of logic · 17th century · Categorical forms · Quantification of the predicate · Juan Caramuel y Lobkowitz · Gottfried Ploucquet · Gottfried Wilhelm Leibniz

Mathematics Subject Classification Primary 03A05 · Secondary 03B65 · 03B99

1 Introduction

According to mainstream historiography of logic, the so-called "Theory of the Quantification of the Predicate" (TQP, for short) was invented towards the middle of the 19th century by William Hamilton and/or Augustus De Morgan. The main point of TQP is to expand the traditional theory of the syllogism, which operates with *four* categorical forms 'Every S is P', 'No S is P', 'Some S is P', 'Some S isn't P', into a system of *eight* propositions where the predicate P is modified by explicitly adding quantifier-expressions like 'every', 'all', 'some', or 'any'. This "Doctrine" had been taught by Hamilton at the University of Edinburgh in the late 1830s before it was published, posthumously, in 1861.[1] Since De

[1]Cf. [11, p. 279]: "All A is all B, Any A is not any B, All A is some B, Any A is not some B, Some A is all B, Some A is not any B, Some A is some B, Some A is not some B".

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_17

Morgan had presented a similar schema of eight propositions in 1847, a fierce controversy arose about who was the "real inventor".[2] According to John Venn, however, none of the two authors was the *first* to develop TQP because similar schemata had already been put forward in 1839 by Thomas Solly [23] and in 1827 by George Bentham [2].[3] Furthermore, in Venn's opinion, TQP had at least been *anticipated* already in the 18th century by Gottfried Ploucquet.

In 2006, Michael Franz published an edition of extracts of Ploucquet's logical writings, [10]. Though I had not heard anything about this author before, I decided to give a seminar about this topic at the University of Osnabrück. I soon became fascinated by Ploucquet's work which comprises some smaller papers collected in [4] and two monographs [20, 21]. It turned out that the logic developed there not only contains a lot of naive errors and serious mistakes,[4] but also some interesting and fruitful new ideas. My reconstruction of Ploucquet's system was presented in [13] and [15], and one of the main results said that Ploucquet's TQP basically accords with that of Hamilton.

On the other hand, I had good reasons to believe that TQP—albeit in a totally different interpretation—had been developed a long time *before* Ploucquet by Gottfried Wilhelm Leibniz. I was therefore inclined to consider *Leibniz* as the "real inventor" of TQP and I defended this claim at a conference in Granada [16]. During that meeting, however, I met Julián Velarde from the University of Oviedo who tried to convince me that TPQ had been discovered more than 50 years *before* Leibniz by *Caramuel*. The aim of this paper is to check whether my Spanish colleague is right.

2 Caramuel's Logical Work

Juan Caramuel y Lobkowitz was born in Madrid in 1606 and he passed away in Vigevano, Italy, in 1682. According to Wikipedia, he had a lot of intellectual interests and talents. Already as a child he was occupied with difficult mathematical problems and later on he worked as a Catholic cleric, philosopher, theologian, astronomer, and mathematician. He is said to have published more than 260 books, of which, however, "only little was of lasting importance".[5] His *logical* oeuvre basically consists of two books only: *Rationalis et realis philosophia* of 1642 and *Theologia rationalis* of 1654.[6]

In the standard works on the history of logic, Caramuel is either ignored or grossly underestimated. Thus he isn't mentioned at all in Bocheński's [3] or in Kneale's [12], while in Risse's survey of *The Logic of Modern Age*, [22, pp. 351–354], he is discredited as "one of the weirdest thinkers whose ideas were full of wit rather than correct". Risse disqualifies Caramuel's logic as "abounding with idiosyncrasies", and he disregards *Theologia rationalis* simply because of the inappropriateness of its title.

[2] Cf. [7, pp. 297–319].

[3] Cf. [26, p. 9], where Venn constrains that "neither Mr. Bentham nor Mr. Solly seem to me to have understood exactly the sense in which their scheme was to be interpreted".

[4] Cf. my contribution to the Montreux conference on the Logical Square [14].

[5] Quoted according to the German Wikipedia article "Juan Caramuel y Lobkowitz", online access on January 17, 2014: "Allerdings behielt kaum etwas davon dauerhafte Bedeutung."

[6] Cf. the Online-bibliography organized by Jacob Schmutz http://web.archive.org/web/20070703162930/http://www.ulb.ac.be/philo/scholasticon/bibcaramuel.html.

It was not before the 20th century that a certain rehabilitation of Caramuel's logical work took place. One of the earliest appreciations was expressed by Pastore [18] who argued, though not very convincingly, that Caramuel invented a version of TQP comparable, if not superior, to that of Hamilton. Similarly, Velarde [25] maintained that Caramuel's TQP is even "more complete and more systematic" than that of Hamilton.[7] These claims will have to be discussed in Sect. 7 below.

Moreover, in recent years, Caramuel's logic has been investigated quite extensively by the Czech authors Karel Berka, Petr Dvořák, and Stanislav Sousedík. The main focus of their research lies on Caramuel's sophisticated theory of "relational logic" which deals with "oblique" propositions such as 'Every man commits some sin' or 'Every elephant is greater than every ant'. This interesting—and perhaps most original—part of Caramuel's logic must, however, remain out of consideration here. The reader is referred to Dvořák's dissertation [8] and to his article for the *Handbook of the History of Logic*, [9].

My reconstruction of Caramuel's TQP is based upon a microfilm copy of *Theologia rationalis* (*Theol*, for short) which was kindly provided to me by the "Staatsbibliothek zu Berlin". Quotations will refer to this edition either by page or by paragraph. For the sake of precision, the following symbols will be used: $\neg$ for negation, $\wedge$ for conjunction, $\vee$ for disjunction, $\rightarrow$ for implication, and $\leftrightarrow$ for equivalence.

3 Principles of "Aristotelian" Syllogistic: Subalternation, Opposition, and Conversion

The traditional theory of the syllogism deals with the four categorical forms of universal affirmative (UA), universal negative (UN), particular affirmative (PA), and particular negative (PN) propositions:

UA	SaP	(Every S is P),	UN	SeP	(No S is P),
PA	SiP	(Some S are P),	PN	SoP	(Some S aren't P).

The principle of *subalternation* states that a universal proposition entails the corresponding particular proposition, i.e. UA entails PA, and UN entails PN:

$$\text{Sub 1}\quad \text{SaP} \rightarrow \text{SiP}, \qquad \text{Sub 2}\quad \text{SeP} \rightarrow \text{SoP}.$$

Caramuel was well aware of these principles although he never formulated them in such a *general* form, but only by way of *examples*. Thus the validity of Sub 2 is stated in § DCLXVIII by pointing out that "these proposition, 'No man is running'; 'Some man is not running' are subalterns".[8]

The two main principles of *opposition* say that the PN is the "contradictory" of the UA, and similarly the PA is the negation of the UN:

$$\text{Opp 1}\quad \text{SoP} \leftrightarrow \neg\text{SaP}, \qquad \text{Opp 2}\quad \text{SiP} \leftrightarrow \neg\text{SeP}.$$

[7] Cf. [25, p. 274]: "La teoriá de Caramuel sobre la cuantificación del predicado no sólo es anterior a la de Hamilton sino que es más completa y sistemática."

[8] Cf. *Theol*, p. 218. In addition to this "subalternation of the *subject*", Caramuel also admits of a "subalternation of the *predicate*": "*Aliquod animal est omnis homo; Ergo aliquod animal est aliquis homo* [...]" (*Theol*, p. 73).

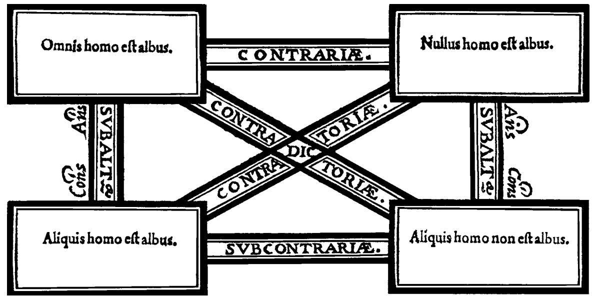


Fig. 1 Traditional square of opposition, from Caramuel [5]

Furthermore, the universal propositions SaP and SeP are said to be *contraries* in the sense that they can't both be true while it is possible that they are both false. Finally, the particular propositions SiP and SoP are considered as *subcontraries* which means that they cannot both be false while they may well both be true. Caramuel subscribed to this traditional view in several places of his work. Thus in §§ CCXXIX + CCXXX he gives clear definitions of the notions 'contrary', 'contradictory', and 'subcontrary';[9] and the entire theory of subalternation plus opposition is summarized on p. 69 in the shape of the traditional *Square of Opposition* (see Fig. 1).

The third group of principles is concerned with the issue of *conversion* where the tradition distinguishes between "simple" and "accidental" conversion.[10] The UN as well as the PA can be converted "simpliciter":[11]

$$\textsc{Conv}\ 1\quad \text{SeP} \rightarrow \text{PeS},\qquad \textsc{Conv}\ 2\quad \text{SiP} \rightarrow \text{PiS}.$$

The UA, however, can only be converted "per accidens", i.e. by "weakening" the quality of the proposition:[12]

$$\textsc{Conv}\ 3\quad \text{SaP} \rightarrow \text{PiS}.$$

The stronger form

$$\textit{Conv 4}\quad SaP \rightarrow PaS$$

is not generally valid (and therefore put in italics). As Caramuel remarks in § DCLXXI,[13] the UN also admits of an "accidental" conversion:

$$\textsc{Conv}\ 5\quad \text{SeP} \rightarrow \text{PoS}.$$

---

9 Cf. *Theol*, p. 69: "Duae contrariae possunt et solent esse simul falsae: non simul verae. Duae subcontrariae solent esse simul verae, non simul falsae. Duae Contradictoriae, nec simul falsae, nec simul verae."

10 At the beginning of § DCLXX, Caramuel defines conversion as a truth-conserving commutation of subject into predicate and predicate into subject, of which there exist three types: "*Primò simpliciter, Secundò per accidens, Tertiò per contrapositionem*".

11 Again, Caramuel doesn't state these laws in an abstract form but only by way of examples. "Universalis negativa converti potest simpliciter, ut si diceres *Nullus homo est argenteus: Nullum argenteum est homo* [...] Particularis affirmativa in omni materiâ convertitur simpliciter; ut *Aliquis homo est album; Aliquod album est homo*" (*Theol*, p. 219).

12 Cf. § DCLXXI: "PRIMO igitur Universalis affirmativa [...] in omni autem materiâ potest converti per accidens. Ut *Omnis homo est animal*: *Aliquod animal est homo*. *Omnis homo est albus*; *Aliquod album est homo*."

13 E.g. "SECUNDO, Universalis negativa [...] poterit etiam converti per accidens: ut *Nullus homo est lapis*: *Aliquis lapis non est homo*".

Let it be mentioned in passing that, in view of the principles of subalternation, the laws of “accidental” conversion can easily be derived from corresponding laws of “simple” conversion.[14] Let it further be noted that, according to the tradition, the PN cannot be converted at all (except for a conversion by *contraposition* to be discussed in Sect. 4). That means, in particular, that the following principle is *invalid*:

$$\textit{C}\textsc{onv}\ 6 \quad SoP \rightarrow PoS.$$

The invalidity may be illustrated by the following example: *Some female are not pregnant*; but it doesn’t follow that *Some pregnant are not female*.[15]

4 Principles of “Scholastic” Syllogistic: Contraposition and Obversion

According to the terminology of [24], the “Scholastic” syllogistic differs from “Aristotelian” syllogistic basically in that it admits *negative* (or “infinite”) *concepts* as subject and predicate of a categorical proposition. Here the *negation* of a *concept* A, Not-A, shall be symbolized by ∼A (in Latin “non-A”). Just like in propositional logic, where $\neg\neg\alpha$ is equivalent to α, the operator of concept negation satisfies the law of double negation

$$\textsc{Neg}\ 1 \quad \sim\sim \mathrm{A} = \mathrm{A}.$$

Further properties of the operator ∼ emerge from principles of contraposition and obversion. In *Grammatica Audax*, Caramuel explains that contraposition is a logical inference where the “quantity” of the proposition is maintained while the “extreme terms”, i.e. subject and predicate, are negated before their position is exchanged.[16] Thus there exist four possible laws of contraposition:

$$\textsc{Contra}\ 1 \quad \mathrm{SaP} \rightarrow \sim\mathrm{Pa}\sim\mathrm{S}, \qquad \textit{C}\textsc{ontra}\ 2 \quad SeP \rightarrow \sim Pe \sim S,$$

$$\textit{C}\textsc{ontra}\ 3 \quad SiP \rightarrow \sim Pi \sim S, \qquad \textsc{Contra}\ 4 \quad \mathrm{SoP} \rightarrow \sim\mathrm{Po}\sim\mathrm{S}.$$

In § DCLXXI, Caramuel systematically investigates all three types of conversions (“simpliciter”, “per accidens”, and “per contrapositionem”) for the different categorical forms. He subscribes to CONTRA 1 by giving the example: “*Omnis homo est substantia*: [Ergo:] *Omne non substantia est non homo*”, and he further tries to defend CONTRA 4 as follows: “Ex propositione, *Antichristus non est bonus*, infertur, *Aliquod non bonum non est Antichristus*.” But this example is rather inapt to demonstrate the validity of CONTRA 4. First of all, the inference does not fit into the schema of CONTRA 4 since the premise ‘Antichrist is not good’ lacks a quantifier expression like ‘some’. Furthermore, ‘Antichrist’

[14]For SeP → PeS by CONV 1 and PeS → PoS by SUB 1, hence CONV 4. Similarly, SaP → SiP by SUB 2 and SiP → PiS by CONV 2, hence CONV 3.

[15]A more powerful proof of the invalidity of *CONV 6* consists in the observation that, if *CONV 6* were valid, then *CONV 4* would become provable. For assume that SaP, then, because of OPP 1, ¬SoP. Now if *CONV 6* were valid, then PoS would entail SoP; hence, by logical contraposition, ¬SoP would entail ¬PoS, so that (again by OPP 1) we would obtain PaS.

[16]“CONTRAPOSITIO afficit omnes propositiones & est quaedam Conversio, quae servat propositionis quantitatem, convertit extrema, & utrumque infinitat” (§ CCXLV, p. 73).

appears to be a *singular term* rather than a (grammatical) *predicate*. As Caramuel remarks, it would be better to abstain from such "barbarian locutions" since a Logician should express himself in such a way that he can be understood by the audience.[17]

A proof of CONTRA 4 is easily obtained if one observes that the fundamental principle CONTRA 1 may be strengthened into an *equivalence*:

$$\text{CONTRA 1*} \quad SaP \leftrightarrow \sim Pa \sim S.$$

For since SaP → ∼Pa∼S, so also ∼Pa∼S → ∼∼Sa∼∼P, hence by the law of double negation, ∼Pa∼S → SaP. But from the thus proved CONTRA 1* one obtains that the negation of SaP must be equivalent to the negation of ∼Pa∼S. Hence (by OPP 1)

$$\text{CONTRA 4*} \quad SoP \leftrightarrow \sim Po \sim S.$$

Let's now consider the remaining principles *CONTRA 2* and *CONTRA 3*. The idea of the foregoing proof may be applied straightforwardly to the case, e.g. of SeP. Thus, *if CONTRA 2* were valid, then the strengthened principle:

$$\textit{CONTRA 2*} \quad SeP \leftrightarrow \sim Pe \sim S$$

would hold as well. But then also the negations of these formulas would be equivalent:

$$\textit{CONTRA 3*} \quad SiP \leftrightarrow \sim Pi \sim S.$$

Now, towards the end of § DCLXXI, Caramuel appears to maintain that *CONTRA 3* is valid, for he explains that besides the UA and the PN also the PA admits of conversion by contraposition.[18] But this might be a slip of the tongue, for in the preceding list of all admissible conversions there is not only no mention of a conversion by contraposition in the case of UN (which may be taken as strong evidence that Caramuel *knew* that CONTRA 2 is invalid). But in that list there is also no mention of a conversion by contraposition in the case of PA, so that—despite the above quotation—Caramuel may have believed that *CONTRA 3* is *invalid*.

On the other hand, in another passage devoted to the topic of contraposition, viz. in § CCXLV of *Grammatica Audax*, Caramuel put forward not only the correct inference "*Omnis homo est animal. Ergo omne non-animal est non-homo*", but also the incorrect inference "*Aliquis homo est animal: Ergo aliquod non-animal est non-homo*" which indicates that he may have considered *CONTRA 3* as *valid*, after all. Be this as it may, the crucial principle *CONTRA 3* (and hence also *CONTRA 2*) *is* invalid, as becomes clear from the following counterexample: Let be S = 'is blond' and P = 'is male'. Suppose that 'All female are blond' but also 'Some male are blond'. Then, of course, 'Some blond are male', SiP, but it doesn't follow that ∼Pi∼S, i.e. that 'Some non-males (i.e. some females) are not-blond'!

The last set of principles to be considered in this section concerns so-called *obversion*.

$$\text{OBV 1} \quad SeP \leftrightarrow Sa\sim P, \qquad \text{OBV 2} \quad SoP \leftrightarrow Si\sim P,$$
$$\text{OBV 3} \quad SiP \leftrightarrow So\sim P, \qquad \text{OBV 4} \quad SaP \leftrightarrow Se\sim P.$$

---

[17] "Sed hae barbarae locutiones sunt, à quibus debemus abstinere; quia Dialecticus debet saltem ita loqui, ut ab audientibus intelligatur." More appropriate examples for CONTRA 4 may, however, be found in other places of *Theol*, e.g. p. 73, § CCXLV: "Aliquis homo non est leo: Ergo aliquod non leo non est non homo".

[18] "Et tandem, [...] universalem affirmativam, & particularem *tam affirmativam quam negativam* [!] converti per contrapositionem."

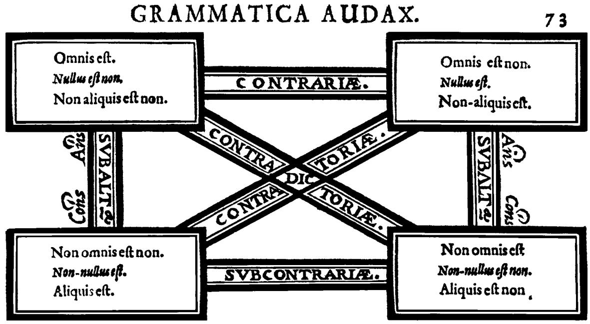


Fig. 2 Caramuel's extended square of opposition, from Caramuel [5]

Caramuel dealt with this issue under the heading "Aequipollentia". In § CCXXX of *Grammatica Audax*, he illustrates the validity of Obv 1 by means of the example 'No man is a stone' which can be transformed first into '*Every man isn't a stone*' and then into 'Every man is a not-stone'.[19] Similarly, the validity of Obv 2 and Obv 3 is illustrated by transforming "Aliquis homo non est lapis" into "Aliquis homo est non lapis", and "Aliquis homo est animal" into "Aliquis homo non est non-animal", respectively.[20] Finally, Obv 4 is at least *indirectly* stated by Caramuel when he transforms "Omnis homo est animal" into "Non aliquis homo est non-animal" (instead of 'Nullus homo est non-animal').[21] These results are integrated in Fig. 2.

For the following investigations, it will be helpful to elaborate this Square a bit further. Let us start with Caramuel's three elliptic representations of the UA, i.e. "Omnis est", "Nullus est non", and "Non aliquis est non". The first, of course, may unambiguously be explicated as 'Omnis S est P', but the second and third admit of two detailed definitions according to whether 'non' is taken to negate the "copula" or the predicate P. In the former case, one better uses the variant 'Nullus S non est P', while the latter variant is best formalized as 'Nullus S est ∼P'. Similarly, "Non aliquis est non" can be understood either as 'Non Aliquis S est ∼P' or as 'Non Aliquis S non est P' where the 'non' in front of the formula must clearly be understood as propositional negation. So all together we get *five* different ways to express the content of a UA:

UA1	Omnis S est P
UA2a	Nullus S est ∼P
UA2b	Nullus S non est P
UA3a	¬(Aliquis S est ∼P)
UA3b	¬(Aliquis S non est P)

Similarly, the elliptic representations of the PN, "Non omnis est", "Non nullus est non", and "Aliquis est non" can be rendered more precise as:

PN1	¬(Omnis S est P)
PN2a	¬(Nullus S est ∼P)

[19] Cf. p. 69: "AEQUIPOLLENTIA *est Enunciationum aequivalentia*: haec enim videntur idem significare [...] Nullus homo est lapis. *Omnis homo non est lapis*. *Omnis homo est non-lapis*."

[20] A general *proof* of Obv 2 and Obv 3 is easily obtained from Obv 1, Obv 4 by means of the laws of opposition, e.g. SoP → ¬(SaP) (by Opp 1) → ¬(Se ∼ P) (by Obv 4) → Si ∼ P (by Opp 2).

[21] Obv 4 easily follows from Obv 1 by substituting ∼P for P.

PN2b	¬(Nullus S non est P)
PN3a	Aliquis S est ∼P
PN3b	Aliquis S non est P

In the case of the UN, however, the explication of "Omnis est non", "Nullus est", and "Non aliquis est" yields only *four* formulas:

UN1a	Omnis S est ∼P
UN1b	Omnis S non est P
UN2	Nullus S est P
UN3	¬(Aliquis S est P)

And, in the case of PA, "Non omnis non", "Non nullum est", and "Aliquis est" correspondingly have to be understood as:

PA1a	¬(Omnis S est ∼P)
PA1b	¬(Omnis S non est P)
PA2	¬(Nullus S est P)
PA3	Aliquis S est P

This asymmetry between *four* versions of UN and PA, and *five* versions of UA and PN, is due to the fact that certain other paraphrases have been overlooked so far. As Caramuel points out in § DCLXXIX, the *singular* proposition 'Petrus est homo' might as well be formulated as 'Petrus non est non-homo'. According to the basic idea of obversion, the latter can first be transformed into 'Petrus est non-non-homo' and then be simplified, by NEG 1, to 'Petrus est homo'.[22]

Similarly, the *universal affirmative* proposition 'Omnis S est P' can be paraphrased by means of two additional negations as 'Omnis S non est ∼P'. Therefore, one obtains also 'Non Omnis S non est ∼P' as another paraphrase of PN. Analogously, 'Nullus S est P' can be expressed in a more complicated way as 'Nullus S non est ∼P', from which one gets ¬(Nullus S non est ∼P) as another version of PA. Similarly, PA3 might be transformed into 'Aliquis S non est ∼P' so that ¬(Aliquis S non est ∼P) emerges as another paraphrase of UN. In this way, the symmetry between all four types of categorical propositions is re-established. Each one possesses *six* different variants, three of which have a positive predicate P, the other three a negative predicate, ∼P; furthermore, three have a positive "copula" 'est', the other three the "negative copula" 'non est':

UA1a	Omnis S est P	UN1a	Omnis S est ∼P
UA1b	Omnis S non est ∼P	UN1b	Omnis S non est P
UA2a	Nullus S est ∼P	UN2a	Nullus S est P
UA2b	Nullus S non est P	UN2b	Nullus S non est ∼P
UA3a	¬(Aliquis S est ∼P)	UN3a	¬(Aliquis S est P)
UA3b	¬(Aliquis S non est P)	UN3b	¬(Aliquis S non est ∼P)

[22]Cf. *Theol*, p. 221, where Caramuel first explains two laws of "Logicae Metamorphoseos": "Prima, *Propositio asserens, cuius praedicatum est infinitum, converti in negativam potest* [. . .] verbi gratia [ex] hâc *Homo est non equus* [. . .] erit *Homo non est equus* [. . .] Secunda, *Cuiuscumque propositionis negativae negatio transeat ad praedicatum* [. . .] ut patet in hâc *Petrus non est canis*, quae transformatur in hanc, *Petrus est non-canis*." With the help of these laws he is then able to transform '*Petrus non est non homo*' as follows: "Respondeo debere juxta secundam regulam ac propterea̋ in hanc, *Petrus est non non-homo*, sive quod idem est, *Petrus est homo*;"

PA1a	¬(Omnis S est ∼P)	PN1a	¬(Omnis S est P)
PA1b	¬(Omnis S non est P)	PN1b	¬(Omnis S non est ∼P)
PA2a	¬(Nullus S est P)	PN2a	¬(Nullus S est ∼P)
PA2b	¬(Nullus S non est ∼P)	PN2b	¬(Nullus S non est P)
PA3a	Aliquis S est P	PN3a	Aliquis S est ∼P
PA3b	Aliquis S non est ∼P	PN3b	Aliquis S non est P

Let us arrange these 24 expressions in a Square of Opposition but simplify the result by dropping all propositions beginning with a '¬':

UA	Omnis S est P	Omnis S est ∼P	UN
	Omnis S non est ∼P	Omnis S non est P	
	Nullus S est ∼P	Nullus S est P	
	Nullus S non est P	Nullus S non est ∼ P	
PA	Aliquis S est P	Aliquis S est ∼ P	PN
	Aliquis S non est ∼P	Aliquis S non est P	

Next, in view of the principles of obversion, all propositions containing a negated predicate ∼P, can be dropped again. Thus one obtains Fig. 3 which shall serve as a basis for the subsequent investigations.

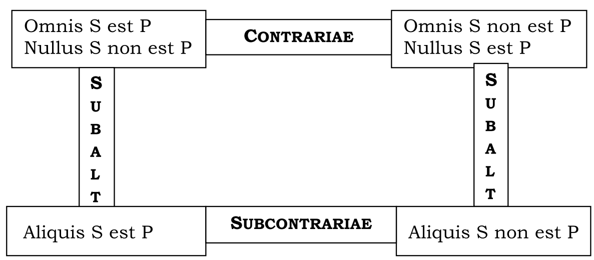


Fig. 3 Revision of Caramuel's extended square of opposition. (For typographical reasons, the diagonal arrows symbolizing that the corresponding propositions are negations of each other are omitted here)

5 Caramuel's TPQ as an Extension of the Theory of Conversion

At the beginning of § DCLXXII, Caramuel explains that the theory of conversion as summarized in the six articles of § DCLXXI only describes "what had come to the mind of the Old Logicians" while he now wants to explore this difficult issue in more detail.[23] For this he makes a complicated subdivision of "simple" conversions into four classes (1a), (1b), (2), and (3):

Simple
- Partial
  - (1) Transubstantiation
    - (a) material
    - (b) formal
  - (2) Transfiguration
- Total
  - (3) Transubstantiation & Transfiguration

However, in the subsequent §§ DCLXXIII-DCLXXVII only (1a), (2) and (3) are further investigated while (1b), i.e. "formal transubstantiation", is no longer mentioned. There-

[23]Cf. *Theol*, p. 219: "Hae dicta sunt ad mentem Veterum sed & placet adhuc penitius hanc difficultatem contemplari".

Fig. 4 Categorical forms with explicitly QP

Omnis S est aliquis P,	Omnis S non est aliquis P,
Nullus S non est aliquis P,	Nullus S est aliquis P,
Aliquis S est aliquis P,	Aliquis S non est aliquis P.

fore, our reconstruction of Caramuel's TQP focuses on these three §§ which offer enough interpretational problems, anyway.

In the subsequent Sects. 5.1–5.3, we will be confronted with some strange and unorthodox propositions which Caramuel obtained by his extended theory of conversion. Unfortunately, the text of *Theol* gives us almost no hint how such propositions as 'Omnis S est omnis P', 'Aliquis S est omnis P', or also 'Nullus S est nullus P' are to be understood. Let us try to develop at least some plausible hypotheses concerning the truth-conditions of these unorthodox propositions. From a purely syntactic point of view, the unqualified predicate P of the "colloquial" formulation of the categorical forms can be modified by adding any of the quantifier expressions 'aliquis', 'nullus', or 'omnis'. In the first case we obtain the schema of Fig. 4.

This schema obviously represents only a variant of Fig. 3. For just like 'Every S is a P' and 'Some S is a P' may be rephrased as 'Every S is some P' and 'Some S is some P', respectively, so also Caramuel's 'Omnis S est aliquis P' and 'Aliquis S est aliquis P' seem to be synonymous with 'Omnis S est P' and 'Aliquis S est P'. Let it be mentioned in passing that these formulations (partially) support the traditional doctrine, advocated, e.g. in the *Logic of Port Royal*, that "the predicate of an *affirmative* proposition is always *particular*"[24] (viz., modified by 'aliquis' or by 'some').

Furthermore, it seems very plausible to assume that Caramuel accepted the following *generalization* of the laws of opposition to propositions with an explicitly quantified predicate:

OPP 1* ¬(Omnis S est aliquis P) ↔ Aliquis S non est aliquis P
¬(Omnis S est omnis P) ↔ Aliquis S non est omnis P
¬(Omnis S est nullus P) ↔ Aliquis S non est nullus P

OPP 2* ¬(Nullus S est aliquis P) ↔ Aliquis S est aliquis P
¬(Nullus S est omnis P) ↔ Aliquis S est omnis P
¬(Nullus S est nullus P) ↔ Aliquis S est nullus P

The negation of the PA 'Aliquis S est aliquis P' then becomes equivalent to 'Nullus S est aliquis P', and the negation of the UA, 'Omnis S est aliquis P' becomes equivalent to 'Aliquis S non est aliquis P'. Therefore, the entire Fig. 4 represents just a variant of Fig. 3. But in Fig. 4 also the negative propositions have a "particular" predicate so that the second part of the traditional doctrine according to which "the predicate of a *negative* proposition is always taken as *universal*" is not corroborated.[25]

The next schema appears to describe quite different states of affairs which cannot be paraphrased in terms of the orthodox propositions UA, UN, PA, and PN (see Fig. 5).

[24] Cf. [1, p. 183]: "L'attribut d'une proposition affirmative [. . .] est toûjours consideré comme pris particulierement [. . .] L'attribut d'une proposition negative est toûjours pris generalement".

[25] In English, the UN might be paraphrased as 'No S is any P' or also as 'Every S isn't any P; and the PN is quite naturally formulated as 'Some S isn't any P'. Here then the predicate may be considered as taken "universally". The Latin counterpart of 'any' would be 'ullus', but Caramuel nowhere considered this quantifier.

Fig. 5 Unorthodox forms with explicitly QP

Omnis S est omnis P,	Omnis S non est omnis P,
Nullus S non est omnis P,	Nullus S est omnis P,
Aliquis S est omnis P,	Aliquis S non est omnis P.

We will not try to make a plausible guess about the intended meaning of the unorthodox propositions of Fig. 5 until we have finished our analysis of Caramuel's TQP in Sect. 5.3. Right now, however, it should be observed that these propositions form another Square of Opposition. For we may assume that Caramuel accepted the following *generalization* of the laws of subalternation:

SUB 1*	Omnis S est aliquis P $\rightarrow$ Aliquis S est aliquis P
	Omnis S est omnis P $\rightarrow$ Aliquis S est omnis P
	Omnis S est nullus P $\rightarrow$ Aliquis S est nullus P
SUB 2*	Nullus S est aliquis P $\rightarrow$ Aliquis S non est aliquis P
	Nullus S est omnis P $\rightarrow$ Aliquis S non est omnis P
	Nullus S est nullus P $\rightarrow$ Aliquis S non est nullus P

Hence 'Omnis S est omnis P' entails 'Aliquis S est omnis P' (by SUB 1*) which in turn entails $\neg$(Nullus S est omnis P) (by OPP 2*). That means that the universal propositions in Fig. 5 are contraries which cannot both be true. But on the other hand, if, e.g. 'Aliquis S non est omnis P' is not true, then 'Omnis S est omnis P' must be true (by OPP 1*), and hence also 'Aliquis S est omnis P' is true (according to SUB 1*). In other words, the two particular propositions of Fig. 5 are subcontraries, which cannot both be false.

Let us now consider the last schema depicted in Fig. 6.

Fig. 6 Categorical forms with emphatically QP

Omnis S est nullus P,	Omnis S non est nullus P,
Nullus S non est nullus P,	Nullus S est nullus P,
Aliquis S est nullus P,	Aliquis S non est nullus P.

To say, e.g. that 'Some S is *no* P' seems to be just an emphatic way to express that 'Some S isn't *any* P', i.e. 'Some S isn't P'. Similarly, 'Some S *isn't no* P' is just an emphatic way to express that 'Some S *is* (a) P', after all. These linguistic observations also appear to hold for the Latin expressions of Fig. 6. Therefore, the propositions on the left hand side turn out to be just unusual, emphatic versions of the UN and PN, respectively; while the propositions on the right hand side represent emphatic versions of the UA and PA, respectively.

While Fig. 5 contains four unorthodox propositions, the truth-conditions of which still have to be explored, the following combination of Figs. 4 and 6 represents the usual Square of Opposition (albeit in an unusual cloth) (see Fig. 7).

UA	Omnis S est aliquis P	Omnis S non est aliquis P	UN
	Nullus S non est aliquis P	Nullus S est aliquis P	
	Omnis S non est nullus P	Omnis S est nullus P	
	Nullus S est nullus P	Nullus S non est nullus P	
PA	Aliquis S est aliquis P	Aliquis S non est aliquis P	PN
	Aliquis S non est nullus P	Aliquis S est nullus P	

Fig. 7 Categorical forms with QP

5.1 Material Transubstantiation

In § DCLXXIII,[26] Caramuel explains the notion of the substance ("substantia materialis") of a term. In a proposition like 'Omnis homo est aliquod animal', the substance of the *subject* is 'homo' while the substance of the *predicate* is 'animal'. The other expressions denote the quantity ("quantitas"): 'omnis' is the quantity of the subject and 'aliquod' that of the predicate. The logical process of "material transubstantiation" is then defined as taking place whenever the substance of the subject is exchanged with that of the predicate while both quantities remain unchanged.

Now Caramuel in general considers not only the traditional quantities "universal" (U) and "particular" (P), but also "singular" (S) and "indefinite" (I). However, indefinite terms shall be ignored here because they do not seem to possess a fourth quantity besides U, P, and S. The characterization of a term (or a proposition) as "indefinite" just indicates that it is left open, or *undetermined*, which quantity the term (or the proposition) actually has.[27] For reason of space, quantity S will also remain out of consideration here, although Caramuel's theory of singular propositions certainly is worthwhile being studied elsewhere. Hence out of his lists of 16 affirmative and 16 negative combinations, only the following eight have to be examined:

TRANSSUB 1 "Omnis homo est omne animal; [ergo:]
Omne animal est omnis homo"

TRANSSUB 2 "*Omnis homo est aliquod animal*; [ergo:]
Omne animal est aliquis homo"

TRANSSUB 3 "Aliquis homo est omne animal; [ergo:]
Aliquod animal est omnis homo"

TRANSSUB 4 "*Aliquis homo est aliquod animal*; [ergo:]
Aliquod animal est aliquis homo"

TRANSSUB 5 "*Nullus homo est omnis lapis*; [ergo:]
Nullus lapis est omnis homo"

TRANSSUB 6 "*Nullus homo est aliquis lapis*; [ergo:]
Nullus lapis est aliquis homo"

TRANSSUB 7 "*Aliquis homo non est omnis lapis*; [ergo:]
Aliquis lapis non est omnis homo"

TRANSSUB 8 "*Aliquis homo non est aliquis lapis*; [ergo:]
Aliquis lapis non est aliquis homo"

As the diligent reader may perhaps have noticed, some, but not all, premises and some, but not all, conclusions of the above principles are set in italics. As Caramuel explains towards the end of § DCLXXIV, all propositions printed in *italics* are (taken to be) *true*, while all others are (taken to be) false.[28] Therefore, TRANSSUB 2 must be *invalid*, since

[26]The paragraph starting at the bottom of the right column of p. 219 bears the headline "De Transsubstantiatione materiali". While Caramuel normally writes the number of a § on the margin, # DCLXXIII is missing here!

[27]In § CCCCLXVIII, Caramuel defines an indefinite proposition as "one which has an undetermined quality, i.e. which has no quantifier expression". And he goes on to explain that, e.g. 'Homo est animal' is sometimes equivalent to a universal and sometimes equivalent to a particular proposition.

[28]"Maioris claritatis gratiâ propositiones veras à falsis distinxi charactere: illas enim scripsi cursivo, has romano." Pastore [18] appears to have overlooked this important point even though Caramuel had written

from the true premise 'Every man is some animal' the false conclusion 'Every animal is *some* man' cannot be inferred "salva veritate". A fortiori the general scheme underlying this inference, *CONV 4*, is not truth-conserving. This is just the core of the traditional wisdom that the UA cannot be converted "simpliciter", but only "per accidens".

But what about the other seven cases? Does the mere fact that the "salva veritate"-requirement is *not violated*, entitle us to conclude that the inference in question is *valid* (or at least considered by Caramuel as valid)? TRANSSUB 4 and TRANSSUB 6 seem to speak in favor of such an assumption. From the true premise 'Aliquis homo est aliquod animal' it follows that 'Aliquod animal est aliquis homo', since, according to CONV 2, SiP can generally be converted into PiS. Similarly, from the true premise 'Nullus homo est aliquis lapis' one can infer 'Nullus lapis est aliquis homo', since, according to the CONV 1, SeP may be converted into PeS.

Unfortunately, in the case of TRANSSUB 8, the situation is different. In the given *example*, the premise 'Some man is not some stone' may be converted—so to speak "accidentally"—into the conclusion 'Some stone is not some man', since both propositions *happen to be true*. Yet, as was argued in connection with *CONV 6* above, the PN SoP cannot *generally* be converted!

Towards the end of § DCLXXIV, Caramuel maintains that his table determines "where and when propositions may be transubstantiated *salva veritate*", and that it is therefore not necessary to state general rules.[29] But this view is all too optimistic! Except for the simple case of a true premise, P_1, and a false conclusion, C_1, the mere assignment of truth-values to a particular *example* does not suffice to determine whether the general inference $P \rightarrow C$ is valid or not. Each other distribution of truth values, say $P_2 = t$ and $C_2 = t$ (as in TRANSSUB 5 and 7), or $P_3 = f$ and $C_3 = t$ (as in TRANSSUB 3), or $P_4 = f$ and $C_4 = f$ (as in TRANSSUB 1), leaves it open whether there exist counterexamples such that $P_5 = t$ but $C_5 = f$! So on the basis of examples alone, we can't tell whether the principles TRANSSUB 1, 3, 5, and 7 are valid. Generalizing from Caramuel's examples, these principles take the following form:

TRANSSUB 1*	Omnis S est omne P → Omnis P est omne S
TRANSSUB 3*	Aliquis S est omne P → Aliquis P est omne S
TRANSSUB 5*	Nullus S est omne P → Nullus P est omne S
TRANSSUB 7*	Aliquis S non est aliquis P → Aliquis P non est aliquis S

Now there is at least an *indirect* argument to show that TRANSSUB 3* can't be *generally valid*. Since Caramuel claims to have put all and only the true propositions in italics, the premise of TRANSSUB 3, 'Aliquis homo est omne animal', should be taken as false while the conclusion 'Aliquod animal est omnis homo' is taken to be true. But then one may argue that TRANSSUB 3* is *not formally valid*, since if we had started instead with the *true* proposition 'Aliquod animal est omnis homo' as a *premise*, the inference schema would have lead to the *false conclusion* 'Aliquis homo est omne animal'!

on the margin: "Nota id bene." Otherwise Pastore's reconstruction of the theory of transubstantiation consists only in listing the 16 possible combinations of quantities without reproducing also the corresponding propositions. Furthermore, he entirely dismisses § DCLXXIV which, allegedly, "only repeats the list of the 16 quoted cases". Pastore overlooks that Caramuel here deals with the conversion of *negative* propositions!

[29] "Ex dictis constat ubi & quando possint propositiones transsubstantiari servata veritate; ac ideo non est opus speciales regulas multiplicare."

Next observe that in view of the generalized laws of oppositions, principle TRANSSUB 5* is provably equivalent to TRANSSUB 3* and hence must be invalid, too. For assume, e.g. that TRANSSUB 5* is *valid* and that the antecedent of TRANSSUB 3*, 'Aliquis S est omne P', is *true*. Because of OPP 2*, 'Nullus S est omne P' then must be *false*. But according to TRANSSUB 5*, the assumption 'Nullus P est omne S' entails that 'Nullus S est omne P'. Hence by way of logical contraposition, the negation of the latter, i.e. ¬(Nullus S est omne P) entails ¬(Nullus P est omne S), i.e. (again by OPP 2*) 'Aliquis S est omnis P'. Hence, if TRANSSUB 5* were valid, so would be TRANSSUB 3*!

As the reader may easily verify, also principles TRANSSUB 1* and TRANSSUB 7* are provably equivalent to each other. So the only open question concerning the theory of "transubstantiation" is whether TRANSSUB 1* (or, equivalently, TRANSSUB 7*) is valid or not. Caramuel's assignment of truth values to the particular *examples* fails to settle this issue. Let us therefore consider the other parts of his theory of QP.

5.2 *Transfiguration*

Caramuel speaks of a transfiguration when the two quantifier expressions are interchanged while the "substances", i.e. subject and predicate, remain unchanged. Since a proposition in which the quantity of the subject is the same as the quantity of the predicate yields an identical inference (e.g. from 'Aliquid S est aliquis P' to 'Aliquis S est aliquis P'), the non-trivial principles reduce to $4 \times 3 = 12$ affirmative and 12 negative instances. Neglecting, again, the quantities I and S, only the following four inferences remain to be investigated here:

TRANSFIG 1 "*Omnis homo est aliquod animal*
Aliquis homo est omne animal"
TRANSFIG 2 "Aliquis homo est omne animal
Omnis homo est aliquod animal"
TRANSFIG 3 "*Nullus homo est aliquis lapis*
Aliquis homo est nullum lapis"
TRANSFIG 4 "*Aliquis homo non est omnis lapis*
Omnis homo non est aliquis lapis"

According to Fig. 4, the premise of TRANSFIG 1, 'Omnis homo est aliquod animal', is just another expression of the "colloquial" UA 'Omnis homo est animal'. Hence this proposition has rightly been put by Caramuel in italics in order to indicate that it is *true*. The unorthodox conclusion 'Aliquis homo est omne animal', in contrast, is printed "romano" and hence considered by Caramuel as *false* (although we don't really know *why* this shall be the case). Anyway, the example in question may be taken to show that the general principle

TRANSFIG 1* Omnis S est aliquis P → Aliquis S est omnis P

is *invalid*. The next principle, TRANSFIG 2, is the direct *inversion* of TRANSFIG 1 since from the unorthodox premise 'Aliquis homo est omne animal' (which Caramuel considers as *false*) the *true* conclusion '*Omnis homo est aliquod animal*' is to be inferred. But this doesn't mean that the general principle:

TRANSFIG 2* Aliquis S est omnis P → Omnis S est aliquis P

is *valid*! For if one considers a *different* example by simply exchanging the concepts ‘homo’ and ‘animal’, one obtains the inference from (i) ‘Aliquod animal est omnis homo’ to (ii) ‘*Omne animal est aliquis homo*’. But, as we have seen,[30] (i) is considered by Caramuel as *true* while (ii) is just a paraphrase of the UA ‘Omne animal est homo’ and hence *false*!

As regards the remaining inferences TRANSFIG 3 and TRANSFIG 4, let it be pointed out here that, for which reason so ever, Caramuel formulated them in an asymmetric way which somehow veils that the two inferences are *inverse* to each other. It seems, however, safe to assume that TRANSFIG 3 may be transformed equivalently into:

TRANSFIG 3′ *Omnis homo non est aliquis lapis*,
Aliquis homo non est omnis lapis[31]

In this way, the symmetry exhibited by all other propositions of § DCLXXV is re-established. But we don’t (yet) understand *why* the strange proposition ‘Aliquis homo non est omnis lapis’, which Caramuel had put in italics, shall be considered as true. As long as we don’t know more about the truth-conditions of these propositions, we cannot decide whether

TRANSFIG 3* Omnis S non est aliquis P → Aliquis S non est omnis P
TRANSFIG 4* Aliquis S non est omnis P → Omnis S non est aliquis P

are generally valid or not.

5.3 Total Conversion

A “total conversion” is defined as simultaneous “transmaterialization & transfiguration” which means that the subject “such as it is” and the predicate “such as it is” change their positions. For example, ‘Nullus homo est aliquis lapis’ is to be converted into ‘Aliquis lapis est nullus homo’. Hence, the “subject such as it is” is the conjunction of the “substance”, ‘homo’, and the quantifier ‘nullus’, and similar the “predicate such as it is” is the conjunction of the “substance” ‘lapis’ and the quantifier ‘aliquis’. Next Caramuel gives a list of the $4 \times 4 = 16$ possibilities of a total conversion of *affirmative* propositions out of which only the following four (with the quantity symbols U and P) shall be considered here.

TOTALCONV 1 “Omnis homo est omne animal
Omne animal est omnis homo”
TOTALCONV 2 “*Omnis homo est aliquod animal*
Aliquod animal est omnis homo”
TOTALCONV 3 “Aliquis homo est omne animal
Omne animal est aliquis homo”

[30] Cf. the discussion of principle TRANSSUB 3 in the previous section.

[31] Cf. the passage from *Theol* p. 69 (already quoted in footnote 22) where Caramuel transforms ‘Nullus homo est lapis’ into ‘Omnis homo non est lapis’.

TOTALCONV 4 "*Aliquis homo est aliquod animal*
Aliquod animal est aliquis homo"

Caramuel explains to have "emphasized the valid and invalid inferences so that one can use the distinction if wanted".[32] The emphasis, again, is achieved by typographical means ("cursivo" vs "romano"). But if the above quotation is taken literally, Caramuel now doesn't want to emphasize the truth-value of the single *propositions* but instead the "truth" or "falsity" of the entire *inference* ("illatio")! This means that TOTALCONV 2 and TOTALCONV 4 are (considered by Caramuel as) *valid* while TOTALCONV 1 and TOTALCONV 3 are (considered by Caramuel as) *invalid*. As a matter of fact, the generalized principle

TOTALCONV 4* Aliquis S est aliquis P $\rightarrow$ Aliquis P est aliquis S

is valid, because this inference represents just a paraphrase of the ordinary conversion principle CONV 2. However, it is not at all clear whether TOTALCONV 2 may be regarded as valid, too. To be sure, in accordance with our previous considerations, not only the premise '*Omnis homo est aliquod animal*' (i.e. the ordinary UA 'Omnis homo est animal') is *true*; but, as we have seen, in Caramuel's opinion also the unorthodox proposition '*Aliquod animal est omnis homo*' shall be taken as *true*. Yet, the truth of these *examples* fails to establish the validity of the generalized principle:

TOTALCONV 2* Omnis S est aliquis P $\rightarrow$ Aliquis P est omnis S

A similar remark applies, *mutatis mutandis*, to TOTALCONV 3! Even if we take it for granted that both the premise 'Aliquis homo est omne animal' and the conclusion 'Omne animal est aliquis homo' are *false*, it remains unclear whether (and why) the generalized inference should be invalid:

TOTALCONV 3* Aliquis S est omnis P $\rightarrow$ Omnis P est aliquis S

As regards TOTALCONV 1, the careful reader may have noted that this principle is the same as TRANSSUB 1 (discussed in Sect. 5.1). *There*, in connection with his theory of "transubstantiation", Caramuel obviously stressed (by the same typographical means of putting propositions in "romano") that *both* 'Omnis homo est omne animal' *and* 'Omne animal est omnis homo' are *false*.[33] But now, within his theory of "total" conversion, he appears to maintain that the *principle itself* is *invalid*.

In order to resolve this difficulty, let me now already introduce a hypothesis which will be further discussed in Sect. 6 below. It is the assumption that Caramuel interprets the unorthodox proposition 'Omnis S est omnis P' always in the "collective" sense, i.e. as saying that the *set* of all S coincides with the set of all P.[34] But as a matter of course, the set of all man is not identical with the set of all animals. More generally, no matter which pair of (different) concepts S and P one chooses, the proposition 'Every S is every P' becomes *false* (for any $S \neq P$).

Nevertheless, the crucial *inference* from, say, 'Omnis homo est omne animal' to 'Omne animal est omnis homo', or from 'Omnis leo est omne animal' to 'Omne animal est omnis

[32]Cf. *Theol*, p. 221: "Haec est Conversio totalis affirmativarum [pro]positionum; notavi bonas & malas illationes, ut possis uti discretione, si velis."

[33]This assignment of truth-values nevertheless leaves open the possibility that the general inference scheme is *valid*.

[34]This is also the sense in which Hamilton [11] interprets his formula 'All S is all P'.

leo' remains *formally* valid. *If* the class of all lions were to coincide with the class of all animals, *then* also conversely the class of all animals would coincide with the class of all lions! That is, the corresponding *conversion principle* must be regarded as logically valid:

TOTALCONV 1* Omnis S est omnis P → Omnis P est omnis S.

We now come to the final part of § DCLXXVII where Caramuel sets out to check "in which way the negative propositions can be [totally] converted".[35] He appears to have been quite distracted and lacking in concentration when he composed (or proof-read) this passage. The pertinent list of all possible conversions of *negative* propositions would have to contain, just like in the preceding case of affirmative propositions, $4 \times 4 = 16$ elements. But he lists only 12 combinations because he forgets the last group where the subject has quantity S and the predicate one from {U, I, P, S}. Furthermore, the examples for the third group (where the subject has quantity P) was erroneously formulated with the copula 'est' instead of 'non est' so that they do not represent negative propositions at all![36] If one corrects this obvious error by inserting 'non' in the corresponding propositions of Caramuel's list, the following principles are obtained:[37]

TOTALCONV 5	"*Nullus homo est omne animal*
	Omne animal est nullus homo"
TOTALCONV 6	"*Nullus homo est aliquis leo*
	Aliquis leo est nullus homo"
TOTALCONV 7	"*Aliquod animal [non] est mnis homo*
	Omnis homo [non] est aliquod animal"
TOTALCONV 8	"*Aliquod animal [non] est aliquis homo*
	Aliquis homo [non] est aliquod animal"

According to Fig. 4, the antecedent of TOTALCONV 6, '*Nullus homo est aliquis leo*' is equivalent to the "colloquial" '*Nullus homo est leo*', and this particular *example* of a UN, of course, is *true*—just as indicated by Caramuel's italics. Furthermore, according to Fig. 6, '*Aliquis leo est nullus homo*' is just an emphatic way to express the PN '*Aliquis leo non est homo*'. Again this particular example, in accordance with Caramuel, is *true*. But now one can generalize this result by maintaining that

TOTALCONV 6* Nullus S est aliquis P → Aliquis P est nullus S

is unrestrictedly *valid*, since it expresses the inference from 'No S is P' to 'Some P isn't S', i.e. the "accidental" conversion of the UN, CONV 5!

Next consider TOTALCONV 8! On the background of Fig. 4, both the (corrected) premise '*Aliquod animal* [*non*] *est aliquis homo*' and the (corrected) conclusion 'Aliquis homo [non] est aliquod animal' are familiar examples of a PN. The former, of course, is true, and the latter false, hence the inference is *invalid*. More generally, the (corrected) principle

TOTALCONV 8* Aliquis S non est aliquis P → Aliquis P non est aliquis S

[35] Cf. *Theol* p. 221: "Videamus igitur, quo [modo negativae] convertantur".

[36] All these problems have been overlooked by Pastore who thinks that Caramuel here only "repeats the first twelve schematic letters, together with corresponding examples, of the preceding §." Cf. [18, p. 131].

[37] As usual, the other inferences involving quality I or S are left out of consideration here.

simply expresses the conversion of the PN, *CONV 6*, which has already been found to be invalid in Sect. 3.[38]

TOTALCONV 5 is considered by Caramuel as *invalid*, since the premise ‘Nullus homo est omne animal’ is printed in italics, i.e. taken to be *true*, while the conclusion ‘Omne animal est nullus homo’ is printed “romano”, i.e. taken to be *false*. In view of Fig. 6, ‘Omne animal est *nullus* homo’ is just an emphatic way to express ‘Omne animal *non est aliquis* homo’, and this UN, or more colloquially, ‘Nullus animal est homo’, really is false. It remains, however, unclear what the unorthodox proposition ‘Nullus homo est omne animal’ precisely means and whether therefore the following general principle really is invalid:

TOTALCONV 5* Nullus S est omnis P → Omnis P est nullus S

Let us finally analyze TOTALCONV 7! According to Fig. 4, the conclusion of the (corrected!) inference, *Omnis homo* [*non*] *est aliquod animal*, means the same as the UN, ‘*Nullus homo est animal*’. This example, of course, is *false*; but this doesn’t contradict the fact that Caramuel had put the conclusion of *his* (uncorrected) inference, *Omnis homo est aliquod animal*, in italics. Surely, *this* proposition is the classical example of a true PA! Now the *premise* of Caramuel’s original inference was the ominous proposition *Aliquod animal est omnis homo* which we encountered already several times before and which is considered by Caramuel as *true* (although it still remains hard to understand why). But even if we take it for granted that *this* proposition *is* true, it doesn’t necessarily follow that our “negated” proposition *Aliquod animal non est omnis homo* must be false. After all, it has been argued in connection with Fig. 5 above that these two unorthodox propositions are opposed to each other as “subcontraries”. So we cannot yet decide whether the general principle

TOTALCONV 7* Aliquis S non est omnis P → Omnis P non est aliquis S

is valid or not. Let us try to solve these open problems by considering some remarks which Caramuel added in a separate section “Notanda”.

6 “Vague Species” vs. “Vague Individuals”

In § DCLXXX, Caramuel admits that one might have different opinions about the truth of principles such as:

> “*Omnis homo est hoc animal* [ergo] *Hoc animal est omnis homo* [. . .]
> *Omnis homo est aliquod animal* [ergo] *Aliquod animal est omnis homo”*.

He tries to defend the first inference by distinguishing two possible interpretations of the singular term ‘hoc animal’. If it is understood as denoting a particular *individual*, then the above propositions are false because “there cannot be found any singular animal with which every man can be identified”.[39] On the other hand, the expression ‘hoc animal’

[38]This counterexample escaped the attention of Caramuel because he forgot the negation ‘non’ and thus considered instead the conversion of the (true) PA ‘*Aliquod animal est aliquis homo*’ into the (equally true) PA ‘*Aliquis homo est aliquis animal*’.

[39]Cf. § DCLXXXII: “Quodsi pronomen *hoc* sit individuale et singulare [. . .] utraque propositio erit falsa: non enim reperitur aliquod animal singulare, cui idemtificetur omnis homo”.

might be interpreted as denoting the *essence* ("ecceitas") of this particular animal. And if this animal happens to be a man, i.e. an "animal rationale", then the essence of 'hoc animal' is the essence of 'animal rationale'. Therefore, both propositions are considered by Caramuel as true since they express that the species 'homo' is the same as the species of 'hoc animal (rationale)'.[40]

This crucial idea of interpreting concepts as "essences" is explained at some more length in § DCLXXXI which deals with the second inference, i.e. TOTALCONV 2.[41] Caramuel points out that the quantifier 'aliquod' can be interpreted in two different ways, either as standing for an undetermined, "vague *individual*", or as standing for an undetermined, "vague *species*". If, e.g. 'Aliquod animal' is interpreted as referring to a "vague individual", then the proposition '*Aliquod animal est omnis homo*' is false.[42] For, clearly, there doesn't exist any particular animal y which is identical with *every* human individual. But if, on the other hand, 'Aliquod animal' is interpreted as referring to a "vague *species*", then '*Aliquod animal est omnis homo*' asserts that a certain species Y is (identical with) the species of man. This proposition now is true, since the species "animal rationale" coincides with the species "omnis homo".

Generalizing from this example, one might be tempted to suppose that Caramuel's entire logic is based on such an understanding of 'Aliquis P' as a "vague species", i.e. an undetermined *subset of P*, and 'Omnis S' as the *set of all S*. This would imply, more precisely, that the UA 'Omnis S est aliquis P' is true iff the whole set S coincides with a certain (non-empty) *subset* Y of P. Similarly, the PA 'Aliquis S est aliquis P' would have to be understood as saying that a certain (non-empty) subset Y of S coincides with a certain (non-empty) subset Z of P. Furthermore, the unorthodox proposition 'Omnis S est omnis P' would mean that the entire set S coincides with the entire set P. Finally 'Aliquis S est omnis P' would mean that a certain (non-empty) subset Y of S coincides with the set of all P. In sum, then, we would obtain the interpretation depicted in Fig. 8.

Omnis S est aliquis P	$\exists Z(Z \subseteq P \wedge Z = S)$, i.e. $S \subseteq P$,
Aliquis S est aliquis P	$\exists Y \exists Z(Z \subseteq P \wedge Y \subseteq S \wedge Y = Z)$, i.e. $S \cap P \neq \varnothing$,
Omnis S est omnis P	$S = P$,
Aliquis S est omnis P	$\exists Y(Y \subseteq S \wedge Y = P)$, i.e. $P \subseteq S$.

Fig. 8 Set-theoretical interpretation of affirmative propositions

This interpretation not only accords with the modern understanding of the UA and PA as $\forall x(Sx \rightarrow Px)$ and $\exists x(Sx \wedge Px)$, respectively, but it also provides an explanation why—for Caramuel—all specific instances of the unorthodox schema 'Omnis S est omnis P'

In the previous §, Caramuel had already pointed out that under the individualistic interpretation even proposition '*Homo* est omnis homo' is *false*: "[...] quia pro nullo suppositorum verificari potest. Neque enim Sortes est omnis homo, neque Plato, neque aliquis alius."

40 Cf. *Theol*, p. 222: "[...] si pronomen illud sumitur pro ecceitate, aut potius taleitate specificâ [...] hoc animal, signato homine ut sic, erit idem ac animal rationale: & hoc sensu utraque propositio erit vera."

41 Cf. § DCLXXXI: "Hanc, *Omnis homo est aliquod animal* aio transire posse in hanc, *Aliquod animal* (nimirum *rationale*, sumendo illud *Aliquod* pro specie vaga) *est omnis homo*".

42 Cf. *Theol*, p. 222: "At Ego dico in hâc propositione, *Aliquod animal est omnis homo* illud *Aliquod* vel sumi pro specie vagâ, vel pro individuo vago: & addo eandem in primâ acceptione esse veram, & in secundâ falsam".

turned out to be *false*. He considered only such examples (e.g. “Omnis homo est omne animal” and “Omne animal est omnis homo”) in which the subject S is *different* from the predicate P. Yet, as was already pointed out in Sect. 5.3, the corresponding *conversion* principle, TOTALCONV 1*, must be considered as logically *valid*.[43]

Nevertheless Fig. 8 offers no adequate basis for a complete reconstruction of Caramuel’s logic because there appears to exist no viable way to interpret *negative propositions* analogously with reference to “vague species”. In particular, Caramuel’s third quantifier ‘Nullus’ cannot consistently be integrated into the above approach! Just as ‘Omnis P’ is taken to denote the *entire* set P, and ‘Aliquis P’ is taken to denote a *subset* of all P, so ‘Nullus P’ would have to be taken to denote the set of *no P* at all, i.e. the *empty set*. But there is not the slightest hint in Caramuel’s work that he might ever have taken this interpretation of ‘Nullus P’ as ‘$\varnothing$’ into account. Furthermore, as we have seen, for Caramuel the proposition ‘Nullus homo est aliquis lapis’ is synonymous with the “colloquial” ‘Nullus homo est lapis’; but, clearly, this UN cannot be analyzed as maintaining that the empty set of no man coincides with a certain subset of stones!

In view of these difficulties, one would have to eliminate the quantifier ‘Nullus’ in favor of ‘Omnis’ and express the UN either with the help of negative concepts as ‘Omnis S est aliquis ∼P’, or by means of the “negative copula” as ‘Omnis S non est aliquis P’. Let us briefly consider the first alternative which would imply that ‘Aliquis ∼P’ is interpreted as a “vague negative species”, i.e. as a subset Z of the *set-theoretical complement* of the class of all P. More precisely, we would obtain the following interpretation of UN and PN:

Omnis S est aliquis ∼P $\quad \exists Z(Z \subseteq \overline{P} \wedge Z = S)$,
Aliquis S est aliquis ∼P $\quad \exists Y \exists Z(Z \subseteq \overline{P} \wedge Y \subseteq S \wedge Y = Z)$.

(Here ‘$\overline{P}$’ symbolizes the set-theoretical complement of P.) These condition, though, would be in accordance with the modern understanding of UN and PN as $\forall x(Sx \rightarrow \neg Px)$ and $\exists x(Sx \wedge \neg Px)$, respectively. But again, there is not the slightest hint in Caramuel’s logical work that he might ever have taken the interpretation of ‘Non-P’ as the set theoretical *complement* of P into account. Furthermore, the corresponding interpretation of the unorthodox propositions:

Omnis S est omnis ∼P $\quad S = \overline{P}$,
Aliquis S est omnis ∼P $\quad \exists Y(Y \subseteq S \wedge Y = \overline{P})$

would give rise to serious problems. As we have seen in connection with principle TOTALCONV 6, Caramuel considers, e.g. ‘Aliquis leo est nullus homo’ as true. But this proposition is now to be equated with ‘Aliquis leo est omnis non-homo’ which would be true iff a certain subset Y of the set of all lions coincides with the complement of the set of all men. But then the analogue example ‘Aliquis *lapis* est nullus homo’ would be true iff a certain subset Z of the set of all stones coincides with the complement of the set of all men. In view of the transitivity of ‘=’, it would follow that Y coincides with Z, so that in sum the truth-condition for the PA ‘Aliquis leo est aliquis lapis’ would be satisfied.[44]

[43] As the reader may verify, also TOTALCONV 2–4 and TRANSSUB 1, 4 become valid under the present set-theoretical interpretation, while TRANSFIG 1, 2 and TRANSSUB 2, 3 are invalid.

[44] Caramuel would even consider the two *universal* propositions ‘Omnis lapis est nullus homo’ and ‘Omnis leo est nullus homo’ as true. Hence, given the above truth-conditions, it would follow that the set of all lions is identical with the set of all stones!

So the only possibility to complete Fig. 8 would consist in paraphrasing the negative propositions with the help of the "copula" 'non est' in the sense of '$\neq$'. According to the basic idea of quantifying over "vague species", the PN 'Aliquis S non est aliquis P' would have to be interpreted as saying that a certain species Y of S is different from a certain species Z of P:

$$\text{Aliquis S non est aliquis P} \quad \exists Y \exists Z(Z \subseteq P \wedge Y \subseteq S \wedge Y \neq Z).$$

Similarly, the UN 'Omnis S non est aliquis P' would have to be interpreted as saying that the whole species S is different from some species Z of P:

$$\text{Omnis S non est aliquis P} \quad \exists Z(Z \subseteq P \wedge Z \neq S).$$

But the weak, almost tautological condition for the PN no longer represents the negation of the UA.[45] Although, e.g. "Every man (S) is an animal" (P), there exists another species of animals, say, the set Z of all lions, which is *different* from S (even in the strong sense of $Z \cap S = \varnothing$)! A fortiori Z is different from *any* species $Y \subseteq S$, so that 'Aliquis homo non est aliquis animal' comes out as *true* under the above formal interpretation! Similarly, the above requirement for the UN, $\exists Z(Z \subseteq P \wedge Z \neq S)$, fails to guarantee that $S \cap P = \varnothing$. Hence the UN is no longer the *contradictory* of the PA. Furthermore, even if the UA is true, i.e. if $S \subseteq P$, it remains possible that the UN is true as well, i.e. that there exists a subset Z of P such that $Z \neq S$. In other words, UA and UN would no longer be *contrary* to each other!

7 The "Real Inventor" of TQP

Let us now return to the question which logician should be considered as the "real inventor" of TQP. Pastore is certainly right when he maintains that Caramuel was *the first* to propound "the lucky [?] idea of assigning to the predicate of each proposition a determined quality" [18, p. 134]. Caramuel's investigations about the possible kinds of a QP may also be considered as *more comprehensive* and *more systematic* than those of Hamilton, De Morgan, and Ploucquet. In particular, Caramuel took account not only of the usual qualities U and P, but also of "indefinite" and singular propositions, and he ended up with an "exhaustive list of all possible cases"[46] of a conversion. Moreover, he outlined at least a rough sketch how to interpret the quantifiers either as ranging over *individuals* or as ranging over *sets*! Nevertheless, there are several reasons to deny that Caramuel *invented* the theory of QP!

Developing a *theory* of QP requires more than only presenting a list of syntactically possible propositions. One has to develop a sufficiently precise *semantics* which allows one to grasp the meaning of these propositions and hence to determine their *truth-conditions*. As was shown above, in this decisive respect Caramuel's logic suffers from

[45] Somewhat surprisingly, this formula is even compatible with the stronger proposition 'Omnis S est omnis P'. If $S = P$ is any set with at least two individuals $\{x, y\}$, then choosing $Z = \{x\}$ and $Y = \{y\}$ shows that $\exists Y \exists Z(Z \subseteq P \wedge Y \subseteq S \wedge Y \neq Z)$.

[46] Cf. [18, p. 133]: "La grande preoccupazione del Caramuel è appunto il calculo e per esso *l'esaustione di tutti i casi possibili*".

serious shortcomings. His "method" of giving a few *examples* which are characterized as true or false by setting them "cursivo" or "romano" doesn't suffice to determine the formal validity of the numerous principles of conversion discussed in Sects. 5.1–5.3. Also, Caramuel's *hint* to the alternative interpretability of quantifiers as either referring to individuals or referring to sets remains too vague to provide a satisfactory semantics of the various propositions with a QP.

Caramuel's idea of interpreting 'Omne P' as the set of all P and 'Aliquis P' as a subset of P was transformed into a real *theory* of QP only about a century later by *Ploucquet*. The logician from the University of Tübingen used 'O.S' and 'Q.P' as symbols for 'Omne S' and 'Quoddam P', respectively, and '–' and '>' as symbols for 'est' and 'non est'.[47] While the "copula" '–' simply means set-theoretical *identity*, the "negative copula" '>' for Ploucquet signifies not just weak diversity in the sense of '$\neq$', but a *strong diversity* in the sense of $S \cap P = \varnothing$! Thus the affirmative propositions can be formalized as shown in Fig. 9.

Omne S est omne P	O.S – O.P	$S = P$,
Omne S est quoddam P	O.S – Q.P	$\exists Z(Z \subseteq P \wedge Z = S)$,
Quoddam S est omne P	Q.S – O.P	$\exists Y(Y \subseteq S \wedge Y = P)$,
Quoddam S est quoddam P	Q.S – Q.P	$\exists Y \exists Z(Z \subseteq P \wedge Y \subseteq S \wedge Y = Z)$.

Fig. 9 Affirmative propositions à la Ploucquet

Hence these propositions receive exactly *the same interpretation as in* Fig. 8. But for the *negative* propositions see Fig. 10.

Omne S non est omne P	O(S) > O(P)	$S \cap P = \varnothing$,
Omne S non est quod. P	O(S) > Q(P)	$\exists Z(Z \subseteq P \wedge S \cap Z = \varnothing)$,
Quod. S non est omne P	Q(S) > O(P)	$\exists Y(Y \subseteq S \wedge Y \cap P = \varnothing)$,
Quod. S non est quod. P	Q(S) > Q(P)	$\exists Y \exists Z(Y \subseteq S \wedge Z \subseteq P \wedge Y \cap Z = \varnothing)$.

Fig. 10 Negative propositions à la Ploucquet

Note that in this schema the first and the third proposition represent the orthodox UN and PN, respectively. In particular, the requirement that a certain subset Y of S must be disjoint from P, $\exists Y(Y \subseteq S \wedge Y \cap P = \varnothing)$, is provably equivalent to the modern condition $\exists x(Sx \wedge \neg Px)$.[48]

To conclude let it be pointed out that about half a century after Caramuel (and about half a century before Ploucquet) *Leibniz* had scrutinized the alternative interpretation of a QP in terms of "vague individuals". In the important fragment "Mathesis rationis", he put forward the following logical analysis of the categorical forms:

> "(3) When I say, Every S is P', I understand that any one of those which are called S is the same as some of those which are called P. This proposition is called a Universal Affirmative.

[47] Cf. [19, p. 3]: "O. praefixum denotat omnitudinem positive sumtam [...] Q. vel q. praefixa denotant particularitatem [...] S – P denotat: S est P. S > P denotat: S non est P". In some places he also uses 'N.P' for the negative quantifier 'Nullus P'.

[48] Note also that Ploucquet's system fully accords with the traditional doctrine that the predicate of an affirmative proposition is always particular while the predicate of a negative proposition is always universal.

(4) When I say 'Some S is P', I understand that some one of those which are called S is the same as some of those which are called P; this is a Particular Affirmative proposition.
(5) When I say 'No S is P', I understand that any one of those which are called S is different from any one of those which are called P; this is a Universal Negative proposition.
(6) Finally, when I say 'Some S is not P', I understand that some one of those which are called S is different from any one of those which are called P; this is called a Particular Negative."[49]

These conditions may straightforwardly be formalized as follows:

$$\begin{array}{ll} \text{UA} & \forall x\big(Sx \rightarrow \exists y(Px \wedge y = x)\big), \\ \text{PA} & \exists x\big(Sx \wedge \exists y(Px \wedge y = x)\big), \\ \text{UN} & \forall x\big(Sx \rightarrow \forall y(Px \rightarrow y \neq x)\big), \\ \text{PN} & \exists x\big(Sx \wedge \forall y(Px \rightarrow y \neq x)\big). \end{array}$$

Immediately after the analysis of the four categorical forms, Leibniz considers four *unorthodox* propositions:

(7) "It could also be considered that Every S is every P, i.e. that all those which are called S are the same as all those which are called P [...], but this is not used in our languages. In the same way, we do not say that Some S are the same as all P, for we express that when we say that All P are S. But it would be superfluous to say that No S is some P, i.e. that any of those which are called S is different from some one of those which are called P; for this is self-evident, unless P is unique. Much more would it be superfluous to say that someone of those which are called S is different from some one of those which are called P."

Although these propositions normally do not "occur in our languages", they have perfectly clear truth-conditions. Thus, 'Every S is every P' means that every individual having property S is identical with every individual having property P. Hence Leibniz's interpretation of the universal affirmative with explicitly quantified predicate, UA_{QP}, for short, can be formalized as follows:

$$UA_{QP} \quad \forall x\big(Sx \rightarrow \forall y(Py \rightarrow y = x)\big).$$

This very strong condition will never be satisfied unless the sets S and P are singletons which contain exactly one and the same object. Similarly, the corresponding proposition 'Some S is every P' means that there exists an object x having property S which is identical with every object y having property P:[50]

$$PA_{QP} \quad \exists x\big(Sx \wedge \forall y(Py \rightarrow y = x)\big).$$

Again this very strong condition can't be true unless the set P is a singleton. The next unorthodox proposition says that each object having property S is different from at least one object having property P:

$$UN_{QP} \quad \forall x\big(Sx \rightarrow \exists y(Py \wedge y \neq x)\big).$$

Finally, we have the proposition that at least one object having property S is different from at least one object having property P:

$$PN_{QP} \quad \exists x\big(Sx \wedge \exists y(Py \wedge y \neq x)\big).$$

[49] [17, p. 95]; cf. also the original version in [6, p. 193].

[50] Note, incidentally, that Leibniz commits a fallacy when he maintains that PA_{QP} might be expressed by "Omne P esse quoddam S". The latter proposition is a UA and thus it has to be formalized as $\forall x(Px \rightarrow \exists y(Sy \wedge y = x))$. As modern logic shows, however, one may not simply *interchange* the two quantifiers within such a formula!

The two formulas UN_{QP} and PN_{QP} now express very weak, almost tautological statements the truth of which is self-evident unless, again, P is a singleton (or, as Leibniz himself put it: "nisi P sit unicum"). It is an easy exercise in first order logic to check that PN_{QP} represents the negation of UA_{QP}, and similarly, PA_{QP} is the negation of UN_{QP}. Hence the four unorthodox formulas constitute another "square of opposition".

References

1. Arnauld, A., Nicole, P.: La Logique ou L'Art de Penser, 5th edn. (1683); reprint Paris (Presses universitaires de France), 1965
2. Bentham, G.: Outlines of a New System of Logic [. . .]. Hunt & Clarke, London (1827)
3. Bocheński, J.M.: Formale Logik. Alber, Freiburg (1956)
4. Bök, A.F. (ed.): Sammlung der Schriften, welche den logischen Calcul Herrn Prof. Ploucquets betreffen, mit neuen Zusätzen, Frankfurt, Leipzig (1766); reprint Stuttgart, 1970
5. Caramuel, J.: Theologia Rationalis. Schönwetter, Frankfurt (1654)
6. Couturat, L.: Opuscules et fragments inédits de Leibniz. Paris (1903)
7. De Morgan, A.: Formal Logic, or, the Calculus of Inference, Necessary and Probable. Taylor & Walton, London (1847)
8. Dvořák, P.: Juan Caramuel Lobkowitz: Selected Aspects of Formal and Applied Logic. Oikumene, Prague (2006) (in Czech)
9. Dvořák, P.: Relational logic of Juan Caramuel. In: Gabbay, D.M., Woods, J. (eds.) Medieval and Renaissance Logic. The Handbook of the History of Logic, vol. 2, pp. 645–665. Elsevier, Amsterdam (2008)
10. Franz, M. (ed.): Gottfried Ploucquet—Logik. Hildesheim, Zürich, New York (2006)
11. Hamilton, W.: Lectures on Metaphysics and Logic. Edinburgh, London (1861), ed. by H.L. Mansel and J. Veitch; reprint Stuttgart, Bad Cannstadt, 1969
12. Kneale, W., Kneale, M.: In: The Development of Logic. Clarendon, Oxford (1962)
13. Lenzen, W.: Der "logische Calcul Herrn Prof. Ploucquets" (1). Arch. Gesch. Philos. **90**, 74–114 (2008)
14. Lenzen, W.: Ploucquet's 'refutation' of the traditional square of opposition. Log. Univers. **2**, 43–58 (2008)
15. Lenzen, W., von Wulfen, H.: Der "logische Calcul Herrn Prof. Ploucquets" (2). In: Lenski, W., Neuser, W. (eds.) Logik als Grundlage von Wissenschaft, pp. 17–51. Winter, Heidelberg (2010)
16. Lenzen, W.: The quantification of the predicate—Leibniz, Ploucquet and the (double) square of opposition. In: Nicolás, J.A. (ed.) Leibniz und die Entstehung der Modernität, pp. 179–191. Steiner, Stuttgart (2010)
17. Parkinson, G.H.R.: Leibniz—Logical Papers. Oxford University Press, Oxford (1966)
18. Pastore, V.A.: Giovanni Caramuel di Lobkowitz e i primordi della teoria della quantificazione del predicato. Class. Neo-Lat. **1**, 125–135 (1905)
19. Ploucquet, G.: Fundamenta Philosophiae Speculativae. Tübingen (1759); quoted according to the "Extracta e Fundamentis Philosophiae Speculativae" in Bök (1766), 3–14
20. Ploucquet, G.: Institutiones Philosophiae Theoreticae sive De arte cogitandi [. . .]. Tübingen (1772)
21. Ploucquet, G.: Expositiones Philosophiae Theoreticae. Stuttgart (1782)
22. Risse, W.: Die Logik der Neuzeit, vol. 1. Frommann, Stuttgart (1964)
23. Solly, T.: A Syllabus of Logic [. . .]. London (1839)
24. Thom, P.: The Syllogism. Philosophia, München (1981)
25. Velarde Lombraña, J.: Aportaciones de Caramuel a la logica. In: Actas: Comunicaciones Presentadas al II^{o} Congreso de Teoria y Metodologia de las Ciencias, Oviedo, pp. 273–277 (1984)
26. Venn, J.: Symbolic Logic. Macmillan, London (1881)

W. Lenzen (✉)
University of Osnabrück, Osnabrück, Germany
e-mail: lenzen@uos.de

Lossy Inference Rules and Their Bounds: A Brief Review

David Makinson and James Hawthorne

Abstract This paper reviews results that have been obtained about bounds on the loss of probability occasioned by applying classically sound, but probabilistically unsound, Horn rules for inference relations. It uses only elementary finite probability theory without appealing to linear algebra, and also provides some new results, in the same spirit, on non-Horn rules. More specifically, it does the following: (i) recalls Adams' well-known sum bound for the rule Right∧+ and shows how it is inherited by the rules CM, CT and Left∨+; (ii) draws attention to lesser known but tighter bounds for CM, CT and Left∨+ due respectively to Bourne & Parsons, Adams, and Gilio, and provides elementary verifications for those that were originally obtained using linear algebra; (iii) shows that the sum bound for Right∧+ and the improved bounds for CM, CT and Left∨+ are all in a natural sense optimal; (iv) distinguishes two kinds of loss for almost-Horn rules, 'distributed' and 'pointed'; and (v) finds bounds for both distributed and pointed loss, optimal in the distributed case, for the specific almost-Horn rules of disjunctive rationality (DR) and rational monotony (RM).

Keywords Uncertain inference · Probabilistic inference · Lossy rules · Preferential inference · Nonmonotonic reasoning · Horn rules · Almost-Horn rules

Mathematics Subject Classification Primary 03B48 · Secondary 03B05 · 03B60

1 Introduction

It is a great pleasure to participate in the *Festschrift* for Jean-Yves Béziau. This paper is in the spirit of what he dubbed 'universal logic': it concerns the probabilistic assessment of rules for uncertain inference, and brings out close connections between quantitative and purely qualitative perspectives.

As is well known, one may model relations $|\!\sim$ of uncertain inference probabilistically, interpreting $a \mid\!\sim x$ as saying that $p(x|a) \geq t$ modulo a given probability function p and threshold value t taken as parameters. It is also well known that a single application of the inference rule of monotony (aka strengthening the antecedent) for an inference relation $|\!\sim$

$$\text{Whenever } a \mid\!\sim x, \quad \text{then } a \wedge b \mid\!\sim x$$

can occasion a total loss of probability. No matter how close the conditional probability $p(x|a)$ (associated with the premise $a \mid\!\sim x$) may be to unity, so long as it doesn't equal

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_18

unity the conditional probability $p(x|a \wedge b)$ (associated with the conclusion $a \wedge b \mid\sim x$) can be zero.

In contrast, some other rules do not lead to any loss of probability, an example being the rule of very cautious monotony (VCM):

$$\text{Whenever } a \mid\sim b \wedge x, \quad \text{then } a \wedge b \mid\sim x.$$

The probability $p(x|a \wedge b)$ associated with its conclusion is always at least as high as the probability $p(b \wedge x|a)$ associated with its premise.

There are also some rules 'in between', whose application may lose some, but not all the probability of the premises; we may describe them as *lossy*. An important example is the rule of conjunction of consequents, also known as Right$\wedge$+:

$$\text{Whenever } a \mid\sim x \text{ and } a \mid\sim y, \quad \text{then } a \mid\sim x \wedge y.$$

So long as the conditional probabilities associated with the premises $a \mid\sim x$ and $a \mid\sim y$ are 'sufficiently high', then that of the conclusion $a \mid\sim x \wedge y$ must be 'reasonably high', with an identifiable bound on how low it can fall.

This paper is about rules of the intermediate kind and their bounds. As well as Right$\wedge$+, we will consider cautious monotony (CM), cumulative transitivity (CT) and unrestricted Left$\vee$+, whose definitions are recalled below. Some of their bounds have been established using linear algebra; we verify them all using only elementary finite probability theory, with some alternative proofs using material from non-monotonic logic. We also begin an investigation of bounds for non-Horn rules, focussing on those of disjunctive rationality (DR) and rational monotony (RM).

After Sect. 2 on background concepts, all sections end with a short summary and all bounds are displayed in boxes. Readers may wish to consult them before deciding whether to plunge into details.

2 Background Concepts

We recall the notion of probabilistic soundness for Horn rules, the formulations of the specific Horn rules mentioned above, and the useful concept of sum soundness.

2.1 Probabilistic Soundness

Consider any n-premise Horn rule for inference relations, authorizing passage from a set of premises $a_i \mid\sim x_i$ $(i \leq n)$ to conclusion $b \mid\sim y$, where the formulae are from classical propositional logic. Note that the premises and the conclusion of such a rule are not themselves formulae of classical logic, but are expressions $a \mid\sim x$ saying that the inference relation $\mid\sim$ holds between classical formulae a and x.

The rule is said to be *probabilistically sound* iff $p(y|b) \geq \min\{p(x_i|a_i) : i \leq n\}$ for every probability function p. In other words iff, for every such function p and every $t \in [0, 1]$, if $p(x_i|a_i) \geq t$ for all $i \leq n$ then also $p(y|b) \geq t$. Here, probability functions are understood as satisfying the usual (finitary) Kolmogorov postulates, and conditional probability $p(x|a)$ is defined in the standard way, putting $p(x|a) = p(a \wedge x)/p(a)$ when $p(a) \neq 0$. It will be convenient to let $p(x|a) = 1$ when $p(a) = 0$, rather than leave it undefined. In the limiting case that $n = 0$, the definition of probabilistic soundness is understood as requiring that $p(y|b) = 1$ for every probability function p.

This definition may be expressed more concisely if we write $a \mid\!\sim_{p,t} x$ for $p(x|a) \geq t$: a rule from premises $a_i \mid\!\sim x_i (i \leq n)$ to conclusion $b \mid\!\sim y$ is probabilistically sound iff $b \mid\!\sim_{p,t} y$ whenever $a_i \mid\!\sim_{p,t} x_i$ for all $i \leq n$. It may also be expressed in terms of improbability: the requirement becomes $1 - p(y|b) \leq \max\{1 - p(x_i|a_i) : i \leq n\}$, that is, $\text{imp}(y|b) \leq \max\{\text{imp}(x_i|a_i) : i \leq n\}$ where we write $\text{imp}(\cdot|\cdot)$ for $1 - p(\cdot|\cdot)$ as is convenient for expressions with a single probability function p.

We have already mentioned one example of a probabilistically sound Horn rule, namely very cautious monotony (VCM). Other important examples, recalled in detail in Sect. 4.1 below, include reflexivity, left classical equivalence, right weakening, exclusive Left$\vee$+, and weak Right$\wedge$+. Taken together with very cautious monotony they constitute the family O, which is nearly, but not quite, complete for probabilistically sound Horn rules.

It was shown in Hawthorne and Makinson [10, Observation 2.4] that probabilistically sound Horn rules are always preferentially sound, that is, they hold in the well-known system **P** of preferential inference. In turn, preferentially sound Horn rules are always classically sound (see, e.g. Kraus, Lehmann & Magidor [11] or the textbook presentation Makinson [13]). But both converses fail. A notorious example of a classically sound Horn rule that is not preferentially sound is monotony (defined above). Salient examples of preferentially sound Horn rules that are not probabilistically sound are conjunction of conclusions (Right$\wedge$+, from $a \mid\!\sim x$, $a \mid\!\sim y$ to $a \mid\!\sim x \wedge y$), cautious monotony (CM, from $a \mid\!\sim x$, $a \mid\!\sim y$ to $a \wedge x \mid\!\sim y$), cumulative transitivity (CT, from $a \mid\!\sim x$, $a \wedge x \mid\!\sim y$ to $a \mid\!\sim y$) and unrestricted disjunction in the premises (Left$\vee$+, from $a \mid\!\sim x$, $b \mid\!\sim x$ to $a \vee b \mid\!\sim x$).

For all of them, and others, the question thus arises: how much probability can be lost in their application? In other words, how *lossy* are they?

2.2 Sum Soundness

In answering this question, the notion of sum soundness has for long played an important role. A Horn rule $a_i \mid\!\sim x_i (i \leq n)/b \mid\!\sim y$ is said to be *sum sound* iff the improbability of y given b is always less than or equal to the sum of the improbabilities of the x_i given their respective a_i. Briefly, iff $\text{imp}(y|b) \leq \sum\{\text{imp}(x_i|a_i) : i \leq n\}$ for every probability function p. Sum soundness thus contrasts neatly with probabilistic soundness: *sum* simply replaces max in the improbability formulation of the latter given above. The notion of sum soundness may also be expressed rather less transparently in terms of probability rather than improbability, holding iff $p(y|b) \geq \sum\{p(x_i|a_i) : i \leq n\} - (n - 1)$.

For $n = 2$, which is the case that will principally concern us, a Horn rule is sum sound iff $\text{imp}(y|b) \leq \text{imp}(x_1|a_1) + \text{imp}(x_2|a_2)$; and when $n = 1$, iff $\text{imp}(y|b) \leq \text{imp}(x_1|a_1)$. In the limiting case that $n = 0$, the definition is understood as requiring $\text{imp}(y|b) = 0$. In terms of probability, the requirements for sum soundness are: for $n = 2$, $p(y|b) \geq p(x_1|a_1) + p(x_2|a_2) - 1$, for $n = 1$, $p(y|b) \geq p(x_1|a_1)$; for $n = 0$, $p(y|b) = 1$.

For zero and one-premise Horn rules, the two concepts of sum soundness and probabilistic soundness thus coincide, and are known also to coincide with preferential soundness as shown, e.g. in Hawthorne and Makinson [10, Observation 4.3]. When more than one premise is in play, the two notions diverge: probabilistic soundness immediately implies sum soundness since $\max\{\text{imp}(x_i|a_i) : i \leq n\} \leq \sum\{\text{imp}(x_i|a_i) : i \leq n\}$; but not in general conversely.

Historically, the concept of sum soundness for Horn rules of inference seems first to have been articulated by Adams [1]. It is a very useful yardstick when comparing the loss of probability that may be incurred when applying inference rules—and of course, the less loss, the better. But the choice of the word 'sound' does not mean that the property is a certificate of acceptability: sum soundness does not play the same kind of normative role as does probabilistic soundness, in other words, zero loss. For a general discussion of the notion of a lossy rule of inference and its use in analyzing the celebrated lottery and preface paradoxes, see Makinson [14].

3 Conjunction of Conclusions (Right∧+)

We show that the rule of conjunction of conclusions is sum sound and that, in a natural sense, this bound is optimal for it.

3.1 Sum Soundness of Right∧+

Like monotony, the rule of conjoining conclusions (from $a \mid\sim x$, $a \mid\sim y$ to $a \mid\sim x \wedge y$) is not probabilistically sound. But it is not difficult to show, as was already done by Adams [1], that it is nevertheless sum sound. That is, $\text{imp}(x \wedge y|a) \leq \text{imp}(x|a) + \text{imp}(y|a)$, equivalently $p(x \wedge y|a) \geq p(x|a) + p(y|a) - 1$.

This follows immediately from a stronger full equality which one might call *exact sum soundness* for Right∧+: $\text{imp}(x \wedge y|a) = \text{imp}(x|a) + \text{imp}(y|a) - \text{imp}(x \vee y|a)$, and which may be established by a short chain of equalities using de Morgan and the well-known probabilistic version of the 'principle of inclusion and exclusion': $\text{imp}(x \wedge y|a) = p(\neg(x \wedge y)|a) = p(\neg x \vee \neg y|a) = p(\neg x|a) + p(\neg y|a) - p(\neg x \wedge \neg y|a) = \text{imp}(x|a) + \text{imp}(y|a) - \text{imp}(x \vee y|a)$.

When expressed 'positively' in terms of probabilities, exact sum soundness says: $p(x \wedge y|a) = p(x|a) + p(y|a) - p(x \vee y|a)$, which is nicely (and curiously) symmetric with the improbability formulation. The positive version may also be established directly: by the 'principle of inclusion and exclusion' $p(x \vee y|a) = p(x|a) + p(y|a) - p(x \wedge y|a)$, and swapping terms gives $p(x \wedge y|a) = p(x|a) + p(y|a) - p(x \vee y|a)$ as desired. For easy reference, both formulations are displayed in the accompanying table 'Sum bound for Right∧+'.

Although exact sum soundness is more precise than plain sum soundness—indeed, being an equality, it is as precise as possible—we would not ordinarily call it a 'bound' on the rule Right∧+. For, to calculate $p(x|a) + p(y|a) - p(x \vee y|a)$ one needs to know, in general, not only the values of $p(x|a)$ and $p(y|a)$ corresponding to the two premises of the rule Right∧+, but also that of $p(x \vee y|a)$. If one is to use the language of bounds, one could perhaps say that exact sum soundness gives a soft bound for the probability of the conclusion of Right∧+, contrasting it with the (hard) bound provided by plain sum soundness.

In general, a *bound* (or hard bound for emphasis) for an n-premise Horn rule passing from premises $a_i \mid\sim x_i$ $(i \leq n)$ to conclusion $b \mid\sim y$ is a function $f : [0, 1]^n$ into $[0, 1]$, expressed in purely arithmetic terms, such that $p(y|b) \geq f(p(x_i|a_i) : i \leq n)$ so long as $0 \neq p(x_i|a_i) \neq 1$ for all $i \leq n$. The reason for the proviso is that it is convenient to allow exceptions for extreme values $0, 1$ of the $p(x_i|a_i)$, as we will do in Sects. 4 and 6 for certain bounds on the rules CM and Left∨+. A soft bound could admit more arguments to the function f. We consider only hard bounds, and call them simply bounds.

In the table on the sum bound for Right∧+, the top row gives the bound in terms of improbability and the second row in terms of probability. The second row also contains the general form of that bound, where r, s are understood to be the probabilities of the premises of Right∧+ and t is the probability of its conclusion. The same pattern will be followed in tables for other rules.

Sum bound for Right∧+
$\text{imp}(x \wedge y\|a) \leq \text{imp}(x\|a) + \text{imp}(y\|a)$
$p(x \wedge y\|a) \geq p(x\|a) + p(y\|a) - 1$ $t \geq r + s - 1$

3.2 Optimality of the Sum Bound for Right∧+

Bounds may be looser or tighter, and we will say that a (hard) bound f for an n-premise Horn rule $a_i \mid\sim x_i (i \leq n)/b \mid\sim y$ is *optimal* (for that rule) iff for all $r_1, \ldots, r_n \in (0, 1)$ there is a probability function p and elements, a_i, x_i, b, y of its domain with each $p(x_i|a_i) = r_i$ and $p(y|b) = \max\{0, f(r_1, \ldots, r_n)\}$. Again, as indicated, we allow exceptions for extreme values 0, 1 of the $p(x_i|a_i)$.

In this sense, the bound $f(r, s) = r + s - 1$ for the rule Right∧+ is optimal: for all $r, s \in (0, 1)$ there is a probability function p and elements, a, x, y of its domain with $p(x|a) = r$, $p(y|a) = s$, and $p(x \wedge y|a) = \max\{0, r + s - 1\}$. The idea of the verification is to choose p, etc. in such a way as to force $p(x \wedge y|a)$ to be 0 or $r + s - 1$ according as $r + s \leq 1$ or $r + s > 1$, and then apply exact sum soundness.

Specifically, take the domain of p to be 2^S (the set of all subsets of S) where $S = \{1, 2, 3\}$, and consider two cases. *Case* 1. Suppose $r + s \leq 1$. Put $p(\{1\}) = r$, $p(\{3\}) = s$, $p(\{2\}) = 1 - (r + s)$. This sums to one and so extends to a probability function on 2^S. Put $a = \{1, 2, 3\}$, $x = \{1\}$, $y = \{3\}$. Then $p(x \wedge y|a) = p(a \wedge x \wedge y)/p(a) = 0 = \max\{0, r + s - 1\}$ as desired. *Case* 2. Suppose $r + s > 1$. Put $p(\{1\}) = 1 - s$, $p(\{3\}) = 1 - r$, $p(\{2\}) = 1 - [(1 - s) + (1 - r)] = (r + s) - 1$. This also sums to one and so extends to a probability function on 2^S. Put $a = \{1, 2, 3\}$, $x = \{1, 2\}$, $y = \{2, 3\}$. Then $p(x \wedge y|a) = p(a \wedge x \wedge y)/p(a) = p(\{2\}) = \max\{0, r + s - 1\}$ and the verification is complete.

In presenting this verification we have, for convenience, mixed the propositional and field-of-subsets ways of speaking of the domain of a probability function; but it should be clear how one could pedantically write out the proof in one or the other mode alone.

3.3 Summary on Lossiness for **Right∧+**

Thus, the loss of probability occasioned by an application of Right∧+ is known to be limited: the improbability in the conclusion is never more than the sum of the improbabilities in the two premises, and the bound is, in a natural sense, optimal. The former feature may be seen as providing a justification for monitored application of the rule in probabilistic contexts, preferably beginning from premises whose probability is well above the threshold that is taken as the norm for the context in which one is working.

In other words, the rule Right∧+ is 'lossy', rather like procedures for compressing the data for a digital image. The amount of loss is variable, but we have a relatively small maximal loss on each application, and can tolerate it if the initial images/premises are of high quality and we do not apply the compression/inference too often.

4 Inheriting the Sum Bound from Right∧+

The bound for ∧+ prompts a search for bounds on other Horn rules that are preferentially but not probabilistically sound. In particular, cautious monotony CM, cumulative transitivity CT, and disjunction in the premises Left∨+ (all formulated in Sect. 2.1) are in that situation. How lossy are they? The first point to observe is that they too are sum sound. This was shown by Bourne & Parsons [5, 6], mainly using linear algebra; in this section, we give elementary verifications.

4.1 Two Methods of Verification

In fact, there are two different kinds of elementary verification for the sum soundness of CM, CT, and Left∨+. One is by direct arithmetic calculation. The other is by taking a system of qualitative rules for probabilistic inference relations, and analyzing a suitable derivation of the rule in the system augmented by Right∧+. We display both methods.

For some of the verifications by arithmetic calculation, we will be making use of the following 'product trumps addition' fact: for $r, s \in [0, 1]$, $r \cdot s \geq r + s - 1$. Indeed, more strongly, if $r, s < 1$ then $r \cdot s > r + s - 1$ while in the limiting case that $r = 1$ or $s = 1$, $r \cdot s = r + s - 1$. The limiting case is immediate. To check the principal case, suppose $r \cdot s \leq r + s - 1$. Then $1 - s \leq r - r \cdot s = r \cdot (1 - s)$ which implies $r = 1$ or $s = 1$.

For the verifications by analysis of derivations, we refer to Hawthorne's system **O**, all of whose rules are probabilistically sound. We recall from [8, 10] that it is made up of a zero-premise rule of reflexivity, one-premise rules of left logical equivalence and right weakening and a pair of rather special two-place rules, exclusive Left∨+ and weak Right∧+, as follows:

$a \mid\sim a$	(reflexivity),
Whenever $a \mid\sim x$ and $a \dashv\vDash b$, then $b \mid\sim x$	(LCE, left classical equivalence),
Whenever $a \mid\sim x$ and $x \vDash y$, then $a \mid\sim y$	(RW, right weakening),
Whenever $a \mid\sim x \wedge y$, then $a \wedge x \mid\sim y$	(VCM, very cautious monotony),
Whenever $a \mid\sim x, b \mid\sim x$ and $a \vDash \neg b$, then $a \vee b \mid\sim x$	(XOR, exclusive Left$\vee$+),
Whenever $a \mid\sim x$ and $a \wedge \neg y \mid\sim y$, then $a \mid\sim x \wedge y$	(WAND, weak Right$\wedge$+).

Here $\vDash$ (in RW and XOR) is classical logical consequence and $\dashv\vDash$ (in LCE) is classical equivalence. They are auxiliary relations: notwithstanding their presence, LCE and RW are regarded as one-premise rules, since only one premise contains $\mid\sim$. Weak Right$\wedge$+ (WAND) is strictly weaker than plain Right$\wedge$+ (whenever $a \mid\sim x$ and $a \mid\sim y$, then $a \mid\sim x \wedge y$). Exclusive Left$\vee$+ (XOR) differs from unrestricted Left$\vee$+ (whenever $a \mid\sim x$ and $b \mid\sim x$, then $a \vee b \mid\sim x$) in that it is subject to the condition that a is classically inconsistent with b, and in the context of system **O** it is strictly weaker than the unrestricted version. Given the other rules of **O**, XOR may equivalently be expressed as: whenever $a \wedge b \mid\sim x$, $a \wedge \neg b \mid\sim x$, then $a \vee b \mid\sim x$. In that form it is also known as weak Left$\vee$+ or weak OR (WOR); for more see [8, 10].

Replacing the weakened rules exclusive Left$\vee$+ and weak Right$\wedge$+ by their plain counterparts gives us the system **P** of preferential consequence. Indeed, just replacing weak Right$\wedge$+ by plain Right$\wedge$+ suffices for that. Thus we know that CM, CT, Left$\vee$+, being preferentially valid, are derivable from **O** supplemented by Right$\wedge$+. That alone does not imply that any lower bound for Right$\wedge$+ also serves as such for those rules, for the derivation could apply Right$\wedge$+ several times. But, as we will see, inspection of standard derivations in the literature reveals that in each of them $\wedge$+ is applied only once, so that the bound is indeed inherited.

4.2 Cautious Monotony (CM)

The sum bound for CM may be obtained from that for Right$\wedge$+ by elementary arithmetic as follows. We want to show that $p(y|a \wedge x) \geq p(x|a) + p(y|a) - 1$. By the bound for Right$\wedge$+, we know that $p(x \wedge y|a) \geq p(x|a) + p(y|a) - 1$, so we need only show that $p(y|a \wedge x) \geq p(x \wedge y|a)$. In the limiting case that $p(a \wedge x) = 0$, $p(y|a \wedge x) = 1$ and we are done. In the case that $p(a \wedge x) \neq 0$, we also have $p(a) \neq 0$ so that $p(y|a \wedge x) = p(a \wedge x \wedge y)/p(a \wedge x)$ while $p(x \wedge y|a) = p(a \wedge x \wedge y)/p(a)$, and we conclude by noting that $p(a) \geq p(a \wedge x)$.

Sum bound for CM
$\text{imp}(y\|a \wedge x) \leq \text{imp}(x\|a) + \text{imp}(y\|a)$
$p(y\|a \wedge x) \geq p(x\|a) + p(y\|a) - 1$ $t \geq r + s - 1$

For a verification of the same bound for CM using logical means, we need a suitable notation. For any $\varepsilon \in [0, 1]$, write $a \mid\sim_{\varepsilon} x$ as shorthand for $\text{imp}(x|a) \leq \varepsilon$, that is, $p(x|a) \geq$

$1 - \varepsilon$, and note that since the rules of **O** are probabilistically sound we may apply them taking the index of the conclusion to be the maximum of the indices of the premises. In particular, this means that for the zero-premise rule of reflexivity the index is 0 and for the one-premise rules the index stays unchanged. For applications of Right$\wedge$+ the index goes up from $\varepsilon_1, \varepsilon_2$ to $\varepsilon_1 + \varepsilon_2$, as already established.

With these indices we decorate a very short derivation of CM in **O**-plus-Right$\wedge$+. Given $a \mid\!\sim_{\varepsilon 1} x$ and $a \mid\!\sim_{\varepsilon 2} y$ we need to get $a \wedge x \mid\!\sim_{\varepsilon 1+\varepsilon 2} y$. From the two assumptions we have by Right$\wedge$+ that $a \mid\!\sim_{\varepsilon 1+\varepsilon 2} x \wedge y$ and so $a \wedge x \mid\!\sim_{\varepsilon 1+\varepsilon 2} y$ by VCM and we are done.

4.3 Cumulative Transitivity (CT)

Recall that CT authorizes passage from $a \mid\!\sim x$ and $a \wedge x \mid\!\sim y$ to $a \mid\!\sim y$; it is thus a partial converse of CM. We can give an arithmetic verification of its sum soundness without even using that for Right$\wedge$+. For, as already observed by Adams [3, Sect. 6.6, p. 128] we have $p(y|a) \geq p(x \wedge y|a) = p(x|a) \cdot p(y|a \wedge x)$, and by substitution in the 'product trumps sum' principle mentioned above, $p(x|a) \cdot p(y|a \wedge x) \geq p(x|a) + p(y|a \wedge x) - 1$. Putting these together, $p(y|a) \geq p(x|a) + p(y|a \wedge x) - 1$ as desired. Later we will see how this arithmetic argument may be refined to give an improved bound for CT.

Sum bound for CT
$\text{imp}(y\|a) \leq \text{imp}(x\|a) + \text{imp}(y\|a \wedge x)$
$p(y\|a) \geq p(x\|a) + p(y\|a \wedge x) - 1$ $t \geq r + s - 1$

For a 'logical' verification of the bound, we analyze a standard derivation of CT in the system **O**-plus-Right$\wedge$+. Suppose $a \mid\!\sim_{\varepsilon 1} x$ and $a \wedge x \mid\!\sim_{\varepsilon 2} y$; we need to get $a \mid\!\sim_{\varepsilon 1+\varepsilon 2} y$. By reflexivity and right weakening, $a \wedge \neg x \mid\!\sim_0 \neg x \vee y$. On the other hand, $a \wedge x \mid\!\sim_{\varepsilon 2} y$ gives $a \wedge x \mid\!\sim_{\varepsilon 2} \neg x \vee y$ by right weakening. Combining these by XOR and LCE we have $a \mid\!\sim_{\varepsilon 2} \neg x \vee y$. Applying Right$\wedge$+ to that and $a \mid\!\sim_{\varepsilon 1} x$ gives $a \mid\!\sim_{\varepsilon 1+\varepsilon 2} x \wedge (\neg x \vee y)$, and so by right weakening $a \mid\!\sim_{\varepsilon 1+\varepsilon 2} y$ as desired.

Interestingly, while the arithmetic verification did not appeal to the sum-soundness of Right$\wedge$+, this logical one does apply that rule, and the application cannot be dispensed with since CT is not probabilistically sound and thus not derivable in the system **O** alone.

4.4 Disjunction in the Premises (**Left$\vee$+**)

The sum soundness of Left$\vee$+ was obtained by Bourne & Parsons [6], using linear programming. It can also be given a direct elementary verification, as we now show. As in the case of CT, we do not need to use the sum soundness of Right$\wedge$+.

We first note that $p(x|a) \leq p(x \vee \neg a|a \vee b)$ and $p(x|b) \leq p(x \vee a|a \vee b)$. For the former, it suffices to show $1 - p(x \vee \neg a|a \vee b) \leq 1 - p(x|a)$ as follows:

$$1 - p(x \vee \neg a|a \vee b) = p(\neg x \wedge a|a \vee b) = p(\neg x|a) \cdot p(a|a \vee b) \leq 1 - p(x|a).$$

For the latter, it likewise suffices to show $1 - p(x \vee a|a \vee b) \leq 1 - p(x|b)$ by the following chain:

$$\begin{aligned} 1 - p(x \vee a|a \vee b) &= p(\neg x \wedge \neg a|a \vee b) = p(\neg x \wedge \neg a \wedge b|a \vee b) \leq p(\neg x \wedge b|a \vee b) \\ &= p(\neg x|b) \cdot p(b|a \vee b) \leq 1 - p(x|b). \end{aligned}$$

Thus $p(x|a) + p(x|b) \leq p(x \vee \neg a|a \vee b) + p(x \vee a|a \vee b)$. But the right side of this inequality equals $1 + p(x|a \vee b)$, as shown by the chain:

$$\begin{aligned} 1 &= p\big((x \vee \neg a) \vee (x \vee a)|a \vee b\big) \\ &= p(x \vee \neg a|a \vee b) + p(x \vee a|a \vee b) - p\big((x \vee \neg a) \wedge (x \vee a)|a \vee b\big) \\ &= p(x \vee \neg a|a \vee b) + p(x \vee a|a \vee b) - p(x|a \vee b). \end{aligned}$$

Putting these together, $p(x|a) + p(x|b) \leq 1 + p(x|a \vee b)$, that is $p(x|a \vee b) \geq p(x|a) + p(x|b) - 1$, as desired.

Sum bound for Left∨+
$\text{imp}(x\|a \vee b) \leq \text{imp}(x\|a) + \text{imp}(x\|b)$
$p(x\|a \vee b) \geq p(x\|a) + p(x\|b) - 1$ $t \geq r + s - 1$

For a 'logical' verification of the same bound, we analyze a standard derivation of Left∨+ in the system **O**-plus-Right∧+ (with Right∧+ applied once). Suppose $a \mid\!\sim_{\varepsilon 1} x$ and $b \mid\!\sim_{\varepsilon 2} x$. We need to show that $a \vee b \mid\!\sim_{\varepsilon 1+\varepsilon 2} x$. Applying LCE to the first supposition, $(a \vee b) \wedge a \mid\!\sim_{\varepsilon 1} x$, so $(a \vee b) \wedge a \mid\!\sim_{\varepsilon 1} x \vee \neg a$ (RW); also $(a \vee b) \wedge \neg a \mid\!\sim_{0} x \vee \neg a$ (reflexivity and RW); thus $a \vee b \mid\!\sim_{\varepsilon 1} x \vee \neg a$ (XOR and LCE). Similarly, applying LCE to the second supposition, $(a \vee b) \wedge b \mid\!\sim_{\varepsilon 2} x$, so $(a \vee b) \wedge b \mid\!\sim_{\varepsilon 2} x \vee a$ (right weakening); also $(a \vee b) \wedge \neg b \mid\!\sim_{0} x \vee a$ (reflexivity and RW); thus $a \vee b \mid\!\sim_{\varepsilon 2} x \vee a$ (XOR and LCE). Putting these together by Right∧+ we have $a \vee b \mid\!\sim_{\varepsilon 1+\varepsilon 2} (x \vee a) \wedge (x \vee \neg a)$ so by RW $a \vee b \mid\!\sim_{\varepsilon 1+\varepsilon 2} x$ as desired.

4.5 Summary on Inheritance of the Sum Bound

The sum bound for Right∧+ is thus inherited by the rules CM (cautious monotony), CT (cumulative transitivity) and Left∨+ (disjunction of premises), as may be shown by two different kinds of argument. One proceeds by analyzing the derivations of those rules in the system **O**-plus-Right∧+ for qualitative inference relations. Each derivation necessarily makes use of Right∧+, but with only one application required. The sum soundness of Right∧+ is then invoked, which is why we speak of inheritance of the property. The other kind of argument is directly arithmetic. Under this method of verification, appeal to the sum soundness of Right∧+ was made only in the case of CM so, from this more refined perspective the fortunes of CT and Left∨+ are actually self-earned rather than inherited. In the next section, we will see how the arithmetic arguments may be refined to yield improved bounds for CM, CT, Left∨+.

5 Improved Bounds for CM, CT, Left∨+

For all three derived rules, the sum bound may be tightened, and in this section we show how it may be done. We have reached the limit of the 'logical' method, and our arguments for the improved bounds are all arithmetic, although still elementary. They have a common feature: they all make use of the product or division of conditional probabilities, whereas the verifications of sum soundness used only addition and subtraction.

5.1 Improved Bound for CM

Using linear algebra, Bourne & Parsons [5, 6] established an improved bound for CM: when $p(x|a) \neq 0$, then $\text{imp}(y|a \wedge x) \leq [\text{imp}(x|a) + \text{imp}(y|a)]/p(x|a)$. In other words, it divides the sum bound, already known to hold for CM, by $p(x|a)$. Equivalently, in terms of probabilities: when $p(x|a) \neq 0$, then $p(y|a \wedge x) \geq [p(x|a) + p(y|a) - 1]/p(x|a)$. We call this the 'divided sum' bound for CM, and write its general form as $t \geq (r + s - 1)/r$, with the understanding that r is the probability $p(x|a)$ of the 'major' premise $a \mid\sim x$.

Clearly, whenever $p(x|a) < 1$, this is better than the plain sum bound, and in the limiting case that $p(x|a) = 1$ they are equal. When $p(x|a) = 0$ the quotient is undefined, but then $p(a \wedge x) = 0$ and thus, using the limiting case convention in the definition of conditional probability, $p(y|a \wedge x) = 1$ and $\text{imp}(y|a \wedge x) = 0$.

We verify the improved bound by elementary means. Recall that CM authorizes passage from $a \mid\sim x$ and $a \mid\sim y$ to $a \wedge x \mid\sim y$. Suppose $p(x|a) \neq 0$. We may assume wlog that $p(a \wedge x) \neq 0$, for otherwise $p(y|a \wedge x) = 1$ and we are done. Hence $p(y|a \wedge x) = p(a \wedge x \wedge y)/p(a \wedge x) = p(x \wedge y|a)/p(x|a)$, and we can replace the top $p(x \wedge y|a)$ by the sum bound for Right∧+ to get $p(y|a \wedge x) \geq [p(x|a) + p(y|a) - 1]/p(x|a)$ as desired.

Improved bound for CM When $p(x\|a) \neq 0$:
$\text{imp}(y\|a \wedge x) \leq [\text{imp}(x\|a) + \text{imp}(y\|a)]/p(x\|a)$
$p(y\|a \wedge x) \geq [p(x\|a) + p(y\|a) - 1]/p(x\|a)$ $t \geq (r + s - 1)/r$

The divided sum bound is optimal for CM, in the sense defined in Sect. 2. That is, for any $r, s \in (0, 1)$ there is a probability function p and elements, a, x, y of its domain with $p(x|a) = r$, $p(y|a) = s$, and $p(y|a \wedge x) = \max(0, (r + s - 1)/r)$. The verification is quite similar to that of the optimality of sum soundness for Right∧+ in Sect. 3.2. Let $r, s \in (0, 1)$, put $S = \{1, 2, 3\}$ as before, and again consider two cases. *Case* 1. Suppose $r + s \leq 1$. Put p, a, x, y as before. Then $p(y|a \wedge x) = p(a \wedge x \wedge y)/p(a \wedge x) = 0 \geq (r + s - 1)/r$ as desired. *Case* 2. Suppose $r + s > 1$. Put p, a, x, y as before. Then $p(y|a \wedge x) = p(a \wedge x \wedge y)/p(a \wedge x) = p(\{2\})/p(\{1, 2\}) = (r + s - 1)/r$ and the verification is complete.

5.2 *Improved Bound for CT*

Recall again that CT authorizes passage from $a \mid\!\sim x$ and $a \wedge x \mid\!\sim y$ to $a \mid\!\sim y$. In Sect. 4.3, we noted that $p(y|a) \geq p(x \wedge y|a) = p(x|a) \cdot p(y|a \wedge x)$, which is already a bound for CT. We also noted in the same section that this is at least as good as the sum bound, since by the 'product trumps sum' principle we have $p(x|a) \cdot p(y|a \wedge x) \geq p(x|a) + p(y|a \wedge x) - 1$. But that principle tells us a little more; it says that the last inequality is strict except when $p(x|a) = 1$ or $p(y|a \wedge x) = 1$. So $p(y|a) > p(x|a) + p(y|a \wedge x) - 1$ except when one of $p(x|a)$, $p(y|a \wedge x)$ equals 1.

By how much does the left exceed the right? We can calculate the improvement as follows. Let $p(x|a) = (1 - \varepsilon_1)$ and $p(y|a \wedge x) = (1 - \varepsilon_2)$. Then $p(x|a) \cdot p(y|a \wedge x) = (1 - \varepsilon_1) \cdot (1 - \varepsilon_2) = [1 - (\varepsilon_1 + \varepsilon_2)] + (\varepsilon_1 \cdot \varepsilon_2) = [(1 - \varepsilon_1) + (1 - \varepsilon_2) - 1] + (\varepsilon_1 \cdot \varepsilon_2) = [p(x|a) + p(y|a \wedge x) - 1] + [(1 - p(x|a)) \cdot (1 - p(y|a \wedge x))]$. The part in the left square parentheses is the sum bound, and the part in the right square parentheses is the improvement. Its general form is $t \geq r \cdot s = (r + s - 1) + [(1 - r) \cdot (1 - s)]$.

Interestingly, although the rule CT is not symmetric around its two premises, the bound is nevertheless symmetric around their probabilities—as indeed was the sum bound. The improved bound can be also expressed more concisely in terms of improbabilities: $\text{imp}(y|a) \leq \text{imp}(x|a) + \text{imp}(y|a \wedge x) - \text{imp}(x|a) \cdot \text{imp}(y|a \wedge x)$.

Improved bound for CT
$\text{imp}(y\|a) \leq \text{imp}(x\|a) + \text{imp}(y\|a \wedge x) - \text{imp}(x\|a) \cdot \text{imp}(y\|a \wedge x)$
$p(y\|a) \geq p(x\|a) \cdot p(y\|a \wedge x) = [p(x\|a) + p(y\|a \wedge x) - 1] + [(1 - p(x\|a)) \cdot (1 - p(y\|a \wedge x)]$ $t \geq r \cdot s = (r + s - 1) + [(1 - r) \cdot (1 - s)]$

The bound is optimal for CT. That is, for any $r, s \in (0, 1)$ there is a probability function p and elements a, x, y of its domain with $p(x|a) = r$, $p(y|a \wedge x) = s$, such that $p(y|a) = r \cdot s$. To verify this, let $r, s \in (0, 1)$, again take $S = \{1, 2, 3\}$, put $p(\{1\}) = r \cdot s$, $p(\{2\}) = r - r \cdot s$, $p(\{3\}) = 1 - r$. These sum to 1 and so determine a probability function on 2^S. Put $a = S$, $x = \{1, 2\}$, $y = \{1\}$. Then $p(x|a) = p(a \wedge x)/p(a) = r$, $p(y|a \wedge x) = p(a \wedge x \wedge y)/p(a \wedge x) = r \cdot s/r = s$, while $p(y|a) = p(a \wedge y)/p(a) = r \cdot s$ and we are done.

5.3 *Comparison of the Improved Bounds for CM and CT*

What is the relationship between the loss levels of the rules Right∧+, CM, and CT? In general terms, one can say that Right∧+ is lossier than either of CM and CT, which are incomparable in this respect.

To verify this, first recall that sum soundness is an *optimal* bound for Right∧+ and also a bound for CM and CT; so that Right∧+ is at least as lossy as the latter two. We also know that CM satisfies the 'divided sum' bound $(r+s-1)/r$, and that CT satisfies the 'product of premises' bound $r \cdot s$. So it will suffice to show that neither Right∧+ nor CM satisfies a product of premises bound, and that neither Right∧+ nor CT satisfies a divided sum bound.

For the former, we need to check that it can happen that for each of the rules Right∧+ and CM, the conclusion of the rule has probability *less* than $r \cdot s$ where r, s are the probabilities of the two premises of the rule. For a simple example covering both rules, let p be any probability function with each $p(\pm x \wedge \pm y) > 0$ and put $a = \neg(x \wedge y)$. With r, s being the values $p(x|a)$, $p(y|a)$ of the premises of the rule, we have $r, s > 0$ so $r \cdot s > 0$, while the values $p(x \wedge y|a)$, $p(y|a \wedge x)$ of the respective conclusions of the two rules both equal 0.

For the latter, we need to check that it can happen that for each of the rules Right∧+ and CT, the conclusion of the rule has probability less than $(r+s-1)/r$ where $r = p(x|a)$ is the probability of the premise $a \mid\!\sim x$ (common to both rules) and s is the probability of the other premise of the rule, that is, $s = p(y|a)$ in the case of Right∧+ and $s = p(y|a \wedge x)$ in the case of CT. An example covering Right∧+ may be obtained as follows: let $S = \{1, 2, 3, 4\}$ and $p(\{i\}) = 1/4$ for each $i \in S$. Put $a = S$, $x = \{1, 2, 3\}$, $y = \{3, 4\}$. Then $r = p(x|a) = 3/4$, $s = p(y|a) = 1/2$, $p(x \wedge y|a) = 1/4$, so that $p(x \wedge y|a) < 1/3 = (r+s-1)/r$. For CT we take the same values for S, p, a, x but put $y = \{2, 3\}$. Then $r = p(x|a) = 3/4$, $s = p(y|a \wedge x) = 2/3$, $p(y|a) = 1/2$, so that $p(y|a) < 5/9 = (r+s-1)/r$.

This relationship between the bounds for Right∧+, CM and CT is what one might have anticipated given the qualitative connections between those rules. While each of CM and CT is derivable in the system **O**-plus-Right∧+, neither Right∧+ nor CM is derivable in **O**-plus-CT, nor are Right∧+ or CT derivable in **O**-plus-CM, as shown in Hawthorne and Makinson [10, Observation 4.1]. To be sure, this consonance is a matter of confirmed expectation rather than implication since, apart from anything else, Hawthorne's system **O** for probabilistic inference between Horn rules is known not to be complete, as was shown by Paris and Simmonds [15] with a brief overview in Hawthorne [9].

5.4 Improved Bound for **Left∨+**

Adams [2, Table I and Appendix] gave an intricate and indirect proof that Left∨+ also satisfies the product bound $p(x|a \vee b) \geq p(x|a) \cdot p(x|b)$ and thus, in terms of improbability, $\mathrm{imp}(x|a \vee b) \leq \mathrm{imp}(x|a) + \mathrm{imp}(x|b) - \mathrm{imp}(x|a) \cdot \mathrm{imp}(x|b)$. This is already an improvement on the plain sum bound on Left∨+ derived in Sect. 4.4 from the sum bound on Right∧+. The same result can also be obtained from the product bound for CT by a careful analysis of the derivation of Left∨+ from **O**-plus-CT that is given in Hawthorne and Makinson [10, Appendix, Fact 4.1.4].

But a better bound was obtained by Gilio [7, p. 23] using linear algebra applied to de Finetti's theory of coherent probability assessments. His bound is:

$$p(x|a \vee b) \geq p(x|a) \cdot p(x|b) / \big[p(x|a) + p(x|b) - p(x|a) \cdot p(x|b)\big]$$

whenever this ratio is well-defined, that is, when either $p(x|a) \neq 0$ or $p(x|b) \neq 0$. This is the product bound divided by $p(x|a) + p(x|b) - p(x|a) \cdot p(x|b)$, and has the symmetric general form $r \cdot s/(r + s - r \cdot s)$. By the 'product trumps sum' principle, the divisor is less than 1 whenever both $p(x|a)$ and $p(x|b)$ are less than 1, so this bound for Right∨+ is better than the simple product $p(x|a) \cdot p(x|b)$ except in the limiting case that one of $p(x|a)$, $p(x|b)$ equals 1, in which case the two bounds are equal. As Gilio notes (p. 26), his bound may likewise be written using improbabilities, in the rather more complex expression also recorded in the display.

Gilio's Improved bound for Left∨+ When either $p(x|a) \neq 0$ or $p(x|b) \neq 0$
$p(x|a \vee b) \geq p(x|a) \cdot p(x|b)/[p(x|a) + p(x|b) - p(x|a) \cdot p(x|b)]$ $r \cdot s/(r + s - r \cdot s)$
$\text{imp}(x|a \vee b) \leq [\text{imp}(x|a) + \text{imp}(x|b) - 2(\text{imp}(x|a) \cdot \text{imp}(x|b))]/1 - (\text{imp}(x|a) \cdot \text{imp}(x|b))$

We show how Gilio's bound for Left∨+ may be verified without either linear algebra or detour through de Finetti's rather non-standard theory of coherent probability assessments. The argument is elementary but rather intricate; readers more interested in the result than its proof may wish to skip to the next section.

First, we may assume wlog that $p(a \vee b) \neq 0$, for otherwise $p(x|a \vee b) = 1$ and the Gilio inequality holds trivially. Indeed, we may assume wlog that $p(a \vee b) = 1$. For, given $p(a \vee b) \neq 0$, if LHS < RHS for a probability function p then, putting $q = p_{a \vee b}$, we have $q(x|c) = p(x|c)$ for all $c \models a \vee b$ so that for q we have LHS < RHS while $q(a \vee b) = 1$.

Thus, under the assumption $p(a \vee b) = 1$, we need only show $p(x) \geq p(x|a) \cdot p(x|b)/[p(x|a) + p(x|b) - p(x|a) \cdot p(x|b)]$, that is,

$$p(x) \cdot \big[p(x|a) + p(x|b) - p(x|a) \cdot p(x|b)\big] \geq p(x|a) \cdot p(x|b). \tag{1}$$

Again, we dispose of some limiting cases. When $p(a) = 0$ then (1) holds. For suppose $p(a) = 0$. Then $p(x|a) = 1$ so RHS(1) $= p(x|b)$; also $p(b) = p(a \vee b) + p(a \wedge b) - p(a) = p(a \vee b) = 1$, so that $p(x) = p(x \wedge b) + p(x \wedge \neg b) = p(x \wedge b) = p(x|b)$, and thus also LHS(1) $= p(x|b)$. Similarly, (1) holds when $p(b) = 0$. So we may assume wlog that $p(a) \neq 0 \neq p(b)$. Under that assumption, when $p(x) = 0$ we have LHS(1) $= 0 =$ RHS(1), so we may further suppose that $p(x) \neq 0$.

Next, we restate the problem. Multiplying both sides of (1) by the non-zero $p(a) \cdot p(b)$, it suffices to show:

$$p(x) \cdot p(a) \cdot p(b)\big[p(x|a) + p(x|b) - p(x|a) \cdot p(x|b)\big] \geq p(x|a) \cdot p(x|b) \cdot p(a) \cdot p(b).$$

Distributing, turning the left minus into a right plus, and eliminating conditional probabilities, it thus suffices to show:

$$p(x) \cdot \big[p(a \wedge x) \cdot p(b) + p(b \wedge x) \cdot p(a)\big] \geq p(a \wedge x) \cdot p(b \wedge x) \cdot p(x) + p(a \wedge x) \cdot p(b \wedge x). \tag{2}$$

To complete the verification, we show that in (2), the left and right sides may be expressed as $\alpha_1 + \alpha_2 + \alpha_3$ and $\beta_1 + \beta_2 + \beta_3$ respectively, with each $\alpha_i \geq \beta_i$. We make use of the fact that since $p(a \vee b) = 1$, we can 'get rid' of $p(x)$ by the equality $p(x) = p((a \vee b) \wedge x) = p((a \wedge x) \vee (b \wedge x)) = p(a \wedge x) + p(b \wedge x) - p(a \wedge b \wedge x)$.

Substituting for $p(x)$ on the left, distributing and regrouping, $\text{LHS}(2) = \alpha_1 + \alpha_2 + \alpha_3$ where:

$$\alpha_1 = p(a \wedge x) \cdot p(b) \cdot \big[p(a \wedge x) - p(a \wedge b \wedge x)\big],$$
$$\alpha_2 = p(b \wedge x) \cdot p(a) \cdot \big[p(b \wedge x) - p(a \wedge b \wedge x)\big],$$
$$\alpha_3 = p(a \wedge x) \cdot p(b) \cdot p(b \wedge x) + p(b \wedge x) \cdot p(a) \cdot p(a \wedge x).$$

Substituting for $p(x)$ in the first term $p(a \wedge x) \cdot p(b \wedge x) \cdot p(x)$ on the right of (2) and distributing transforms that term into:

$$p(a \wedge x) \cdot p(b \wedge x) \cdot p(a \wedge x) + p(a \wedge x) \cdot p(b \wedge x) \cdot p(b \wedge x)$$
$$- \, p(a \wedge x) \cdot p(b \wedge x) \cdot p(a \wedge b \wedge x)$$

so that, by subtracting and adding an additional $p(a \wedge x) \cdot p(b \wedge x) \cdot p(a \wedge b \wedge x)$, distributing, and finally adding the second term from the right of (2), we have $\text{RHS}(2) = \beta_1 + \beta_2 + \beta_3$ where:

$$\beta_1 = p(a \wedge x) \cdot p(b \wedge x) \cdot \big[p(a \wedge x) - p(a \wedge b \wedge x)\big],$$
$$\beta_2 = p(a \wedge x) \cdot p(b \wedge x) \cdot \big[p(b \wedge x) - p(a \wedge b \wedge x)\big],$$
$$\beta_3 = p(a \wedge x) \cdot p(b \wedge x) \cdot p(a \wedge b \wedge x) + p(a \wedge x) \cdot p(b \wedge x).$$

Clearly, since $p(b) \geq p(b \wedge x)$ and $p(a) \geq p(a \wedge x)$ we have $\alpha_1 \geq \beta_1$ and $\alpha_2 \geq \beta_2$. It remains to show that $\alpha_3 \geq \beta_3$. Distributing, we have $\alpha_3 = p(a \wedge x) \cdot p(b \wedge x) \cdot [p(a) + p(b)]$. Under our wlog assumption that $p(a \vee b) = 1$, we have $p(a) + p(b) = 1 + p(a \wedge b) \geq p(a \wedge b \wedge x) + 1$. Thus $\alpha_3 \geq p(a \wedge x) \cdot p(b \wedge x) \cdot [p(a \wedge b \wedge x) + 1] = p(a \wedge x) \cdot p(b \wedge x) \cdot p(a \wedge b \wedge x) + p(a \wedge x) \cdot p(b \wedge x) = \beta_3$, and the verification is complete.

5.5 Comments on the Improved Bound for **Left∨+**

It is difficult to believe that there is no simpler elementary verification of Gilio's bound, but the authors have not been able to find one. However, we can show that the bound is optimal, as follows.

Let $r, s \in (0, 1)$ with $r \leq s$; we want to find a probability function p and elements a, b, x in its domain such that (i) $p(x|a) = r$, (ii) $p(x|b) = s$, (iii) $p(x|a \vee b) = r \cdot s/(r + s - r \cdot s)$. We show, more specifically, that this can be done with also $x = a \wedge b$ and $p(a \vee b) = 1$.

First step. Find p, a, b such that $p(a \vee b) = 1$ and (i), (ii) hold with $x = a \wedge b$, that is, (i)′ $p(a \wedge b|a) = r$, (ii)′ $p(a \wedge b|b) = s$. To do this, draw a line of unit length, extend it to the left until $1/(1+\text{left-extension}) = r$, that is, the left-extension is of length $(1-r)/r$; do the same to the right until $1/(1 + \text{right-extension}) = s$, as in the accompanying diagram. Take S to be the line from the leftmost point to the rightmost one, a to be the initial unit line plus its left extension, b to be the same but to the right (notice that $a \wedge b$ is now the initial unit line), and measure probabilities of lines by their length relative to the total length of the extended line. Then $p(a \vee b) = 1$ and (i)$'$, (ii)$'$ hold.

Diagram for optimality of Gilio's bound for Left$\vee$+		
Left-extension $(1-r)/r$	1	Right-extension $(1-s)/s$
a		
	b	

Second step. Show that whenever $p(a \vee b) = 1$ and (i)$'$, (ii)$'$ hold then (iii) holds with $x = a \wedge b$, that is, (iii)$'$ $p(a \wedge b|a \vee b) = p(a \wedge b|a) \cdot p(a \wedge b|b)/[p(a \wedge b|a) + p(a \wedge b|b) - p(a \wedge b|a) \cdot p(a \wedge b|b)]$. To do this, re-run the 'it-suffices-to-show' part of the verification of Gilio's bound, that is, up to display (2), but with $a \wedge b$ in place of x and $=$ in place of $\leq$. It thus suffices to show that (2) holds with those modifications, that is,

$$\big[p(a \wedge b) \cdot p(b) + p(a \wedge b) \cdot p(a)\big] \cdot p(a \wedge b) = p(a \wedge b)^3 + p(a \wedge b)^2.$$

But LHS $= p(a \wedge b)^2 \cdot (p(a) + p(b)) = p(a \wedge b)^2 \cdot (p(a \vee b) + p(a \wedge b)) = p(a \wedge b)^2 \cdot (1 + p(a \wedge b)) =$ RHS as desired, and the verification of optimality is complete.

How does this optimal bound for Left$\vee$+ compare with those obtained for CM and CT? We already remarked at the beginning of Sect. 5.4 that $(r \cdot s)/(r + s - r \cdot s) > r \cdot s$ except in the limiting case that one of r, s equals 1, in which case the two sides are equal, so the bound for Left$\vee$+ is higher than that for CT, in other words, CT is lossier than Left$\vee$+, except in that limiting case.

On the other hand, neither Left$\vee$+ nor CM is lossier than the other. In Sect. 5.3, we already gave a simple example where the conclusion of CM has probability less than $r \cdot s$ and so less than the lower bound $(r \cdot s)/(r + s - r \cdot s)$ for Left$\vee$+. It remains to give an example where the conclusion of Left$\vee$+ has probability less than the lower bound $(r + s - 1)/r$ of CM. Put $S = \{1, 2, 3, 4, 5\}$, $a = \{1, 2, 3, 4\}$, $b = \{2, 3, 4, 5\}$, $x = \{2, 3, 4\}$ and let $p(\{i\}) = 1/5$ for each $i \in S$. Then $r = p(x|a) = 3/4$, $s = p(x|b) = 3/4$ so that $(r + s - 1)/r = 2/3$, while $p(x|a \vee b) = 3/5 < 2/3$ as desired.

These facts are again consonant with the qualitative ones in Hawthorne and Makinson [10, Observation 4.1], where it was shown that while Left$\vee$+ is derivable in the system **O**-plus-CT, it is not derivable in the system **O**-plus-CM, nor is CM derivable in **O**-plus-Left$\vee$+.

5.6 Summary on Improved Bounds for CM, CT, Left$\vee$+

Thus the rules CM, CT and Left$\vee$+ all have optimal bounds that do better than the sum bound, with Left$\vee$+ also doing better than CT. They are: 'divided sum' $(r + s - 1)/r$ for CM due to Bourne & Parsons [5, 6]; 'product' $r \cdot s$ for CT going back to Adams [3]; and the more complex bound $(r \cdot s)/(r + s - r \cdot s)$ due to Gilio [7] for Left$\vee$+. The bounds for CM and Left$\vee$+ were originally established using linear algebra but can also be verified by elementary methods, as done above. In the case of Left$\vee$+, this elementary verification is quite intricate, and it would be agreeable to find a simpler one.

6 Lossiness for Non-Horn Rules

We now turn to some well-known non-Horn rules that are classically correct and are sometimes added to strengthen the Kraus–Lehmann system **P** of qualitative uncertain inference.

6.1 Probabilistic Soundness for Almost-Horn Rules

The rule NR of negation rationality (whenever $a \mid\sim x$, then either $a \wedge b \mid\sim x$ or $a \wedge \neg b \mid\sim x$) is known to be probabilistically sound and so has zero loss—see the easy verification in Hawthorne [8] or Hawthorne & Makinson [10]. But the stronger rules of disjunctive rationality DR (whenever $a \vee b \mid\sim x$, then either $a \mid\sim x$ or $b \mid\sim x$) and rational monotony RM (whenever $a \mid\sim x$, then either $a \mid\sim \neg b$ or $a \wedge b \mid\sim x$) are not probabilistically sound. How lossy are they?

Before attempting to answer this question, we need to be quite clear about the notion of probabilistic soundness for such rules. All three take the form: $a_i \mid\sim x_i$ (for all $i \leq n)/b_j \mid\sim y_j$ (for at least one $j \leq m$), where $n \geq 0$ and $m \geq 1$. We will call rules of that form *almost-Horn*. Thus the right side of an almost-Horn rule may be multiple, but we require that it is never empty, i.e. that $m \geq 1$. We say that an almost-Horn rule is *probabilistically sound* iff $\min_{i \leq n}\{p(x_i|a_i)\} \leq \max_{j \leq m}\{p(y_j|b_j)\}$ for every probability function p. In other words, iff for every such function p and every $t \in [0, 1]$, if $p(x_i|a_i) \geq t$ for all $i \leq n$, then $p(y_j|b_j) \geq t$ for some $j \leq m$. Writing $a \mid\sim_{p,t} x$ for $p(x|a) \geq t$, this abbreviates to: if $a_i \mid\sim_{p,t} x_i$ for all $i \leq n$ then $b_j \mid\sim_{p,t} y_j$ for some $j \leq m$. In the special case that $m = 1$, the almost-Horn rules are exactly the Horn rules, and the definition of probabilistic soundness agrees with the earlier definition. For the three non-Horn rules NR, DR and RM presented above, $n = 1$ and $m = 2$.

The above formulation may be called the *distributed* (or alternate) way of expressing an almost-Horn rule and its probabilistic soundness. As is well known, any such rule may equivalently be expressed with negative premises replacing some (but not all) of the alternates in the conclusion. For example, with suitable choices of which item in the conclusion to shift, NR becomes: whenever $a \mid\sim x$ and $a \wedge b \not\mid\sim x$, then $a \wedge \neg b \mid\sim x$; DR takes the form: whenever $a \vee b \mid\sim x$ and $a \not\mid\sim x$, then $b \mid\sim x$; and RM reads: whenever $a \mid\sim x$ and $a \not\mid\sim \neg b$, then $a \wedge b \mid\sim x$. This may be called the *negative* way of formulating an almost-Horn rule.

When all but one of the alternates of the conclusion has been transformed into a negative premise, we say that the negative formulation is *pointed*. Thus for NR, DR and RM, which have only two alternates in their distributed form, their negative presentations are automatically pointed; while in the general case where there may be more than two alternates, that is not the case. For simplicity, we focus on the distributed and pointed presentations. In general, in its pointed presentation an almost-Horn rule takes the form: $a_i \mid\sim x_i$ (for all $i \leq n$), $b_j \not\mid\sim y_j$ (for all $j < m$, $j \neq k)/b_k \mid\sim y_k$, where $n \geq 0$ and $m \geq 1$.

The pointed presentation gives rise to a trivially equivalent definition of probabilistic soundness: if $\min_{i \leq n}\{p(x_i|a_i)\} > \max_{j \leq m, j \neq k}\{p(y_j|b_j)\}$ then $\min_{i \leq n}\{p(x_i|a_i)\} \leq p(y_k|b_k)$. Equivalently and perhaps more intuitively, for all $t \in [0, 1]$, if $\min_{i \leq n}\{p(x_i| a_i)\} \geq t$ but $\max_{j \leq m, j \neq k}\{p(y_j|b_j)\} < t$ then $p(y_k|b_k) \geq t$.

But, while they are equivalent, the distributed and pointed formulations give rise to quite different ways of measuring loss, as we explain in the next section.

6.2 Distributed vs. Pointed Loss

We begin by illustrating the difference between distributed and pointed loss with the examples of DR and RM, and then state the difference in general form.

For DR, distributed loss is naturally measured by the drop from $p(x|a \vee b)$ to $\max\{p(x|a), p(x|b)\}$, with no loss when $\max\{p(x|a), p(x|b)\} \geq p(x|a \vee b)$. This is straightforward. But pointed loss is more subtle, and moreover depends on which alternate is transformed into a negative premise. Suppose that it is the alternate $a \mid\sim x$ that is transformed into a negative premise $a \not\mid\sim x$. Then a natural notion of pointed loss would consider the fall from $p(x|a \vee b)$ to $p(x|b)$ in situations where $p(x|a \vee b) > p(x|a)$. Similarly with interchanged variables when the alternate $b \mid\sim x$ is transformed into a negative premise $b \not\mid\sim x$.

For RM, distributed loss is again straightforward, measured by the drop from $p(x|a)$ to $\max\{p(\neg b|a), p(x|a \wedge b)\}$. Pointed loss again depends on which alternate is transformed into a negative premise and, as the rule is not symmetric, the dependence is more important. Suppose that it is the alternate $a \mid\sim \neg b$ that is transformed into a negative premise $a \not\mid\sim \neg b$. Then a natural definition of pointed loss would consider the fall from $p(x|a)$ to $p(x|a \wedge b)$ in situations where $p(x|a) > p(\neg b|a)$. Suppose, on the other hand that the alternate $a \wedge b \mid\sim x$ is transformed into a negative premise $a \wedge b \not\mid\sim x$. Then the pointed loss would consider the passage from $p(x|a)$ to $p(\neg b|a)$ in situations where $p(x|a) > p(x|a \wedge b)$.

In general, our proposed definitions for almost-Horn rules are as follows. Consider any almost-Horn rule, written in distributed form as $a_i \mid\sim x_i$ (for all $i \leq n$)/$b_j \mid\sim y_j$ (for at least one $j \leq m$), or equivalently in pointed form as $a_i \mid\sim x_i$ (for all $i \leq n$), $b_j \not\mid\sim y_j$ (for all $j < m, j \neq k$)/$b_k \mid\sim y_k$, where $n \geq 0$ and $m \geq 1$.

- *Distributed loss* considers the drop from $\min_{i \leq n}\{p(x_i|a_i)\}$ to $\max_{j \leq m}\{p(y_j|b_j)\}$. A *bound statement* for distributed loss takes the form $\max_{j \leq m}\{p(y_j|b_j)\} \geq f(\{p(x_i|a_i)_{i \leq n}\})$ for some n-argument function f expressed in purely arithmetic terms.
- *Pointed loss* considers the drop from the same $\min_{i \leq n}\{p(x_i|a_i)\}$ but to $p(y_k|b_k)$ in situations where $\min_{i \leq n}\{p(x_i|a_i)\} > \max\{p(y_j|b_j)_{j \leq m, j \neq k}\}$. A *bound statement* for it would take the form $p(y_k|b_k) \geq g(\{p(x_i|a_i)_{i \leq n}, p(y_j|b_j)_{j \leq m, j \neq k}\})$ for some $(n + m - 1)$-argument function g expressed in purely arithmetic terms.

Care should be taken about conflating n-argument functions f for distributed bounds with $(n + m - 1)$-argument functions g for a pointed ones. We will see shortly how this plays out for RM and DR.

Which measure of loss is more appropriate for an almost-Horn rule—distributed or pointed and, in the latter case, pointed to which component? There does not appear to be a uniform answer. On the one hand, for RM pointed loss appears to be more natural, since the rule is most naturally and commonly thought of as a weakened version of monotony with a negative premise: whenever $a \mid\sim x$ and $a \not\mid\sim \neg b$, then $a \wedge b \mid\sim x$. On the other hand, distributed loss seems more appropriate for the rule DR, which is normally formulated in a symmetric manner with a multiple conclusion rather than a negative premise.

6.3 Pointed Bound for RM

First, we consider RM from its natural pointed perspective, that is, proceeding from $a \mid\sim x$ and $a \not\mid\sim \neg b$ to $a \wedge b \mid\sim x$. This is closely related to the Horn rule CM which, we recall, goes from $a \mid\sim x$ and $a \mid\sim b$ to $a \wedge b \mid\sim x$. Syntactically, CM can be seen as formed by taking the negative premise $a \not\mid\sim \neg b$ of RM and changing the denial into an affirmation $a \mid\sim b$ with opposite right side. From the probabilistic point of view, this amounts to replacing the requirement that $p(\neg b|a)$ is below a threshold to the requirement that $p(b|a)$ at least reaches it. From the point of view of preferential consequence, it amounts to replacing an existential condition (at least one minimal a-state is a b-state) to the corresponding universal one (all minimal a-states are b-states).

We can exploit this relationship to get a pointed bound for RM directly from any bound for CM. In particular, the improved bound for CM (Sect. 5.1) tells us that when $p(b|a) \neq 0$ then $p(x|a \wedge b) \geq [p(x|a) + p(b|a) - 1]/p(b|a)$. Rewriting $p(b|a)$ as $1 - p(\neg b|a)$ and simplifying a little, this becomes a pointed bound for RM, as in the display.

Pointed bound for RM
When $p(\neg b|a) \neq 1$
$p(x|a \wedge b) \geq [p(x|a) - p(\neg b|a)]/[1 - p(\neg b|a)]$

The proviso, needed to ensure that the RHS is well defined, does not exclude any cases of interest to the rule RM when it is understood as pointed towards $a \wedge b \mid\sim x$. For when so understood, the only case of interest is that where $a \mid\sim x$ while $a \not\mid\sim \neg b$, which implies that $p(x|a) > p(\neg b|a)$ so in particular $p(\neg b|a) \neq 1$.

It is clear from this bound that when $p(x|a) > p(\neg b|a)$, then $p(x|a \wedge b)$ is never zero. Nevertheless, it may get arbitrarily close to zero. To see this, let $\varepsilon > 0$; we want to find p, a, b, x with $p(x|a) > p(\neg b|a)$ and $p(x|a \wedge b) < \varepsilon$. Take any positive integer n with $\varepsilon > 1/n$. Put $S = \{1, \ldots, 2n\}$, put $p(i) = 1/2n$ for all $i \in S$. Take $a = \{1, \ldots, 2n\}$, $b = \{1, \ldots, n\}$, $x = \{n, \ldots, 2n\}$. Then $p(x|a) = (n + 1)/2n > n/2n = p(\neg b|a)$ and $p(x|a \wedge b) = 1/n < \varepsilon$ as required.

6.4 Distributed Bound for RM

RM has a very simple distributed bound, namely: $\max\{p(\neg b|a), p(x|a \wedge b)\} > p(x|a)/2$ except when LHS $= 0 =$ RHS. Unlike the bounds formulated earlier, it is a strict inequality.

Distributed bound for RM
$\max\{p(\neg b|a), p(x|a \wedge b)\} > p(x|a)/2$, except when LHS $= 0 =$ RHS

To prove this, we again apply the improved bound for CM, but with a less immediate argument. Suppose that $\max\{p(\neg b|a), p(x|a \wedge b)\} \neq 0$ and that $p(\neg b|a) \leq p(x|a)/2$; we need to show $p(x|a \wedge b) > p(x|a)/2$.

We begin by disposing of several limiting cases. (i) We may assume wlog that $p(a) \neq 0$; otherwise LHS $= 1 > 1/2 =$ RHS as needed. (ii) We may assume wlog that $p(b|a) \neq 1$; otherwise, using (i), $p(\neg b|a) = 0$ and so by the first supposition, $0 \neq p(x|a \wedge b) = p(x|a) > p(x|a)/2$ as needed. (iii) We have $p(b|a) \geq 0.5 > 0$; otherwise $p(\neg b|a) > 0.5$ so by the second supposition $p(x|a) > 1$. (iv) Finally, $p(x|a) \neq 0$; otherwise $p(x|a)/2 = 0$ so that by the second supposition $p(\neg b|a) = 0$, contrary to (ii).

We can now proceed to the main argument. Since $p(b|a) \neq 0$ as noted in (iii), the improved bound for CM with variables relabeled tells us that $p(x|a \wedge b) \geq [p(x|a) + p(b|a) - 1]/p(b|a)$. By the second supposition, $p(b|a) \geq 1 - p(x|a)/2$, so $p(x|a) + p(b|a) - 1 \geq p(x|a)/2$, so $p(x|a \wedge b) \geq p(x|a)/(2 \cdot p(b|a)) > p(x|a)/2$ since $p(x|a) \neq 0$ by (iv) and $p(b|a) \neq 0$ by (iii) again, and we are done.

It is optimal in the following sense: for every $\varepsilon > 0$ there is a probability function p and elements a, b, x in its domain such that $\max\{p(x|a \wedge b), p(\neg b|a)\} < p(x|a)/2 + \varepsilon$. To verify this, let $\varepsilon > 0$, so there is a positive integer n such that $\varepsilon > 1/n$. Take $S = \{1, 2, 3\}$ with $p(\{1\}) = p(\{3\}) = 1/n$ and $p(\{2\}) = (n-2)/n$, and put $a = \{1, 2, 3\}$, $b = \{2, 3\}$, $x = \{1, 3\}$. Then $p(x|a) = 2/n$ so $p(x|a)/2 = 1/n$ while $p(x|a \wedge b) = 1/n \div (n-1)/n = 1/(n-1)$ and $p(\neg b|a) = 1/n$, so $\max\{p(x|a \wedge b), p(\neg b|a)\} = p(x|a \wedge b) = 1/(n-1)$. It remains to check that $1/(n-1) < 1/n + \varepsilon$, i.e. that $\varepsilon > 1/(n-1) - 1/n$. But $1/(n-1) - 1/n = n - (n-1)/n \cdot (n-1) < 1/n < \varepsilon$ by the choice of n, and we are done.

6.5 Distributed Bound for DR

As is well known, in the context of the preferential system **P** the rule RM is strictly stronger than DR, and one might be tempted to conclude that every lower bound for RM will *ipso facto* be one for DR. But the standard derivation of DR from RM, given in Lehmann and Magidor [12], uses not only RM and the probabilistically sound rules of **O**, but also Right$\wedge$+. It is not clear whether there is a derivation of DR that uses resources from **O** plus just one application of RM, which would be needed even to begin justifying such a conclusion.

Nevertheless, DR does happen to satisfy the same kind of (strict) distributed bound as we found for RM: $\max\{p(x|a), p(x|b)\} > p(x|a \vee b)/2$ unless LHS $= 0 =$ RHS.

The limiting case is immediate. To verify the principal case, assume wlog $p(x|a) \leq p(x|b)$, so $p(x|b) \neq 0$; we need to show that $p(x|a \vee b) < 2 \cdot p(x|b)$. First note that $p(x|a \vee b) \leq p(x|b) \cdot [p(a|a \vee b) + p(b|a \vee b)]$ by the following chain: $p(x|a \vee b) = p((x \wedge a) \vee (x \wedge b)|a \vee b) = p(x|a) \cdot p(a|a \vee b) + p(x|b) \cdot p(b|a \vee b) - p(x \wedge a \wedge b|a \vee b) \leq p(x|b) \cdot p(a|a \vee b) + p(x|b) \cdot p(b|a \vee b) = p(x|b) \cdot [p(a|a \vee b) + p(b|a \vee b)]$. In the case that $p(a|a \vee b) + p(b|a \vee b) < 2$, this gives directly $p(x|a \vee b) < 2 \cdot p(x|b)$ as desired. In the case $p(a|a \vee b) + p(b|a \vee b) = 2$ we have $p(a|a \vee b) = 1 = p(b|a \vee b)$ so $p(x|a) = p(x|a \vee b) = p(x|b)$ and again $p(x|a \vee b) < 2 \cdot p(x|b)$, completing the verification.

Distributed bound for DR
$\max\{p(x|a), p(x|b)\} > p(x|a \vee b)/2$, except when LHS $= 0 =$ RHS

It is optimal in the same sense: for every $\varepsilon > 0$ there is a probability function p and elements a, b, x in its domain such that $\max\{p(x|a), p(x|b)\} < p(x|a \vee b)/2 + \varepsilon$. To verify this, let $\varepsilon > 0$. Then there is a positive integer n such that $\varepsilon > 1/n$. As before, take $S = \{1, 2, 3\}$ with $p(\{1\}) = p(\{3\}) = 1/n$ and $p(\{2\}) = (n-2)/n$, but his time put $a = \{1, 2\}$, $b = \{2, 3\}$, $x = \{1, 3\}$. Then $p(x|a) = p(x|b) = 1/n \div (n-1)/n = 1/(n-1)$ while $p(x|a \vee b) = 2/n$ so $p(x|a \vee b)/2 = 1/n$. But $1/(n-1) < 1/n + \varepsilon$ as already checked, so we are done.

6.6 Pointed Bound for DR

Finally, we consider bounds for pointed disjunctive rationality (DR): whenever $a \vee b \mid\sim x$ and $a \mid\not\sim x$, then $b \mid\sim x$. Just as pointed RM may be seen as a non-Horn counterpart of the Horn rule CM, so too we can regard DR as a non-Horn counterpart of a Horn rule taking us from $a \vee b \mid\sim x$ and $a \mid\sim \neg x$ to $b \mid\sim x$. As far as the authors are aware, this rule has not been studied, named, or even articulated in the literature; we call it *disjunctive choice* (DC). It can easily be derived from **O** plus just one application of Right$\wedge$+ so that it satisfies a sum bound. To check this we use the same subscript notation as we did in Sect. 4 for similar verifications, and the same acronyms for the rules applied.

Suppose $a \vee b \mid\sim_{\varepsilon 1} x$, $a \mid\sim_{\varepsilon 2} \neg x$; we want to get $b \mid\sim_{\varepsilon 1+\varepsilon 2} x$. By LCE and RW on the second supposition we have $(a \vee b) \wedge a \mid\sim_{\varepsilon 2} \neg a \vee \neg x$. Also $(a \vee b) \wedge \neg a \models \neg a \vee \neg x$ so $(a \vee b) \wedge \neg a \mid\sim_0 \neg a \vee \neg x$, so by XOR $a \vee b \mid\sim_{\varepsilon 2} \neg a \vee \neg x$. Hence by Right$\wedge$+ with the first supposition, $a \vee b \mid\sim_{\varepsilon 1+\varepsilon 2} (\neg a \vee \neg x) \wedge x$ so by RW $a \vee b \mid\sim_{\varepsilon 1+\varepsilon 2} \neg a \wedge x$. Clearly, $(a \vee b) \wedge \neg(a \vee b) \mid\sim_0 (a \vee b)$, so by WAND, $a \vee b \mid\sim_{\varepsilon 1+\varepsilon 2} (\neg a \wedge x) \wedge (a \vee b)$ and thus by RW $a \vee b \mid\sim_{\varepsilon 1+\varepsilon 2} b \wedge x$ and finally by VCM $(a \vee b) \wedge b \mid\sim_{\varepsilon 1+\varepsilon 2} x$ so that $b \mid\sim_{\varepsilon 1+\varepsilon 2} x$ by LCE.

Thus $\text{imp}(x|b) \le \text{imp}(x|a \vee b) + \text{imp}(\neg x|a)$, in other words DC satisfies the sum bound. In positive terms, $p(x|b) \ge p(x|a \vee b) + p(\neg x|a) - 1$. Rewriting $p(\neg x|a)$ as $1 - p(x|a)$, we thus have a simple bound for DR: $p(x|b) \ge p(x|a \vee b) - p(x|a)$.

But one can do better, obtaining the result on display by deploying Gilio's improved bound for Left$\vee$+. It is a better bound than $p(x|b) \ge p(x|a \vee b) - p(x|a)$ because its bottom is always less than 1 except when $p(x|a) = 0$, in which limiting case the two bounds coincide.

Pointed bound for DR
$p(x|b) \ge [p(x|a \vee b) - p(x|a)]/[p(x|a) \cdot \{p(x|a \vee b) - 2\} + 1]$

To derive the displayed bound, we may suppose wlog that $p(\neg x|b) \ne 0$; otherwise $p(x|b) = 1$ and the inequality holds. Take Gilio's bound for Left$\vee$+ in its improbability version, substitute $\neg x$ for x throughout, and rewrite $\text{imp}(\neg \cdot|\cdot)$ as $p(\cdot|\cdot)$. Observing that the precondition is satisfied since $p(\neg x|b) \ne 0$, we thus have

$$p(x|a \vee b) \le \left[p(x|a) + p(x|b) - 2(p(x|a) \cdot p(x|b)\right]/\left[1 - p(x|a) \cdot p(x|b)\right]$$

so that

$$p(x|a \vee b) \cdot \big[1 - p(x|a) \cdot p(x|b)\big] \leq p(x|a) + p(x|b) - 2 \cdot p(x|a) \cdot p(x|b)$$

and thus, distributing on the left,

$$p(x|a \vee b) - p(x|a \vee b) \cdot p(x|a) \cdot p(x|b) \leq p(x|a) + p(x|b) - 2 \cdot p(x|a) \cdot p(x|b).$$

Rearranging,

$$p(x|a \vee b) - p(x|a) \leq p(x|b) \cdot \big[p(x|a) \cdot \big\{p(x|a \vee b) - 2\big\} + 1\big],$$

giving the desired

$$p(x|b) \geq \big[p(x|a \vee b) - p(x|a)\big] / \big[p(x|a) \cdot \big\{p(x|a \vee b) - 2\big\} + 1\big].$$

When $p(x|a \vee b) > p(x|a)$, as must be the case if $a \vee b \mid\sim x$ while $a \not\mid\sim x$ in an application of DR pointed towards $b \mid\sim x$, then $p(x|b) \neq 0$. However, $p(x|b) \neq 0$ may come arbitrarily close to zero, as can easily be verified. Let $\varepsilon > 0$. We want to find p, a, b, x with $p(x|a \vee b) > p(x|a)$ and $p(x|b) < \varepsilon$. Take any positive integer n with $\varepsilon > 1/n$. Let $S = \{1, \ldots, n^2 + 1\}$ and put $p(i) = 1/(n^2 + 1)$ for all $i \in S$. Put $a = \{1, \ldots, n^2\}$, $b = \{n^2 - (n-1), \ldots, n^2 + 1\}$, $x = \{1, \ldots, n^2 - n\} \cup \{n^2 + 1\}$. Then $p(x|a \vee b) = (n^2 - n + 1)/(n^2 + 1) > (n^2 - n)/n^2 = p(x|a)$ and $p(x|b) = 1/n < \varepsilon$ as required.

Of course, this bound on pointed DR immediately supplies one for its Horn transform DC—we need only rewrite $p(x|a)$ as $1 - p(\neg x|a)$ to get $p(x|b) \geq [p(x|a \vee b) - (1 - p(\neg x|a))]/[(1 - p(\neg x|a)) \cdot \{p(x|a \vee b) - 2\} + 1]$.

6.7 Other Almost-Horn Rules

There are many other rules, both Horn and non-Horn, whose lossiness one may wish to determine, but they are less central than those that have been considered in this paper. We mention briefly just a few, omitting verifications.

One such Horn rule, cautious contraposition (CC), is like plain contraposition with a parameter a and an extra 'cautionary premise' $a \mid\sim \neg x$. It says: whenever $a \wedge b \mid\sim x$ and $a \mid\sim \neg x$ then $a \wedge \neg x \mid\sim \neg b$. Its pointed almost-Horn counterpart is obtained by replacing the cautionary premise $a \mid\sim \neg x$ by $a \not\mid\sim x$: whenever $a \wedge b \mid\sim x$ and $a \not\mid\sim x$, then $a \wedge \neg x \mid\sim \neg b$. Neither version is probabilistically sound. CC has the bound $p(\neg b|a \wedge \neg x) \geq [p(\neg x|a) + p(x|a \wedge b) - 1]/p(\neg x|a)$ with equality when $p(b|a) = 1$. This is the same as the improved bound for CM itself, with functional form $t \geq (r + s - 1)/r$ where r and s are the probabilities of the premises in the right order—which for CC means $r = p(\neg x|a)$ and $s = p(x|a \wedge b)$. This bound on CC (and thus also on its non-Horn counterpart) can be obtained either by direct calculation or by analyzing a derivation of CC in the system **O** plus a single application of CM.

Another rule of possible interest is cautious transitivity (CAT), which is like plain transitivity but with an added cautionary premise. It says: whenever $a \mid\sim b$, $b \mid\sim x$ and

$b \mid\sim a$ then $a \mid\sim x$. This too has its pointed almost-Horn counterpart, obtained by replacing the cautionary premise $b \mid\sim a$ by $b \not\mid\sim \neg a$. They have the bound $p(x|a) \geq p(b|a) \cdot p(x|b \wedge a) \geq p(b|a) \cdot [p(x|b) + p(a|b) - 1]/p(a|b)$, which again may be obtained either by direct calculation or by analyzing a derivation—this time a very short one in system O plus one application of each of CM and CT.

A word of warning about terminology should be made here. The almost-Horn counterparts of cautious contraposition and transitivity mentioned above, which were originally articulated in Hawthorne [8], are *not* the same as those called 'rational' contraposition and transitivity in Bezzazi et al. [4]. The latter are also pointed almost-Horn rules but with quite different negative premises. They, and other rules studied in [4], are in various respects stronger than rational monotony, but are of little interest from a probabilistic point of view. For example, the rule of determinacy preservation (DP) says, in its distributive form: whenever $a \mid\sim x$, then either $a \wedge b \mid\sim x$ or $a \wedge b \mid\sim \neg x$. This has a constant function as its trivial (and optimal) distributive bound: $\max\{p(x|a \wedge b), p(\neg x|a \wedge b)\} \geq 1/2$, with value 1 in the limiting case that $p(x|a) \in \{0, 1\}$. In its pointed form it says: whenever $a \mid\sim x$ and $a \wedge b \not\mid\sim x$ then $a \wedge b \mid\sim \neg x$, and has the even more trivial exact bound $p(\neg x|a \wedge b) = 1 - p(x|a \wedge b)$ except in the limiting case that $p(a \wedge b) = 0$ where LHS $= 1 =$ RHS.

6.8 Summary on Lossiness for Almost-Horn Rules

Although the distributed and pointed formulations of almost-Horn rules are equivalent as regards their probabilistic soundness, they give rise to different kinds of bounds function, which we call distributive and pointed. For the specific almost-Horn rules DR (disjunctive rationality) and RM (rational monotony) both distributed and pointed bounds may be established. The distributive bounds may be shown to be optimal for the rule, but the loss can be quite large—up to half. The pointed bounds admit the possibility of even larger loss; indeed, they can leave the conclusion with probability arbitrarily close to zero even when the probabilities of the premises are high.

Acknowledgements Thanks to Hanti Lin for kick-starting the investigation of optimality, and to Hans Rott for comments on a draft.

References

1. Adams, E.W.: Probability and the logic of conditionals. In: Hintikka, J.K.K., Suppes, P. (eds.) Aspects of Inductive Logic, pp. 265–316. North-Holland, Amsterdam (1966)
2. Adams, E.W.: Four probability-preserving properties of inferences. J. Philos. Log. **25**, 1–24 (1996)
3. Adams, E.W.: A Primer of Probability Logic. CSLI, Stanford (1998)
4. Bezzazi, H., Pino Pérez, R., Makinson, D.: Beyond rational monotony: on some strong non-Horn conditions for nonmonotonic inference operations. J. Log. Comput. **7**, 605–632 (1997)
5. Bourne, R., Parsons, S.: Propagating probabilities in system P. In: Proceedings of the 11th International FLAIRS Conference. AAAI Org, pp. 440–445 (1998)
6. Bourne, R., Parsons, S.: On proofs in system **P**. Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **8**, 203–233 (2003)

7. Gilio, A.: Probabilistic reasoning under coherence in system P. Ann. Math. Artif. Intell. **34**, 5–34 (2002)
8. Hawthorne, J.: On the logic of nonmonotonic conditionals and conditional probabilities. J. Philos. Log. **25**, 185–218 (1996)
9. Hawthorne, J.: New Horn rules for probabilistic consequence: is O^+ enough? In: Hansson, S.O. (ed.) David Makinson on Classical Methods for Non-classical Problems, pp. 157–166. Springer, Dordrecht (2014)
10. Hawthorne, J., Makinson, D.: The quantitative/qualitative watershed for rules of uncertain inference. Stud. Log. **86**, 249–299 (2007)
11. Kraus, S., Lehmann, D., Magidor, M.: Nonmonotonic reasoning, preferential models and cumulative logics. Artif. Intell. **44**, 167–207 (1990)
12. Lehmann, D., Magidor, M.: What does a conditional knowledge base entail? Artif. Intell. **55**, 1–60 (1992)
13. Makinson, D.: Bridges from Classical to Nonmonotonic Logic. College Publications, London (2005)
14. Makinson, D.: Logical questions behind the lottery and preface paradoxes: lossy rules for uncertain inference. Synthese **186**, 511–529 (2012)
15. Paris, J., Simmonds, R.: **O** is not enough. Rev. Symb. Log. **2**, 298–309 (2009)

D. Makinson
Department of Philosophy, Logic and Scientific Method, London School of Economics,
Houghton St, London WC2A 2AE, UK
e-mail: d.makinson@lse.ac.uk

D. Makinson (✉)
e-mail: david.makinson@gmail.com

J. Hawthorne
Department of Philosophy, University of Oklahoma, 605 Dale Hall Tower, Norman, OK 73019,
USA
e-mail: hawthorne@ou.edu

Nonsets

Daniel Parrochia

"Logic is not logic" (Béziau 2010).

Abstract Our topics are set theory and the problem of the definition of a set. With the existence of non-collectivizing relations and inconsistent multiplicities, the usual definition meets some limits. But, what is a nonset? Making a difference between sets and collections could appear as a subtle concern. In fact, this does not work very well. Indeed, philosophers show a lot of examples when naive sets and collectivizing relations fail, and modern mathematicians, from Cantor and Dedekind to Aczel suspect themselves that rising objections to the uncritical use of collectivizing relations is not unreasonable. A solution of the problem may be found in a rational model (logic but not Logic, as Beziau would say) that can formalize the notion of uncomparability between objects. An example is given with the ethical relation of "absolute alterity" developed by the French philosopher Emmanuel Levinas. Then, non-transitivity, trellis and weakly associative structures can be used to formalize such a situation.

Keywords Sets · Non-collectivizing relations · Uncomparability · Ethical relations · Pseudo-ordered sets · Trellis

Mathematics Subject Classification 03E99 · 03B80 · 06D99

1 Collectivizing Relations

In usual set theories, one often think of sets as displaying the following characteristics (among others):

1. No set is a member of itself;[1]
2. Sets (unlike properties) have their extensions essentially; hence no set can exist if one of its members has not;

[1]Of course, we know that some particular theories like, for instance, Finsler's set theory (see [4]), and the revival of it in Aczel (see [1]) or in Barwise and Etchemendy (see [2]), admit circular sets. Some others even assume the existence of a "universal set" (see [16]). But it is not the case in ZF.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_19

3. Sets form an iterated structure: at the first level, sets whose members are nonsets, at the second, sets whose members are nonsets or first level sets, etc.[2]

Cantor also inclined to think of sets as collections, i.e. things whose existence depends upon a certain sort of intellectual activity, a collecting or "thinking together". Here is Cantor's definition of sets in 1895 (see [6, p. 282]):

> By a "set" we understand any collection M into a whole of definite well-distinguished objects x of our intuition or our thought (which will be called the "elements" of M).

For doing so, Bourbaki [5, E, II, p. 3] suggests that there must exist a relation R such that the x in question are put together into the whole. This one is named a "collectivizing relation" and noted as follows: $Coll_x R$ says that "R is collectivizing in x".

2 The Problem of "Collectivizing"

One of the problems that can be raised about such a definition is that many objects or beings in the world can form a collection possibly without the result being understood as a whole. Very often, it is not the case.

Of course, a basic feature of reality is that there exist many things that we can collect. And when a multitude of given objects can be collected together, we arrive at a set.

For example, there are two tables in this room. We are ready to view them as given both separately and as a unity, and justify this by pointing to them or looking at them or thinking about them either one after the other or simultaneously.

Somehow the viewing of certain objects together suggests a loose link which seems to tie the objects together in our intuition. If the so-called objects are simple ones, there is no problem at all. If they are exactly the same, you can even add them.

But suppose these objects are very different, or very complex, or separated by some infinite gap (for instance, the gap existing between real numbers and natural numbers); or assume that the member of a set is the same that the set itself, so that the set is effectively a member of itself; or imagine again that not all the nonsets of a set can be collected together because some of them, for instance, still do not exist at the moment. What happens in all these cases? It seems obvious that you have got real nonsets, but none of them can be considered as elements of a set, whatever it can be.

3 Non-collectivizing Relations

Take the example of the Russell's set $E = \{x : x \notin x\}$. As Schoenfiels (see [27, p. 238]) writes, a closer examination of the paradox ($E \in E \iff E \notin E$) shows that it does not really contradict the intuitive notion of a set. According to this notion, a set E is formed by gathering together certain objects to form a single object, which is the set E. Thus before the set E is formed, we must have available all of the objects which are to be

[2]It is not only the case of the famous Russell's theory of types. ZF theory, as also Gödel's theory $V = L$ (see [19] or, easier to read [13]), carries on a hierarchical vision of the world.

members of E. But though these objects exist in a plain acception, in fact, there is no set such like E. Using the Bourbaki language, it means that the relation between all the objects of E is not a "collectivizing relation".

Another example of a non-collectivizing relation is what Cantor called in 1899, in a letter to Dedekind, "absolutely infinite or inconsistent multiplicities":

> A (definite) multiplicity is a system or a totality of things. But a multiplicity may be such that the assumption that all of its elements "are together" leads to a contradiction, so that it is impossible to conceive of the multiplicity as a unity, as a "one finished thing" [29, pp. 113–117].

One example, Cantor said, is the "totality of everything thinkable". Other ones are the system Ω of all the ordinals and the system T of all the alephs.

Nelson's non-standard analysis gave recently another example of an inconsistent multiplicity. In that theory, some objects or sets are said to be internal or external, according to the fact that they can be formalized or not within the Zermelo–Fraenkel axiomatics (ZF). A new predicate P, meaning "to be standard", is then introduced, which does not stand in ZF. Then an internal or external formula f is defined, so that the set $F = \{x : f(x)\}$ will be itself either an external or internal set. In the first case, one cannot apply the rules of ZF axiomatics. F is not a true set and f is not a collectivizing relation. For instance, the set of all non-standard integers is not a usual set of ZF. It is an external set. Of course, non-standard analysis introduces a "standardization axiom" so that, nevertheless, we may get some access to a substitute of this set. But we cannot work with the set itself.

4 Some Well-Known Theorems

We give here some examples of well-known theorems about non-collectivizing relations and their consequences.

Theorem 4.1 (Bourbaki) *Non* $\mathrm{Coll}_x(x \notin x)$.

Proof Assume $x \notin x$ is a collectivizing relation. Then there must exist a set E so that E is the set of all the x such that $x \notin x$. But now, a question: Do we have $E \in E$, or not? If $E \in E$, E is one of the x such that $x \notin x$. So $E \in E$. But if $E \in E$, E is not one of the x such that $x \notin x$. So E is not in E, and we have $E \notin E$. So we get $E \in E \iff E \notin E$, a contradiction. □

Theorem 4.2 (Cantor) $(\nexists X)(\forall x)(x \in X)$.

Proof The theorem says that there does not exist a set X whose all objects are elements. If it were the case, all relations would be collectivizing ones. But we have seen (Theorem 4.1) that $x \notin x$ is not collectivizing. □

Theorem 4.3 (Cantor) *Let* $\Omega = \omega_0, \omega_1, \ldots, \omega_n, \omega_{n+1}, \ldots, \omega_{\omega_0}, \ldots$, *the sequence of infinite ordinals. Ω is not a set, so we have*: non $\mathrm{Coll}_{\omega_\alpha}(\omega_\alpha \in \Omega)$.

Theorem 4.4 (Cantor) *Let* $T = \aleph_0, \aleph_1, \ldots, \aleph_n, \aleph_{n+1}, \ldots, \aleph_{\omega_0}, \ldots$, *the sequence of infinite cardinals. T is not a set, so we have*: non $\mathrm{Coll}_{\aleph_\alpha}(\aleph_\alpha \in T)$.

The problem of inconsistent multiplicities is that they are, in fact, contradictory entities. So we cannot use them to prove anything, as Cantor quickly realized. In fact, we cannot use them at all.

Nevertheless, a French philosopher Quentin Meillassoux [23] tried to make use of them in philosophy of science: Thinking that experience is pure contingency, he assumes that it is an inconsistent multiplicity. As a consequence, we cannot apply there the concepts of Probability Calculus because experience—or, in other words, the multiplicity of facts—*is not a set*! So we cannot assure the existence of invariants or prove the stability of scientific laws by some inductive reasoning (that probability, otherwise, could have justified).

But solving the epistemological problem of scientific laws at this cost is too much expansive. And there is no convincing evidence that experience, even if it is taken through the course of time, might be identified with a so problematic entity.

5 Collectivities, Sets, Collections

As far as I know, most of "working mathematicians" would agree with me that taking non-collectivizing relations as a basis for a theory is not a solution to set-theoretic problems.

A lot of them, in front of the question raised above, will rather make some difference between *collectivities* and *sets*. In this case, they will give to collectivities, not to sets, the best part, that is, the property of being built from true relations (see [22]).

Indeed, we can observe collectivities in the not living world (universe galaxies, solar systems, crystalline units) as in the living world (ant hills, bee swarms, nations). What properties are behind the relations who tie the collectivities? Maybe is the gravity, the symmetry or the survival instinct? In a word, it is the structural self-organization. The self-organization can be structural and functional. The paper [22] refers to the structural self-organization applied to the interconnected collectivities.

First, let us define the collectivity. Therefore, we must answer another question: What is a set? A set can be selected by a membership, or by a relation which substantiates the membership, or by bringing in the set field elements which fulfill the relation.

Because Bourbaki names "collectivizing relation" the relation defining a set, the authors name "collectivities" only the sets selected or built up by the help of the relations. Therefore, they exclude the sets selected by the membership, the most general. A collectivity means not a set made, for example, of a star, a planet, a crystal, an ant, a bee and a man. The relation that substantiates the membership of a collectivity is connected with its structure: a collectivity is made of the least structural entities, e.g. an interconnection means nodes and links, equivalent to the graph definition.

As it is said, an assemble grouping very different things such a star, a planet, a crystal, an ant, a bee and a man, etc. can still be viewed as a "set", but it is certainly not a "collectivity" because no clear relations exist between these elements. Then, to be the member of a set is not sufficient to make a collectivity. So, for such authors, it is clear that a set is not a collectivity, even if forming a set supposes the existence of a "collectivizing" relation R in the sense of Bourbaki.

So we must strongly differentiate *collectivities* and *collections*. A set is not necessarily a collectivity because the whole that their members belong to has not necessarily a well-define structure. But a set is an actual collection, that is, a collection of objects that we can enumerate, if it is finite, and to which corresponds some aleph, if it is infinite.

Now a collection cannot be necessarily a set, and the problem is that we do not know, for the moment, how to study collections that are not sets, i.e. those entities that we may call, using a negative and very imprecise term, "nonsets".

Making a difference between sets and collections could appear as a subtle concern. But we must remember that Cantor's (or Dedekind's) view was that a set is in fact a "system"—in the sense of Kant, i.e. a platonistic idea:

> Par un "ensemble" ou "système", j'entends en effet de façon générale toute multiplicité qui peut être pensée comme une unité, c'est-à-dire toute collection d'éléments déterminés qui peut être par une loi combinée en un tout : je crois définir ainsi quelque chose d'apparenté à l'"image" ou l'"idée" platonicienne.

6 Philosophy

In the past, for the German philosopher I. Kant, some problems already raised with the use of concepts like God, World or Soul (see [20]), which mean strictly nothing from a cognitive point of view. The main argument of Kant was that we have no intuitions (in particular, no sense data) of these objects (which are not, moreover, space–time entities). Besides, as knowledge was, for him, a kind of "synthesis" between concepts and intuitions, we cannot get positive ideas about such things. God and Soul are made of an invisible stuff and so, stay beyond our perception. But even the World itself, in what we are, cannot be viewed as a whole: we always see *a part of* the world, but not *the whole* world because, for seeing such a big entity, we should have to get out of it and move, for instance, in the direction of some extra-world, from which it should be possible to see this one as a whole. So our World cannot be identified with the "set of all things". In the same way, our soul cannot be the "set of all the ideas of our mind". As Dedekind had already shown—and Cantor says nothing different—the system of all the ideas is an inconsistent multiplicity.[3]

In modern philosophy, God himself may be seen as the "Absolute Other", that is, not only a transcendent point that could be located at a finite distance in a geometric space and that we could reach by a combination of finite algebra operations with homogeneous coordinates, as we can see in projective geometry over finite fields (see [15, 18]), but an absolutely infinite entity which does not enter any set operations. If the relation to God becomes a model for the relations between men, as in Levinas' philosophy, then, even two men cannot be put together into a set (see [21, p. 9]). So, the axiom of pairing does not work in affective or social relations, that always are relations between "absolute others". And no more the axiom of arithmetic, which precisely allows us to make collections. Of course, adding certain elements to others does obey to some rule fixing the extent of that possibility: we only add entities of the same stuff. But the question is: Is there, in the world, any entity of the same stuff? Of course, there is no problem if you add balls, books, cats, etc. But can you add two human beings, especially when they have complex relationships? Everybody may understand that a couple is not a pair, even when its members have the same sex.

---

[3] A consequence is that the notion of an "infinite mode" (of thought), in Spinoza's Ethics, is certainly a contradiction.

Table 1 Profile of merit

Wars	2
Legs	1
Wives	2
Children	4
Wounds	2
Total	11

In front of these problems, some people feel inclined to get an extension of the addition law. Philip J. Davis and Reuben Hersh [11, p. 71] recall the story of a beggar in Time Square, who pined on himself the following information, proving that the addition rule is not so well known—or perhaps must we say it might be transgressed if we decided to define a profile of merit based on some new computation of troubles and offenses (see Table 1).

For sure, this apologue will make every mathematician laugh. And I have been laughing as well for a long time.

But, one day, I came across this famous letter from Cantor to Dedekind (Hahnenklee 1899, August 28th) [7, p. 447–448] where Cantor asks about the consistency of finite and infinite well-orderings. And suddenly, I stop laughing because Cantor wrote the worrying following lines, where he said that even finite multiplicities cannot be proved to be consistent. The fact of their consistency is a simple unprovable truth—"the axiom of arithmetic (in the old sense)"; the fact of the consistency of infinite multiplicities, i.e. multiplicities that have an aleph as their cardinal number, is in exactly the same way an axiom, "the axiom of the extended transfinite arithmetic":

> One must throw up the question, where from I know that the well ordered multiplicities or sequences to which I ascribe the cardinal numbers,
>
> $$\aleph_0, \aleph_1, \ldots, \aleph_{\omega_0}, \ldots, \aleph_{\omega-1}, \ldots,$$
>
> are really "sets" in the sense of the word explained above, i.e. "consistent multiplicities". Could not we conceive that already *those* multiplicities be "inconsistent" ones, but that the contradiction which there is to suppose a "simultaneous existence (Zusammensein) of all their elements" does not get noticed yet? My answer is that this question *can be also widened to the finite multiplicities*, and that, if we think about it exactly, we end then in the result: even for finite multiplicities, *there is no* proof of their "consistency". In other words: the fact of the "consistency" of the finite multiplicities is a simple, unprovable truth, it is the "*axiom* of the arithmetic" (in the old sense of the word). And, also, the "consistency" of the multiplicities to which I attribute the alephs as cardinal numbers, is "the axiom of the transfinite widened arithmetic".

Suppose now a person who is unreasonable enough[4] for refusing, at less in some cases, the very pragmatic (though completely unprovable) "axiom of arithmetic (in the old sense)". That is exactly the position of the French philosopher Emmanuel Levinas. As Levinas says, one cannot add the other to me, there is no number to be associated with such a relation, which is a non-symmetric relation, and, in fact, an *ethical* one.

[4]If you wonder there can exist some, you may think of an imaginary animal like the tortoise of Lewis Carroll (see [8]), which never admitted a quite simple deduction relation before it has been set up into a group of axioms.

7 The Myth of "Naive Sets"

In front of these facts, we are led to a tricky question: Are collectivizing relations an error? But where do they come from? Indeed, raising such a problem is asking in fact about the origin of logic and rational thought.

For what reason, at a particular date in the history of man, some members of our species began to need gathering things and putting them into the same whole cluster, constituting what we call now "naive sets"?

The least we can say is that there is no evidence that the first grouping or classification[5] that has been constructed by men all around the World is particularly clear and undisputable. For example, Nietzsche thought it was the result of some handicap or infirmity:

> *Origin of the logical.*—How did logic come into existence in man's head? Certainly out of illogic, whose realm originally must have been immense. Innumerable beings who made inferences in a way different from ours perished: for all that, their ways might have been truer! Those, for example, who did not know how to find often enough what is "equal" as regards both nourishment and hostile animals, who subsumed things too slowly and cautiously, were favored with a lesser probability of survival than those who guessed immediately upon encountering similar instances that they must be equal. The dominant tendency, however, to treat as *equal* what is merely *similar*, an illogical tendency—for nothing is really equal—is what first created any basis for logic. In order that the concept of substance could originate—which is indispensable for logic although in the strictest sense nothing real corresponds to it—it was likewise necessary that for a long time one did not see nor perceive the changes in things; the beings that did not see so precisely had an advantage over those that saw everything "in flux" [24, § 111].

As Nietzsche says, men who did not see very well, and so, could not distinguish small details and changes in their perceptions of objects or other entities of the world, were led to identify most of them and came to consider that they belong to the same invariant class, the same *species*. Here is the birth of early classifications, but also the beginning of logic, because one of the most famous philosophical class invariant is the concept of "substance", which is also one of the basic concept of the Aristotelian ontology (to which corresponds the concept of "subject" in logic). So, with an uncomparable sense of humor, Nietzsche tells us that logic and classifications could have been based on a kind of shortsightedness, which led to identify—in modern words—*equivalence relations* (reflexive, symmetric and transitive) with similarity relations (only reflexive and symmetric, but not transitive). But it means, in fact, that the main categories of our thought could be based on a prime error, and that putting together things of our environment is not a trivial operation. In fact, there are no "naive" sets.

8 The Modern View

Let us return now to Cantor's letter to Dedekind (Hahnenklee 1899, August 28th).

In his famous book about the cantorian construction of set theory, the French philosopher Jean Cavaillès commented the quotation of Cantor concerning the lack of a true evidence for consistent and inconsistent multiplicities, and the existence of implicit (and

[5]For an overview on classifications problems, see [26].

quite unprecise) axioms, respectively named the “axiom of arithmetic” and the axiom of transfinite arithmetic, in the following way:

> To these axioms “in the old sense” whose no evidence is a guarantee, the successors of Cantor will try, for maintaining the essentials of the abstract theory, to substitute, as far as it is possible, the consistency of a modern axiomatics. As a beginning for their endeavors, were naturally introduced the systematic construction and the methods, by which the algebraic spirit of Dedekind had initiated a fragment of set theory that is the less intuitive possible [9, p. 118].

In particular, Von Neumann [30] admitted that a nonset could be an ordinal number. But then, all its segments must be true sets, since they stand for an argument in the function by which is defined a correspondence between the segment and the element which determines it. The result is that Ω, the system of all the ordinals, is the only one ordinal number which is not a set.

9 Problems of Non-collectivizing

The main problem of set theorists was to avoid paradoxes. Different strategies have been set up. Russell’s theory of types, Zermelo–Fraenkel axiomatics (with the Schema of Separation), and finally, Von Neumann’s theory (with the axiom of foundation). But there is a cost to pay. Russell’s theory of types yields a hierarchical universe with a lot of internal problems; ZF-axiomatics leads to an artificial difference between sets and proper classes (classes that are not sets); and Von Neuman’s theory assumes an axiom which is reputed completely unnecessary (the axiom of foundation). When replaced by its negation (the anti-foundation axiom), it makes possible a new theory, the theory of circular sets, whose origin, as we have seen, lays in Finsler’s theory of sets of 1925 (see [4]), and its later developments in Aczel (see [1]), or in Barwise and Etchemendy (see [2]).

So raising objections to the uncritical use of collectivizing relations is not, from the viewpoint of pure theory, so unreasonable.

Having said this, we must, however, observe that we can hardly get rid of collectivizing relations and save a reasonable mathematical theory.

The main fact is that when we speak of more than one thing at a time, we implicitly consider the things we talk about as a consistent multiplicity. And if we take apart the questions of the Russell’s set, of the sets of all ordinals or cardinals, and of the sets of all sets, we are generally logically right.

So, if we want to suggest that the things we put together in the same class are absolutely different from one another (so that they make a collection but not necessarily a collectivity), it will probably be better to indicate that in a definition.

In this case, defining a particular relation on such a set, may reveal some hidden structure, to which we can associate, further on, some kind of algebra, or order, or topology, so that we can continue to investigate the properties or things in the collection with the help of reason.

If not, I am afraid, we could not do any mathematics at all.

So, as a solution to the problem raised by the Levinas’ counter-examples to set theory, I allow myself to introduce the following rational model.

10 A Rational Model

Let $\parallel$ be a relation of *absolute alterity*. We say that $a \parallel b$ (that we read as "a is absolutely different from b", if a and b cannot be connected in any way. This kind of relation is very closed to the uncomparability relation, where $a \parallel b$ is interpreted as "a and b are uncomparable elements" (which means that we have neither $a \leq b$ nor $b \leq a$).

Let us also introduce a new relation $\otimes$ which will be named a *hostage relation*. This relation $a \otimes b$ means that "a is the hostage of b" or that "b has taken a hostage".

10.1 Properties of $\parallel$ *and* $\otimes$

Let us first explain the properties of $\parallel$:

- $\parallel$ is irreflexive (we do not have $a \parallel a$ because a is not absolutely different from itself, except in the case of the "simulacrum"[6]);
- $\parallel$ is symmetric ($a \parallel b$ iff $b \parallel a$, for all $a \neq b$);
- $\parallel$ is not transitive (if $a \parallel b$ and $b \parallel c$, then we do not have in general $a \parallel c$, even for all $a \neq c$). This means that $\parallel$ is not an equivalence relation (because of irreflexivity) or an order or a preorder relation (because of non-transitivity).

Assume now we have, for $\otimes$, the following properties:

- $a \otimes a$ (reflexivity, a may be a hostage of himself).
- $\neg(a \otimes b \Rightarrow b \otimes a)$ if $a \neq b$. (If $a \neq b$, the fact that a is the hostage of b does not yield that b is the hostage of a; from the viewpoint of a, it is not required. It is the problem of b.)[7]

From this follows that the relation is not symmetric (as $a \otimes b$ does not imply $b \otimes a$, *a fortiori*, $a \otimes b$ is not equivalent to $b \otimes a$).

But $\otimes$ is not transitive. The fact that $a \otimes b$ and $b \otimes c$ does not necessarily mean that $a \otimes c$.

So $\otimes$ is a reflexive, antisymmetric and non-transitive relation. As we have no transitivity, it is not an order or a preorder relation.

We shall define now a third relation $\trianglelefteq$, that we shall name an "ethical relation" and that we define as the logical sum of the previous ones. Let us set:

$$a \trianglelefteq b =_{df} a \parallel b \oplus a \otimes b.$$

[6]This concept, coming from Plato's images (eikones), is emphasized in Deleuze (see [12]). But we can already find an example of it in Diderot (see [14]) where, for example, the person of Rameau's nephew, constantly differs from itself (see our comments in [25]).

[7]Cf. [21, p. 94]. The following is a part of a dialog between Philippe Nemo (PH. N.) and Emmanuel Levinas (E.L.): "PH. N.—But do not others be, also, responsible for me? E.L.—Maybe, but this is *their* business. One of the fundamental themes (...) of *Totalité et Infini* is that the intersubjective relation is not a symmetric relation. This way, I am responsible for others without waiting for the reverse, had this to cost me the life. It is exactly because, between the others and me, the relation is not mutual that I am subjected to the others; and I am a "subject" essentially this way."

10.2 Properties of Relation $\trianglelefteq$

- (Reflexivity) $a \trianglelefteq a =_{df} a \parallel a \oplus a \otimes a = a \otimes a$. But we do not in general[8] have $a \parallel a$. So $a \trianglelefteq a$ almost always means $a \otimes a$.
- As we may have $a \otimes b$ without necessarily having $b \otimes a$, symmetry is not verified. But we can say that if x has an ethical relation with y and if y has an ethical relation with x, then x and y are "equals" in the democratic sense. So, we have a kind of antisymmetry.
- In general, transitivity is not verified.

Proof $a \trianglelefteq b =_{df} a \parallel b \oplus a \otimes b$. Then $b \trianglelefteq c =_{df} b \parallel c \oplus b \otimes c$. But even though $a \parallel b$ and $b \parallel c$ do not yield $a \parallel c$ (non-transitivity of $\parallel$), even for $a \neq c$, as well, $a \otimes b$ and $b \otimes c$ do not imply $a \otimes c$. □

- Some elements of the set may be incomparable.

11 Trellis

We have seen that the "ethical relation" is not a transitive relation. In the 1970s, Helen Skala [28] has developed some attempt to take into account such situations. The material presented in her mathematical memoir contains a foundation for the theory of non-transitive orderings. Let us recall here first some definitions.

Definition 11.1 (Pseudo-ordered sets) Any reflexive and antisymmetric binary relation $\trianglelefteq$ on a set A will be called a *pseudo-order* on A, that is, $x \trianglelefteq x$ for any element x of A, and if $x \trianglelefteq y$ and $y \trianglelefteq x$ then $x = y$.

As Skala said, a natural example of a pseudo-order on the set of real numbers is obtained by setting $x \trianglelefteq y$ iff $0 \leq y - x \leq a$, where a is a positive number.

Another example, connected with classification problems, is the following one. Let X be a set and χ, a family of subsets of X. For any subsets R and S of X, set $R \trianglelefteq S$ iff $R \subseteq S$ and the set-theoretic difference $S - R$ belongs to χ. In this way, 2^X is pseudo-ordered.

Now, the ethical relation between persons, according to Emmanuel Levinas, may be also considered as a pseudo-order.

Let us now introduce the following theorems and definitions:

Theorem 11.1 *Each pseudo-ordered set is isomorphic with a contraction set of a partially ordered set.*

Definition 11.2 (Pseudo-chain, cycle) For each subset B of a pseudo-ordered set A, a transitive and reflexive, but not necessarily antisymmetric relation $\trianglelefteq_B$ can be defined on B by setting $b \trianglelefteq_B b'$ iff there exists a finite sequence $(b_1, \ldots, b_n)$ of elements from B such that $b \trianglelefteq b_1 \trianglelefteq \cdots \trianglelefteq b_n \trianglelefteq b'$. If for each pair of elements b and b' at least one of the relations $b \trianglelefteq b'$ or $b' \trianglelefteq b$ holds, then B will be called a *pseudo-chain*. If both these relations hold for each pair of elements, B is said to be a *cycle*.

[8]Except in the case of the simulacrum.

Fig. 1 A pseudo-ordered set

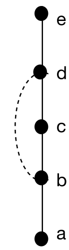


Table 2 Trellis

∧ \ ∨	a	b	c	d	e
a		b	c	d	e
b	a		e	e	e
c	a	b		d	e
d	a	a	a		e
e	a	b	c	d	

Theorem 11.2 *The elements of a finite pseudo-chain can be arranged (with possible repetitions) in a sequence* $(b_1, \ldots, b_n)$ *such that* $b_1 \trianglelefteq \cdots \trianglelefteq b_n$.

Definition 11.3 A pseudo-order on a set A is said to be *linear* if A itself is a pseudo-chain.

It can easily be shown that any pseudo-order can be extended to a linear one, but there is a sharper result:

Theorem 11.3 *Any pseudo-order* $\trianglelefteq$ *on a set* A *can be extended to a linear pseudo-order* $\trianglelefteq'$ *on* A *such that any cycle with respect to* $\trianglelefteq'$ *is also a cycle with respect to* $\trianglelefteq$.

We have also:

Theorem 11.4 *Any pseudo-order is the intersection of all its linear extension.*

Usually, finite pseudo-ordered sets are represented by Hasse-type diagrams with the convention that if x is below y and connected to y without $x \trianglelefteq y$ holding, then x and y will be joined by a dashed curve. For example, the pseudo-ordered set $\{a, b, c, d, e\}$ where $a \trianglelefteq b \trianglelefteq c \trianglelefteq d \trianglelefteq e$, $a \trianglelefteq x \trianglelefteq e$ for each x, would be represented as indicated in Fig. 1.

Definition 11.4 (Trellis) By a trellis, we mean a pseudo-ordered set, any two of whose elements have a least upper bound (l.u.b.) and a greatest lower bound (g.l.b.).

The pseudo-ordered set of Fig. 1, for example, is a trellis. Its l.u.b. and g.l.b. are given in Table 2. (It is understood that the operations are commutative and idempotent).

"Ethical relations" are of this kind. No doubt that they present a lot of interesting properties that have to be carefully explored.

To conclude (and also extend) our main view, we can say that the material presented above as a generalization of the concepts of partial order and lattices, offers many advantages.

1. By starting out with a reflexive and antisymmetric but not necessarily transitive order, one can define the least upper bound and greatest lower bound similarly as for partially ordered sets, thus obtaining the structure called a "trellis", in which these operations are not necessarily associative (see [28]). But with this approach, one can prove nearly all the basic theorems of lattice theory, thus revealing the superfluity of the assumption of associativity, which seems, philosophically, compatible with the possibility of ethical relations. So Ethics (at least, ethics in the sense of Levinas) seems to be a pseudo-ordered domain in which, consequently, nearly the same theorems may hold.
2. Trellis or weakly associated structures can also formalize the concept of a "tournament", i.e. a structure $(T, <)$, where T with a binary relation $<$ is such that for all $a, b \in T$ exactly one of $a = b$, $a < b$, and $b < a$ holds. This structure is nearly the one we got when we defined an $\otimes$ relation on a set, if we admit that $\otimes$ may be extended to signify "Exactly one of the following relations holds: a is the hostage of b or b is the hostage of a, or a is the same as b". It is also an algebra $(T, \vee, \wedge)$ defined by the rule: if $x < y$ then $x = x \wedge y = y \wedge x$ and $y = x \vee y = y \vee x$, and $x = x \wedge x = x \vee x$ for all x. In this algebra $(T, \vee, \wedge)$, neither $\wedge$ nor $\vee$ is associative unless $(T, <)$ is a chain, that is, $<$ is transitive. However, the two operations are idempotent, commutative; the absorption identities hold, and a weak form of the associative identities holds (see [17]).
3. On weakly associative lattices, we can also define some *tolerance relations* (that may be derived as well on lattices). Remember that a relation R on a set E is said to be a *tolerance* if it is reflexive and symmetric (see [10]). So, note that tolerance is exactly the relation $||$ when we admit it is reflexive, that is, in the case mentioned above of the "simulacrum". In France, tolerance relations are rather named "similarity relations". This kind of relations gives covers instead partitions, these being connected with equivalence relations, which suppose transitivity to be verified. A lot of interesting theorems may be derived from that.

References

1. Aczel, P.: Non Well Founded Set Theory. CSLI Lecture Notes, vol. 14. CSLI, Stanford (1988)
2. Barwise, J., Etchemendy, J.: The Liar: An Essay on Truth and Circularity. Oxford University Press, Oxford (1987)
3. Béziau, J.-Y.: Logic is not logic. Abstracta **6**, 73–102 (2010)
4. Booth, D., Ziegler, R.: Finsler Set Theory: Platonism and Circularity. Birkhaüser, Basel (1996). Translation of Paul Finsler's Papers on set theory with introductory comments
5. Bourbaki, N.: Théorie des Ensembles. Hermann, Paris (1970)
6. Cantor, G.: Gesammelte Abhandlungen mathematischen und philosophischen. Springer, Berlin (1932). Ed. Ernst Zermelo
7. Cantor, G.: Abhandlungen mathematischen und philosophischen Inhalts. Olms, Hildesheim (1966). Reprint of [5].
8. Carroll, L. (Dodgson, C.): The Game of Logic. MacMillan, London (1886)
9. Cavaillès, J.: Philosophie Mathématique. Hermann, Paris (1962)
10. Chajda, I., Bohdan, Z.: Weakly associative lattices and tolerance relations. Czechoslov. Math. J. **26**(2), 259–269 (1976)

11. Davis, Ph., Hersh, R.: L'Univers Mathématique. Bordas, Paris (1985). A translation of: The Mathematical Experience, Birkhäuser, Boston, 1982
12. Deleuze, G.: Logique du Sens. Minuit, Paris (1969)
13. Devlin, K.: The Joy of Sets. Fundamental of Contemporary Set Theory, 2nd edn. New York, Berlin (1990)
14. Diderot, D.: Le neveu de Rameau. In: D. Diderot, Oeuvres. Gallimard, Paris (1979)
15. Dubarle, D.: Logique et Dialectique. Larousse, Paris (1970)
16. Forster, T.: Set Theory with a Universal Set, Exploring an Untyped Universe, 2nd edn. Clarendon Press, Oxford (1995)
17. Grätzer, G., Kisielewicz, A., Wolk, B.: An equational basis in four variables for the three-element tournament. Colloq. Math. **63**(1), 41–44 (1992)
18. Hirschfield, J.W.P.: Projective Geometries over Finite Fields. Clarendon Press, Oxford (1979)
19. Jech, Th.: Set Theory. Springer, Berlin (2000)
20. Kant, I.: Critique of Pure Reason. Cambridge University Press, Cambridge (1999). New translation
21. Levinas, E.: In: Totalité et Infini (1971). M. Nijhoff, La Haye
22. Lupu, C., Niculiu, T.: Locality and symmetry in interconnecting. Sci. Bull. - "Politeh." Univ. Buchar., Ser. C Electr. Eng. **71**(4), 15–26 (2009)
23. Meillassoux, Q.: L'inexistence divine, essai sur le dieu virtuel. Thèse de Philosophie, Université Paris **1** (1997)
24. Nietzsche, F.: The Gay Science. Vintage Books, New York (1974), translation by W. Kaufmann
25. Parrochia, D.: Philosophie et Musique Contemporaine, le Nouvel Esprit Musical. Champ Vallon, Seyssel (2006).
26. Parrochia, D., Neuville, P.: Towards a General Theory of Classifications. Birkhaüser, Basel (2013)
27. Schoenfiels, J.R.: Mathematical Logic, 2nd revised edn. A K Peters, Natick (2001)
28. Skala, H.: Trellis Theory. AMS, Providence (1972)
29. Van Heijenoort, J.: From Frege to Gödel. Harvard University Press, Cambridge (1967)
30. Von Neumann, J.: Die Axiomatisierung der Mengenlehre. Math. Z. **27**, 669–752 (1928)

D. Parrochia (✉)
University Jean Moulin, Lyon, France
e-mail: daniel.parrochia@wanadoo.fr

A Roadmap to Decidability

João Rasga, Cristina Sernadas, and Amílcar Sernadas

Abstract It is well known that quantifier elimination plays a relevant role in proving decidability of theories. Herein the objective is to provide a toolbox that makes it easier to establish quantifier elimination in a semantic way, capitalizing on the fact that a 1-model-complete theory with algebraically prime models has quantifier elimination. Iteration and adjunction are identified as important constructions that can be very helpful, by themselves or composed, in proving that a theory has algebraically prime models. Some guidelines are also discussed towards showing that a theory is 1-model-complete. Illustrations are provided for the theories of the natural numbers with successor, term algebras (having stacks as a particular case) and algebraically closed fields.

Keywords Quantifier elimination · Decidability of theories

Mathematics Subject Classification (2000) 03B25 · 03C10

1 Introduction

Quantifier elimination is a key property for proving decidability of a first-order theory. Decidable theories play an essential role in computer science applications. For example, the theory of real closed fields (firstly proved to be decidable by Alfred Tarski, see [14]) is of utmost interest when considering, for instance, probabilistic reasoning (see [4]) and hybrid system verification (see [12]), and the theory of algebraically closed fields has an important role in computer algebra systems (see [2]). On the other hand, decidable theories are very relevant in the areas of theorem proving and data abstraction [1, 5, 13, 15].

Proving decidability of new theories as well as finding new algorithms for proving quantifier elimination is still a very active research concern (see, for instance, [10, 11]). Moreover, it is foreseeable that future applications of computer science may require proving decidability of other relevant theories. Thus we need an effective toolbox of techniques for helping computer scientists in this task.

The role of quantifier elimination for proving decidability is depicted in Fig. 1. From there we can conclude that if we manage to prove that a theory is axiomatizable, has quantifier elimination and has a prime model then we can conclude that it is decidable.

Quantifier elimination has been studied from a symbolic perspective (constructive in the sense that an algorithm is given to compute a quantifier-free formula equivalent to a given formula) as well as from a semantic point of view. In applications, people want to investigate symbolic quantifier elimination techniques since they are more useful in

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_20

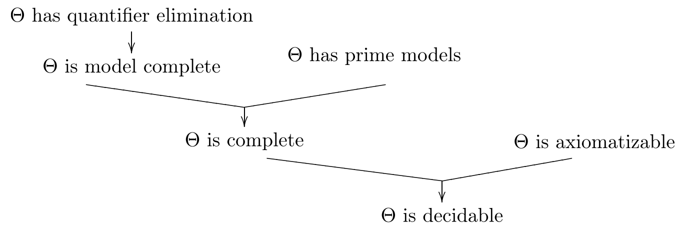


Fig. 1 A roadmap to decidability

algorithms. However, we believe that the semantic perspective is also very important since it may allow to prove in a simpler way that a theory enjoys quantifier elimination. If so, then the effort can be concentrated on the constructive algorithm but already with the knowledge that the theory enjoys quantifier elimination. And, of course, if this is not the case no investment on the constructive part is made. In this paper. we concentrate our attention on general semantic techniques for quantifier elimination.

A well known sufficient condition for a theory Θ to have quantifier elimination (see, for instance, Corollary 3.1.12 of [9]) requires that Θ has algebraically prime models and is 1-model-complete. However, looking at several well known theories, we get the impression that more help could be provided, based on that sufficient condition, for proving quantifier elimination. This is even more important when there are more researchers who want to use (semantic) quantifier elimination.

The objective of this paper is to provide workable sufficient conditions for proving that a theory has algebraically prime models and is 1-model-complete.

We start by generalizing the notion of a theory having *algebraically prime models*, to a theory having *algebraically prime models with respect to* another theory. As we will see in Sect. 3, the extended notion coincides with the original one in some particular cases.

We identify two ways of establishing that a first-order theory Δ has algebraically prime models with respect to a theory Υ:

- (Iteration) Δ has algebraically prime models with respect to Υ whenever there is a map between the models of Δ satisfying particular conditions that guarantee that the image, by this map, of a model of Δ is "closer" to be a model of Υ than the original model and that the successive application of this map converges eventually in the limit to a model of Υ which is an algebraic prime extension of the original model;
- (Adjunction) Δ has algebraically prime models with respect to Υ whenever there is a particular adjunction between their categories of models and embeddings.

The first condition is developed in Sect. 3 and the second in Sect. 4. By composing these two techniques (see the ending of Sect. 4), we have a more practical way to prove that a theory has algebraically prime models. Furthermore, we propose in Sect. 5 a sufficient condition, named adequacy for $\exists$, for a theory to be 1-model-complete.

We illustrate these conditions and techniques throughout the paper on the theory of naturals with successor, the theory of term algebras (having a theory of stacks as a particular case) and on the theory of algebraically closed fields.

Finally, in Sect. 6, we briefly draw some concluding remarks.

2 Preliminaries

We start this section by briefly reviewing some relevant notions. Given a first-order signature, by a *(first-order) theory* over that signature we mean a set of sentences over the signature, closed under semantic entailment.[1] In the sequel, in order to simplify the presentation, we may omit the reference to the signature when there is no ambiguity. A theory Γ is *decidable* if there is an algorithm that when receiving a formula returns 1 if the formula is in Γ and 0 otherwise. Usually the decidability of a theory is not proved directly. It is common to prove that the theory is complete and axiomatizable, since these conditions imply the decidability of the theory. A theory Γ over a signature Σ is *complete* if, for every closed formula φ, either

$$\Gamma \vDash_{\Sigma} \varphi \quad \text{or} \quad \Gamma \vDash_{\Sigma} \neg\varphi$$

and is *axiomatizable* if there is a decidable set of sentences Θ such that $\Theta^{\vDash_{\Sigma}} = \Gamma$. For instance, Th($\mathbb{N}$), the theory of natural numbers with function symbols 0, successor, $+$ and $\times$ and predicate symbol $<$ is not axiomatizable as shown by Gödel in [6]. Herein we only consider axiomatizable theories, and, so, we identify a theory with a decidable set of sentences (the *axioms* of the theory) whose closure under semantic entailment coincides with the theory.

A theory Θ over a signature Σ has *quantifier elimination* if, for each formula φ, there is a quantifier-free formula φ^* such that $\Theta \vDash_{\Sigma} \varphi \Leftrightarrow \varphi^*$, and φ and φ^* have the same free variables. As examples of theories enjoying quantifier elimination note that the first-order theory of the:

- Natural numbers with successor;
- Natural numbers with successor, $+$, $-$, $<$;
- Divisible torsion-free Abelian groups;
- Term algebras;
- Atomless Boolean algebras;
- Algebraically closed fields;
- Ordered real closed fields;
- Differentially closed fields;
- Dense orders without limits

enjoy quantifier elimination. Amongst the results establishing sufficient conditions for quantifier elimination, there is one that is widely used when proving this property by symbolic and constructive techniques. It says that a theory Θ has quantifier elimination providing that for each formula

$$\exists x\, \varphi$$

where φ is a quantifier-free formula, there is a formula φ^* such that: (i) φ is a quantifier-free formula; (ii) φ has the same free variables of φ^* with the exception of x; and (iii) $\Theta \vdash_{\Sigma} (\exists x\, \varphi) \Leftrightarrow \varphi^*$.

[1] In this work, we consider first-order logic with equality $\cong$.

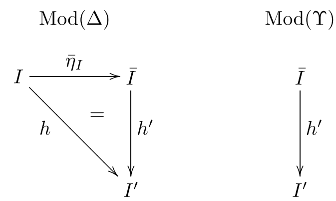


Fig. 2 Δ has algebraically prime models with respect to Υ

3 Iteration

We start by defining when a model $\bar{I}$ of a theory Υ is an *algebraic prime extension of a model* I of a theory Δ over the same signature as Υ and such that $\text{Mod}(\Upsilon) \subseteq \text{Mod}(\Delta)$: this means that there is an embedding $\bar{\eta}_I : I \to \bar{I}$ such that for every embedding $h : I \to I'$ with I' in $\text{Mod}(\Upsilon)$ there is an embedding $h' : \bar{I} \to I'$ with $h = h' \circ \bar{\eta}_I$ (see Fig. 2). In this case, $\bar{I}$ is also said to be *algebraically prime with respect to* I *via* $\bar{\eta}_I$.

So, given first-order theories Δ and Υ over the same signature, we say that Δ has *algebraically prime models with respect* to Υ whenever

- $\text{Mod}(\Upsilon) \subseteq \text{Mod}(\Delta)$; and
- There is a map F that associates to each model I of Δ an *algebraically prime model* $\bar{I} = F(I)$ of Υ.

When Δ is $\Upsilon^{\forall}$, this notion coincides with the notion of Υ having *algebraically prime models*.[2]

In general, the proof that a theory Δ has algebraically prime models with respect to another theory Υ, can be split in the following three steps: (a) definition of map $F : \text{Mod}(\Delta) \to \text{Mod}(\Upsilon)$; (b) for each model I of Δ, the definition of the embedding $\bar{\eta}_I$; (c) verification of the universal property in Fig. 2.

Looking at several examples, we arrived at the conclusion that the map F from $\text{Mod}(\Delta)$ to $\text{Mod}(\Upsilon)$ can be obtained by using an iterative construction over a one-step map E from $\text{Mod}(\Delta)$ to $\text{Mod}(\Delta)$ where $E(I)$ is closer to being algebraically prime with respect to I. By iteratively applying E to I, we obtain $F(I)$ which is an algebraically prime model with respect to I.

The idea is that for showing that a theory has quantifier elimination using this method, it is only necessary to worry about the one step construction. A general result will state that if the one step construction fulfills certain properties then Δ will have algebraically prime models with respect to Υ. With this purpose in mind, assume that we start with:

- First-order theories Δ and Υ contained in $\forall_2$ over a signature Σ and with $\text{Mod}(\Upsilon) \subseteq \text{Mod}(\Delta)$;[3]
- A map $E : \text{Mod}(\Delta) \to \text{Mod}(\Delta)$;

[2]Recall that, for any first-order theory Θ, $\Theta^{\forall}$ is the set of all sentences entailed by Θ of the form $\forall x_1 \ldots \forall x_n \varphi$ where φ is a quantifier-free formula.

[3]Recall that $\forall_2$ is the smallest class of formulas containing $\exists_1$ and closed under $\wedge$, $\vee$ and adding universal quantifiers at the front, where $\exists_1$ is the smallest class of formulas containing the quantifier-free formulas and closed under $\wedge$, $\vee$ and adding existential quantifiers at the front. Observe that every $\forall_2$ formula is equivalent to a $\forall_2$ formula $\forall x_1 \ldots \forall x_n \psi$ with ψ in $\exists_1$ (for more details see, for instance, Sect. 2.4 of [7]). From now on we assume, without loss of generality, that the $\forall_2$ formulas are of this form, i.e. of the form $\forall x_1 \ldots \forall x_n \psi$ with ψ in $\exists_1$.

- A family of embeddings $\eta = \{\eta_I : I \to E(I)\}_{I \in \mathrm{Mod}(\Delta)}$.

We say that E *extends in one step* Δ *towards* Υ *via* η whenever:

- E is *quasi-adjoint for* Υ *via* η, i.e.
  - For every I in $\mathrm{Mod}(\Delta)$, I' in $\mathrm{Mod}(\Upsilon)$ and embedding $h : I \to I'$ there is an embedding h' from $E(I)$ to I' such that $h' \circ \eta_I = h$;
- E *increments local satisfaction for* Υ *via* η, i.e.
  - For every $\forall x_1 \ldots \forall x_n \varphi$ in $\Upsilon \setminus \Delta$ where φ does not have universal quantifiers, I in $\mathrm{Mod}(\Delta)$ and assignment ρ over I, if $I\rho \not\Vdash_\Sigma \varphi$ then $E(I)\eta_I \circ \rho \Vdash_\Sigma \varphi$.[4]

The objective now is to prove, using these assumptions, that Δ has algebraically prime models with respect to Υ. We start by introducing E^n and $\eta^{p,q}$. Let

$$E^n : \mathrm{Mod}(\Delta) \to \mathrm{Mod}(\Delta) \quad \text{for } n \in \mathbb{N}$$

be the family of maps inductively defined as follows: E^0 is $\mathrm{id}_{\mathrm{Mod}(\Delta)}$ and E^{n+1} is $E \circ E^n$. Moreover, given natural numbers p and q with $p \leq q$, let $\eta^{p,q}$ be the family

$$\left\{\eta_I^{p,q} : E^p(I) \to E^q(I)\right\}_{I \in \mathrm{Mod}(\Delta)}$$

of embeddings inductively defined on $q - p$ as follows: $\eta_I^{p,p}$ is $\mathrm{id}_{E^p(I)}$ and $\eta_I^{p,q} = \eta_{E^{q-1}(I)} \circ \eta_I^{p,q-1}$.

Finally, we define the map E^ω as well as the family of embeddings η^ω. Given $I \in \mathrm{Mod}(\Delta)$, let

$$E^\omega(I) = \left(D_\omega, \cdot^{\mathsf{F}_\omega}, \cdot^{\mathsf{P}_\omega}\right)$$

be as follows:

- D_ω is the quotient set of $\biguplus_{n \in \mathbb{N}} D_n$ by the binary relation $\sim$ where

$$d \sim e \quad \text{iff } \eta_I^{k,m}(d) = e$$

  for every $d \in D_k, e \in D_m$ with $k, m \in \mathbb{N}$ and $k \leq m$;
- $f^{\mathsf{F}_\omega}([d_1], \ldots, [d_n]) = [f^{\mathsf{F}_k}(\eta_I^{k_1,k}(d_1), \ldots, \eta_I^{k_n,k}(d_n))]$, whenever $d_i \in D_{k_i}$ for $i = 1, \ldots, n$ and $k = \max\{k_1, \ldots, k_n\}$;

[4]Given a signature Σ, an interpretation structure $I = (D, \cdot^F, \cdot^P)$ over Σ, a variable assignment $\rho : X \to D$, and a first-order formula φ, we denote by $I\rho \Vdash_\Sigma \varphi$ the satisfaction of φ by I and ρ. Recall that this relation is inductively defined as follows:

- $I\rho \Vdash_\Sigma p(t_1, \ldots, t_n)$ whenever $([\![t_1]\!]^{I\rho}, \ldots, [\![t_n]\!]^{I\rho}) \in p^P$ for every terms $t_1, \ldots, t_n$ and n-ary predicate symbol p, where $[\![t]\!]^{I\rho}$ is the interpretation of term t over I and ρ, inductively defined as follows: (a) $[\![x]\!]^{I\rho} = \rho(x)$ for every variable x; and (b) $[\![f(t_1, \ldots, t_n)]\!]^{I\rho} = f^F([\![t_1]\!]^{I\rho}, \ldots, [\![t_n]\!]^{I\rho})$ for every n-ary function symbol f and terms $t_1, \ldots, t_n$;
- $I\rho \Vdash_\Sigma \neg\varphi_1$ whenever $I\rho \not\Vdash_\Sigma \varphi_1$;
- $I\rho \Vdash_\Sigma (\varphi_1 \Rightarrow \varphi_2)$ whenever $I\rho \Vdash_\Sigma \varphi_1$ implies $I\rho \Vdash_\Sigma \varphi_2$;
- $I\rho \Vdash_\Sigma \forall x \varphi_1$ whenever for every ρ' over I with $\rho'(y) = \rho(y)$ for every $y \neq x$, $I\rho' \Vdash_\Sigma \varphi_1$.

- $p^{\mathsf{P}_\omega}([d_1], \ldots, [d_n]) = p^{\mathsf{P}_k}(\eta_I^{k_1,k}(d_1), \ldots, \eta_I^{k_n,k}(d_n))$, whenever $d_i \in D_{k_i}$ for $i = 1, \ldots, n$ and $k = \max\{k_1, \ldots, k_n\}$

and let η_I^ω be the family $\{\eta_I^{n,\omega} : E^n(I) \to E^\omega(I)\}_{n \in \mathbb{N}}$ of embeddings such that $\eta_I^{n,\omega}(d) = [d]$. We denote by η^ω the map $\eta^{0,\omega}$. Observe that $(E^\omega(I), \eta_I^\omega)$ is a direct limit of the directed diagram

$$\left(\left\{E^n(I)\right\}_{n \in \mathbb{N}}, \left\{\eta_I^{p,q} : E^p(I) \to E^q(I)\right\}_{p,q \in \mathbb{N}, p \leq q}\right).$$

The idea is that F is E^ω and $\bar{\eta}_I$ is η_I^ω for each model I of Δ. We start by proving that $E^\omega(I)$ is a model of Υ whenever I is a model of Δ, provided that E satisfies the conditions above.

Proposition 3.1 *Assuming that E extends in one step Δ towards Υ via η, then E^ω is a map from* $\mathrm{Mod}(\Delta)$ *to* $\mathrm{Mod}(\Upsilon)$.

Proof We must show that for each model I of Δ, $E^\omega(I)$ is a model of Υ. Let $\forall x_1 \ldots \forall x_n \varphi$ be a sentence in Υ where φ does not have universal quantifiers, and ρ^ω an assignment over $E^\omega(I)$. Denote by $\{y_1, \ldots, y_m\}$ the set of variables that occur free in φ and for each $j = 1, \ldots, m$, let k_j be a natural number such that $\rho^\omega(y_j) \in \eta_I^{k_j,\omega}(E^{k_j}(I))$. Moreover, denote by k the maximum of $\{k_1, \ldots, k_m\}$ and let ρ^k be an assignment over $E^k(I)$ such that

$$\eta_I^{k,\omega}\big(\rho^k(y_j)\big) = \rho^\omega(y_j)$$

for $j = 1, \ldots, m$. Observe that such an assignment exists since

$$\eta_I^{k,\omega} \circ \eta_I^{k_j,k} = \eta_I^{k_j,\omega}.$$

One of the following two cases hold:

(a) $\forall x_1 \ldots \forall x_n \varphi \in \Delta$. Then $I \Vdash_\Sigma \forall x_1 \ldots \forall x_n \varphi$ and so $E^\omega(I) \Vdash_\Sigma \forall x_1 \ldots \forall x_n \varphi$ since satisfaction of $\forall_2$ sentences is preserved by directed limits, see Theorem 2.4.4 and Theorem 2.4.6 in [7].

(b) $\forall x_1 \ldots \forall x_n \varphi \notin \Delta$. Then, one of the following two cases hold:

  (i) $E^k(I), \rho^k \Vdash_\Sigma \varphi$. Then $E^\omega(I), \eta_I^{k,\omega} \circ \rho^k \Vdash_\Sigma \varphi$ since φ is in the closure for $\wedge$ and $\vee$ of the class of formulas $\exists_1$, $\eta_I^{k,\omega}$ is an embedding from $E^k(I)$ to $E^\omega(I)$, and embeddings preserve satisfaction of such formulas. Therefore, $E^\omega(I), \rho^\omega \Vdash_\Sigma \varphi$.

  (ii) $E^k(I), \rho^k \nVdash_\Sigma \varphi$. Then, since E via η increments local satisfaction for Υ,

$$E^{k+1}(I), \eta_{E^k(I)} \circ \rho^k \Vdash_\Sigma \varphi$$

  and so, since φ is in the closure for $\wedge$ and $\vee$ of the class of formulas $\exists_1$, $\eta_I^{k+1,\omega}$ is an embedding from $E^{k+1}(I)$ to $E^\omega(I)$, and embeddings preserve satisfaction of such formulas,

$$E^\omega(I), \eta_I^{k+1,\omega} \circ \eta_{E^k(I)} \circ \rho^k \Vdash_\Sigma \varphi.$$

  Therefore, $E^\omega(I), \rho^\omega \Vdash_\Sigma \varphi$ since $\eta_I^{k+1,\omega} \circ \eta_{E^k(I)} = \eta_I^{k,\omega}$. □

We now show that Δ has algebraically prime models with respect to Υ, under the conditions above.

Theorem 3.2 *Assuming that E extends in one step Δ towards Υ via η, then for each model I of Δ, $E^\omega(I)$ is algebraically prime with respect to I via η_I^ω. Hence, Δ has algebraically prime models with respect to Υ.*

Proof Let I be a model of Δ, I' a model of Υ and h an embedding of I into I', Consider a family $\{g^k : E^k(I) \to I'\}_{k\in\mathbb{N}}$ of embeddings where:

- g^0 is h;
- g^k is such that $g^k \circ \eta_{E^{k-1}(I)} = g^{k-1}$ (there is such an embedding since E via η is quasi-adjoint for Υ).

Observe that

$$(\dagger) \quad g^q \circ \eta_I^{p,q} = g^p$$

for any natural numbers p and q with $p \le q$, as can be shown by induction on $q - p$.

Let $h' : E^\omega(I) \to I'$ be such that

$$h'(b) = g^k(a)$$

where $k \in \mathbb{N}$ and $a \in E^k(I)$ are such that $b = \eta_I^{k,\omega}(a)$. Then:

1. h' is well defined.
   Let n be a natural number and d in $E^n(I)$ such that $b = \eta_I^{n,\omega}(d)$. Suppose without loss of generality that $k \le n$. Then, $\eta_I^{k,n}(a) = d$, and, so, by ($\dagger$), $g^k(a) = g^n(\eta_I^{k,n}(a)) = g^n(d)$.
2. h' is injective.
   Assume that $h'(b_1) = h'(b_2)$. Let k_1 and k_2 be natural numbers, and a_1 and a_2 elements of $E^{k_1}(I)$ and $E^{k_2}(I)$ respectively, such that $\eta_I^{k_1,\omega}(a_1) = b_1$ and $\eta_I^{k_2,\omega}(a_2) = b_2$. Then $g^{k_1}(a_1) = h'(b_1) = h'(b_2) = g^{k_2}(a_2)$ taking into account the definition of h'. Without loss of generality, assume that $k_1 \le k_2$. Then $g^{k_2}(\eta_I^{k_1,k_2}(a_1)) = g^{k_1}(a_1)$ by ($\dagger$). Hence $\eta_I^{k_1,k_2}(a_1) = a_2$ since g^{k_2} is injective. Therefore, $\eta_I^{k_1,\omega}(a_1) = \eta_I^{k_2,\omega}(a_2)$, and so $b_1 = b_2$.
3. h' is an homomorphism. Straightforward.
4. $h' \circ \eta_I^{0,\omega} = h$. Indeed, let d be an arbitrary element of I. Then $h'(\eta_I^{0,\omega}(d)) = g^0(d)$ by definition of h'. The thesis follows since $g^0(d)$ is $h(d)$. □

We now illustrate these notions and results with several examples involving the theory of natural numbers with successor, the theory of term algebras and the theory of algebraically closed fields.

Natural Numbers with Successor Let Σ_S be the signature with $F_0 = \{0\}$ and $F_1 = \{\mathbf{S}\}$, for the theory Θ_S containing the sentences:

S1 $\forall x(\neg(\mathbf{S}\, x \cong 0))$;
S2 $\forall x \forall y((\mathbf{S}\, x \cong \mathbf{S}\, y) \Rightarrow (x \cong y))$;
S3 $\forall y((\neg(y \cong 0)) \Rightarrow (\exists x(y \cong \mathbf{S}\, x)))$;
S4 $\forall x(\neg(\mathbf{S}^n\, x \cong x))$ for each $n \in \mathbb{N}^+$

for the natural numbers with successor, see Sect. 3.1 of [3]. Taking into account Theorem 3.2, in order to show that $\Theta_S^\forall$ has algebraically prime models with respect to Θ_S, it is enough to show the following conditions:

1. The sentences of $\Theta_S^\forall$ and Θ_S are in $\forall_2$;
2. $\mathrm{Mod}(\Theta_S) \subseteq \mathrm{Mod}(\Theta_S^\forall)$;
3. There is a map $E : \mathrm{Mod}(\Theta_S^\forall) \to \mathrm{Mod}(\Theta_S^\forall)$ and a family of embeddings $\eta = \{\eta_I : I \to E(I)\}_{I \in \mathrm{Mod}(\Theta_S^\forall)}$ such that
   a. E via η is quasi-adjoint for Θ_S;
   b. E via η increments local satisfaction for Θ_S.

It is immediate to see that Conditions 1 and 2 hold. We now show that Condition 3 also holds. Let $E : \mathrm{Mod}(\Theta_S^\forall) \to \mathrm{Mod}(\Theta_S^\forall)$ be such that

$$E\big((D, \cdot^F, \cdot^P)\big) = \big(D^\bullet, \cdot^{F^\bullet}, \cdot^{P^\bullet}\big)$$

where

- $D^\bullet$ is $D \cup \{d^\bullet : d^\bullet \notin D, d \in D \setminus \{0^F\},$ there is no e in D with $\mathbf{S}^F(e) = d\}$;
- $0^{F^\bullet} = 0^F$;
- $\mathbf{S}^{F^\bullet}(e) = \begin{cases} \mathbf{S}^F(e) & \text{if } e \text{ is in } D, \\ d & \text{if } e \text{ is } d^\bullet; \end{cases}$

and η be the family $\{\eta_I : I \to E(I)\}_{I \in \mathrm{Mod}(\Theta_S^\forall)}$ where each η_I is the inclusion of I into $E(I)$.

Clearly, for each model I of $\Theta_S^\forall$, $E(I)$ is also a model of $\Theta_S^\forall$. In fact, observe that I is a substructure of a model of Θ_S, see Exercise 2.5.10 of [9]. Hence, $E(I)$ is also a substructure of that model, and, so, it is a model of $\Theta_S^\forall$.

Proposition 3.3 *The map E via η is quasi-adjoint for Θ_S.*

Proof Let $I = (D, \cdot^F, \cdot^P)$ be a model of $\Theta_S^\forall$, $I' = (D', \cdot^{F'}, \cdot^{P'})$ a model of Θ_S, h an embedding from I to I', and h' a map from $E(I)$ to I' defined as follows:

$$h'(e) = \begin{cases} h(e) & \text{if } e \text{ is in } D, \\ (\mathbf{S}^{F'})^{-1}(h(d)) & \text{if } e \text{ is } d^\bullet. \end{cases}$$

Then

1. h' is well defined. It is enough to see that if e is $d^\bullet$ then there is one and only one e' in D' with $\mathbf{S}^{F'}(e') = h(d)$. Indeed, by S3, there is one e' in D' with $\mathbf{S}^{F'}(e') = h(d)$ since $h(d)$ is not $0^{F'}$ because d is not 0^F and h is an embedding. There is at most one e' in D' with $\mathbf{S}^{F'}(e') = h(d)$ since $\mathbf{S}^{F'}$ is one to one by S2.
2. h' is one to one. Since h is an embedding, it is enough to show that
   (i) For each e_1 of the form $d_1^\bullet$, $h'(e_1) \notin h(D)$. Suppose by contradiction that $h'(e_1) \in h(D)$ and let $d \in D$ be such that $h'(e_1) = h(d)$. Then $(\mathbf{S}^{F'})^{-1}(h(d_1)) = h(d)$. Hence $h(d_1) = \mathbf{S}^{F'}((\mathbf{S}^{F'})^{-1}(h(d_1))) = \mathbf{S}^{F'}(h(d)) = h(\mathbf{S}^F(d))$. Since h is one to one then $d_1 = \mathbf{S}^F(d)$ which contradicts the fact that there is no e in D with $\mathbf{S}^F(e) = d_1$.

(ii) For each e_1 and e_2 of the forms $d_1^\bullet$ and $d_2^\bullet$, respectively, if $h'(e_1) = h'(e_2)$ then $e_1 = e_2$. Suppose that $h'(e_1) = h'(e_2)$. Then $(\mathbf{S}^{F'})^{-1}(h(d_1)) = (\mathbf{S}^{F'})^{-1}(h(d_2))$ and so $h(d_1) = \mathbf{S}^{F'}((\mathbf{S}^{F'})^{-1}(h(d_1))) = \mathbf{S}^{F'}((\mathbf{S}^{F'})^{-1}(h(d_2))) = h(d_2)$. Since h is an embedding, $d_1 = d_2$, and so $e_1 = e_1$.

3. $h'(0^{F^\bullet}) = 0^{F'}$. Indeed, $h'(0^{F^\bullet}) = h'(0^F) = h(0^F) = 0^{F'}$ since h is an embedding.
4. $h'(\mathbf{S}^{F^\bullet}(e)) = \mathbf{S}^{F'}(h'(e))$. In order to show this, consider two cases:

(i) $e \notin D$. Suppose e is $d^\bullet$. Then $h'(\mathbf{S}^{F^\bullet}(e)) = h'(d) = h(d) = \mathbf{S}^{F'}((\mathbf{S}^{F'})^{-1}(h(d))) = \mathbf{S}^{F'}(h'(e))$;
(ii) e is in D. Then $h'(\mathbf{S}^{F^\bullet}(e)) = h'(\mathbf{S}^F(e)) = h(\mathbf{S}^F(e)) = \mathbf{S}^{F'}(h(e)) = \mathbf{S}^{F'}(h'(e))$.

5. $h' \circ \eta_I = h$. Indeed, let $d \in D$. Then $h'(\eta_I(d)) = h'(d) = h(d)$. □

Proposition 3.4 *The map E via η increments local satisfaction for Θ_{S}.*

Proof Let $I = (D, \cdot^F, \cdot^P)$ be a model of $\Theta_{\mathrm{S}}^{\forall}$ and ρ an assignment over I. Observe that $\Theta_{\mathrm{S}} \setminus \Theta_{\mathrm{S}}^{\forall} = \{\mathrm{S3}\}$. Suppose $I\rho \not\Vdash_{\Sigma_{\mathrm{S}}} (\neg(y \cong 0)) \Rightarrow (\exists x(y \cong \mathbf{S}x))$. Then $I\rho \Vdash_{\Sigma_{\mathrm{S}}} \neg(y \cong 0)$ and $I\rho \not\Vdash_{\Sigma_{\mathrm{S}}} \exists x(y \cong \mathbf{S}x)$. Hence $\rho(y) \neq 0^F$ and there is no e in D with $\rho(y) = \mathbf{S}^F(e)$. Observe that $\rho(y)^\bullet \in E(I)$, $\rho(y) = \mathbf{S}^{F^\bullet}(\rho(y)^\bullet)$ and $\eta_I(\rho(y)) = \rho(y)$. Therefore, $E(I)\eta_I \circ \rho \Vdash_{\Sigma_{\mathrm{S}}} \exists x(y \cong \mathbf{S}x)$ and so $E(I)\eta_I \circ \rho \Vdash_{\Sigma_{\mathrm{S}}} (\neg(y \cong 0)) \Rightarrow (\exists x(y \cong \mathbf{S}x))$. □

So, taking into account that E via η is quasi-adjoint for Θ_{S} and increments local satisfaction for Θ_{S}, see Propositions 3.3 and 3.4, respectively, we can use Theorem 3.2 to conclude that the following theorem holds.

Theorem 3.5 *Theory Θ_{S} has algebraically prime models.*

Term Algebras Consider the first-order theory for term algebras induced by a given first-order signature with no predicate symbols, see [7]. We show that it has algebraically prime models using our iteration criteria.

Given a signature Σ with no predicate symbols, let Σ_{ta} be the signature induced by Σ with $F_{\mathrm{ta}_1} = \{f_i : n \in \mathbb{N}, f \in F_n \text{ and } i \in \{1, \dots, n\}\}$ and $P_{\mathrm{ta}_1} = \{\mathbf{Is}_c : c \in F_0\} \cup \{\mathbf{Is}_f : n \in \mathbb{N} \text{ and } f \in F_n\}$ for the theory Θ_{ta} containing the sentences:

(T1) $\exists^1 x\, \mathbf{Is}_c(x)$
for each constant symbol c in F_0;
(T2) $\forall x_1 \dots \forall x_n \exists^1 x (\mathbf{Is}_f(x) \wedge f_1(x) \cong x_1 \wedge \cdots \wedge f_n(x) \cong x_n)$
for each function symbol f in F_n and $n \in \mathbb{N}$;
(T3) $\forall x\, \neg(\mathbf{Is}_{e_1}(x) \wedge \mathbf{Is}_{e_2}(x))$
where e_1 and e_2 are distinct constant or function symbols;
(T4) $\forall x((\neg\, \mathbf{Is}_f(x)) \Rightarrow f_i(x) \cong x)$
for each function symbol f in F_n, $n \in \mathbb{N}$ and $i = 1, \dots, n$;
(T5) $\forall x((t(f_i(x)) \cong x) \Rightarrow \neg\, \mathbf{Is}_f(x))$
for each function symbol f in F_n, $n \in \mathbb{N}$, and sequence t of function symbols over Σ_{ta}, and $i = 1, \dots, n$.

The idea is that, given a model of Θ_{ta}, the values in its carrier set that correspond to terms of Σ are the ones for which one of the predicates **Is_** hold. For example, if a value of the domain satisfies predicate $\mathbf{Is}_f$ for a function symbol f in Σ, then that value is the denotation of a term of Σ whose main constructor is f. With this in mind, observe that T1 guarantees that each constant of Σ has a unique value corresponding to it, and T2 does the same for each function symbol of Σ and domain values (corresponding or not to terms of Σ) as arguments. Axiom T3 ensures that each value in the domain of a model of Θ_{ta} corresponds to at most one term of Σ, and axiom T4 requires that if a domain value is not a representative of a term with main constructor f then its projection along f is itself. Finally, axiom T5 requires that any representative of a non-constant term is different from each of its arguments, as well as from each of the arguments of its arguments, and so on.

Note that when Σ is the signature Σ^{stc} with $F_0^{\text{stc}} = \mathbb{N}$, $F_2^{\text{stc}} = \{\mathsf{push}\}$ and the other sets are empty, then $\Theta_{\text{ta}}^{\text{stc}}$ is a first-order theory for stacks. Observe that, in this case, push_1 and push_2 correspond to the usual stack operations of pop and top.

Taking into account Theorem 3.2, in order to show that $\Theta_{\text{ta}}^{\forall}$ has algebraically prime models with respect to Θ_{ta}, it is enough to show the following conditions:

1. The sentences of $\Theta_{\text{ta}}^{\forall}$ and Θ_{ta} are in $\forall_2$;
2. $\text{Mod}(\Theta_{\text{ta}}) \subseteq \text{Mod}(\Theta_{\text{ta}}^{\forall})$;
3. There are a map $E : \text{Mod}(\Theta_{\text{ta}}^{\forall}) \to \text{Mod}(\Theta_{\text{ta}}^{\forall})$ and a family of embeddings $\eta = \{\eta_I : I \to E(I)\}_{I \in \text{Mod}(\Theta_{\text{ta}}^{\forall})}$ such that
   (a) E via η is quasi-adjoint for Θ_{ta};
   (b) E via η increments local satisfaction for Θ_{ta}.

It is immediate to see that Conditions 1 and 2 hold. We now show that Condition 3 also holds.

Let $E : \text{Mod}(\Theta_{\text{ta}}^{\forall}) \to \text{Mod}(\Theta_{\text{ta}}^{\forall})$ be such that

$$E\big((D, \cdot^F, \cdot^P)\big) = \big(D^*, \cdot^{F^*}, \cdot^{P^*}\big)$$

where

- D^* is the union of D with
  - $\{d_c : c \text{ in } F_0, d_c \notin D$ and there is no e in D with $\mathbf{Is}_c^F(e) = 1\}$;
  - $\{\langle d_1, \dots, d_n\rangle_f : d_1, \dots, d_n \in D, n \in \mathbb{N}, f \in F_n, \langle d_1, \dots, d_n\rangle_f \notin D$ and there is no e in D with $\mathbf{Is}_f^F(e) = 1$ and $f_1^F(e) = d_1, \dots, f_n^F(e) = d_n\}$;
- For every $c \in F_0$,

$$\mathbf{Is}_c^{F^*}(d) = \begin{cases} \mathbf{Is}_c^F(d) & \text{if } d \in D,\\ 1 & \text{if } d \text{ is } d_c,\\ 0 & \text{otherwise;}\end{cases}$$

- For every $f \in F_n$,

$$\mathbf{Is}_f^{F^*}(d) = \begin{cases} \mathbf{Is}_f^F(d) & \text{if } d \in D,\\ 1 & \text{if } d \text{ is } \langle d_1, \dots, d_n\rangle_f,\\ 0 & \text{otherwise;}\end{cases}$$

- for every $f \in F_n$ and $i \in \{1, \ldots, n\}$,

$$f_i^{F^*}(d) = \begin{cases} f_i^F(d) & \text{if } d \in D \\ d_i & \text{if } d \text{ is } \langle d_1, \ldots, d_n \rangle_f \\ d & \text{otherwise;} \end{cases}$$

and η be the family $\{\eta_I : I \to E(I)\}_{I \in \mathrm{Mod}(\Theta_{\mathrm{ta}}^{\forall})}$ where each η_I is the inclusion of I into $E(I)$.

Clearly, for each model I of $\Theta_{\mathrm{ta}}^{\forall}$, $E(I)$ is also a model of $\Theta_{\mathrm{ta}}^{\forall}$, since I is a substructure of a model of Θ_{ta}, by Exercise 2.5.10 of [9], and, so, by definition, $E(I)$ is also a substructure of a model of Θ_{ta}. Hence, see Exercise 2.5.10 of [9], $E(I)$ is a model of $\Theta_{\mathrm{ta}}^{\forall}$.

Proposition 3.6 *The map E via η is quasi-adjoint for Θ_{ta}.*

Proof Let $I = (D, \cdot^F, \cdot^P)$ be a model of $\Theta_{\mathrm{ta}}^{\forall}$, $I' = (D', \cdot^{F'}, \cdot^{P'})$ a model of Θ_{ta}, h an embedding from I to I', and h' a map from $E(I)$ to I' defined as follows:

$$h'(e) = \begin{cases} h(e) & \text{if } e \text{ is in } D, \\ d' & \text{if } e \text{ is } d_c \text{ and } d' \text{ is such that } \mathbf{Is}_c^{F'}(d') = 1, \\ d' & \text{if } e \text{ is } \langle d_1, \ldots, d_n \rangle_f \text{ and } d' \text{ is such that } \mathbf{Is}_f^{F'}(d') = 1 \text{ and} \\ & f_i^{F'}(d') = h(d_i) \text{ for } i = 1, \ldots, n. \end{cases}$$

Then

1. h' is well defined. It is enough to see that if e is d_c then by axiom T1 there is a unique d' in D' with $\mathbf{Is}_c^{F'}(d') = 1$, and if e is $\langle d_1, \ldots, d_n \rangle_f$ then by axiom T2 there is a unique d' in D' with $\mathbf{Is}_f^{F'}(d') = 1$ and $f_i^{F'}(d') = h(d_i)$ for $i = 1, \ldots, n$.
2. h' is one to one. Since h is an embedding, it is enough to show that:
   (i) For each e in D^* of the form d_c, $h'(e) \notin h(D)$. Suppose by contradiction that $h'(e) \in h(D)$ and let $d \in D$ be such that $h'(e) = h(d)$. Observe that $\mathbf{Is}_c^{F'}(h'(e)) = 1$ and thus $\mathbf{Is}_c^{F'}(h(d)) = 1$. Hence $\mathbf{Is}_c^F(d) = 1$ since h is an embedding which contradicts the existence of d_c in D^*.
   (ii) For each e in D^* of the form $\langle d_1, \ldots, d_n \rangle_f$, $h'(e) \notin h(D)$. The proof of this case is similar to the proof of case (i) and so we omit it.
   (iii) For each e_1 and e_2 in D^* of the forms d_{c_1} and d_{c_2} respectively, if $h'(e_1) = h'(e_2)$ then $e_1 = e_2$. Suppose that $h'(e_1) = h'(e_2)$. Then by definition of h' the constants c_1 and c_2 are the same and so e_1 and e_2 are the same.
   (iv) For each e_1 and e_2 in D^* of the forms $\langle d_1^1, \ldots, d_n^1 \rangle_{f_1}$ and $\langle d_1^2, \ldots, d_n^2 \rangle_{f_2}$ respectively, if $h'(e_1) = h'(e_2)$ then $e_1 = e_2$. The proof proceeds as in (iii) and so we omit it.
   (v) For each e_1 and e_2 of the forms $\langle d_1, \ldots, d_n \rangle_f$ and d_c, respectively, $h'(e_1) \neq h'(e_2)$. Immediate by definition of h' taking also into account axiom T3.
3. $h'(f_i^{F^*}(e)) = f_i^{F'}(h'(e))$ for every $e \in D^*$. Consider the following three cases:
   (i) e is in D. Then $h'(f_i^{F^*}(e)) = h'(f_i^F(e)) = h(f_i^F(e)) = f_i^{F'}(h(e)) = f_i^{F'}(h'(e))$.

(ii) e is not in D and is not of the form $\langle d_1, \ldots, d_n \rangle_f$ for some $d_1, \ldots, d_n$ in D. Then $h'(f_i^{F^*}(e)) = h'(e) = f_i^{F'}(h'(e))$ by axiom T4.

(iii) e is of the form $\langle d_1, \ldots, d_n \rangle_f$. Then $h'(f_i^{F^*}(e)) = h'(d_i) = h(d_i) = f_i^{F'}(h'(e))$ by definition of h'.

4. $\mathbf{Is}_f^{F^*}(e) = 1$ iff $\mathbf{Is}_f^{F'}(h'(e)) = 1$. Consider the following three cases:

   (i) $e \in D$. Then $\mathbf{Is}_f^{F^*}(e) = 1$ iff $\mathbf{Is}_f^{F}(e) = 1$ iff $\mathbf{Is}_f^{F'}(h(e)) = 1$ iff $\mathbf{Is}_f^{F'}(h'(e)) = 1$.

   (ii) $e \notin D$ and is not of the form $\langle d_1, \ldots, d_n \rangle_f$ for some $d_1, \ldots, d_n$ in D. Then $\mathbf{Is}_f^{F^*}(e) = 0 = \mathbf{Is}_f^{F'}(h'(e))$ by definition of h'.

   (iii) e is of the form $\langle d_1, \ldots, d_n \rangle_f$. Then $\mathbf{Is}_f^{F^*}(e) = 1 = \mathbf{Is}_f^{F'}(h'(e))$ by definition of h'.

5. $\mathbf{Is}_c^{F^*}(e) = 1$ iff $\mathbf{Is}_c^{F'}(h'(e)) = 1$. The proof of this case is omitted since it is similar to the proof of case 4.
6. $h' \circ \eta_I = h$. Indeed, $h'(\eta_I(d)) = h'(d) = h(d)$ for every $d \in D$. □

Proposition 3.7 *The map E via η increments local satisfaction for Θ_{ta}.*

Proof Let $I = (D, \cdot^F, \cdot^P)$ be a model of $\Theta_{\mathrm{ta}}^{\forall}$ and ρ an assignment over I. Observe that, in order to simplify the presentation of Θ_{ta}, the axioms $T1$ and $T2$ use the quantifier $\exists^1$. This quantifier is used as an abbreviation of two sentences, one having an existential quantifier and the other with only universal quantifiers. In the context of this proof, it is important that we not use these abbreviations and so we are seeing axiom T_1 as two axioms, and the same for axiom T_2. Then, $\Theta_{\mathrm{ta}} \setminus \Theta_{\mathrm{ta}}^{\forall}$ is the set $\{\forall x_1 \ldots \forall x_n \exists x (\mathbf{Is}_f(x) \wedge f_1(x) \cong x_1 \wedge \cdots \wedge f_n(x) \cong x_n), \exists x_1\, \mathbf{Is}_c(x_1)\}$.

1. Suppose $I\rho \not\Vdash_{\Sigma_{\mathrm{ta}}} \exists x_1\, \mathbf{Is}_c(x_1)$. Then $\mathbf{Is}_c^F(d) = 0$ for every d in D. Hence $d_c \in D^*$ and moreover $\mathbf{Is}_c^{F^*}(d_c) = 1$. Thus $E(I)\eta_I \circ \rho \Vdash_{\Sigma_{\mathrm{ta}}} \exists x_1\, \mathbf{Is}_c(x_1)$.
2. Suppose $I\rho \not\Vdash_{\Sigma_{\mathrm{ta}}} \exists x (\mathbf{Is}_f(x) \wedge f_1(x) \cong x_1 \wedge \cdots \wedge f_n(x) \cong x_n)$. Then there is no e in D with $\mathbf{Is}_f^F(e) = 1$ and $f_1^F(e) = \rho(x_1), \ldots, f_n^F(e) = \rho(x_n)$. Hence $\langle \rho(x_1), \ldots, \rho(x_n) \rangle_f \in D^*$ with $f_i^{F^*}(\langle \rho(x_1), \ldots, \rho(x_n) \rangle_f) = \rho(x_i)$ for every $i = 1, \ldots, n$ and $\mathbf{Is}_f^{F^*}(\langle \rho(x_1), \ldots, \rho(x_n) \rangle_f) = 1$. Thus $E(I)\eta_I \circ \rho \Vdash_{\Sigma_{\mathrm{ta}}} \exists x (\mathbf{Is}_f(x) \wedge f_1(x) \cong x_1 \wedge \cdots \wedge f_n(x) \cong x_n)$. □

So, taking into account that E via η is quasi-adjoint for Θ_{ta} and increments local satisfaction for Θ_{ta}, see Propositions 3.6 and 3.7, respectively, we can use Theorem 3.2 to conclude that the following theorem holds.

Theorem 3.8 *Theory Θ_{ta} has algebraically prime models.*

Algebraically Closed Fields—From Θ_{f} to Θ_{acf} Let Σ_{f} be the signature for fields, that is: $F_0 = \{0, 1\}$, $F_1 = \{-\}$, and $F_2 = \{+, \times\}$, Θ_{f} the theory containing the field axioms, and Θ_{acf} the theory over Σ_{f} containing the axioms for algebraically closed fields, see [7–9]. We start by stating a well known theorem, Theorem 2.5 of [8], used extensively in this example.

Theorem 3.9 *Let k be a field. Then there exists an algebraically closed field containing k as a subfield.*

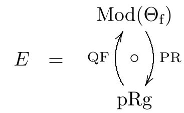


Fig. 3 Definition of E as QF ∘ PR

Proof This is Theorem 2.5 in Chap. 5 of [8]. □

Moreover, let $E = (\mathrm{QF} \circ \mathrm{PR})$ where PR is a map that associates to each field I a polynomial ring $I[V_I]$ where V_I is a set of names v_q in bijection with the set of all polynomials q in $I[x]$ with degree greater than 0, and QF associates to each polynomial ring $I[V_I]$ a field generated by the quotient with a maximal ideal containing the ideal generated by all polynomials $q(v_q)$ in $I[V_I]$ (see Fig. 3—for details consult the proof of Theorem 3.9—Theorem 2.5 in Chap. 5 of [8]). Observe that E associates with each field an algebraic extension of it where every polynomial in one variable of degree at least one with coefficients in the field has a root.

Take η to be the family

$$\{\eta_I : I \to \mathrm{QF}(\mathrm{PR}(I))\}_{I \in \mathrm{Mod}(\Theta_{\mathrm{f}})}$$

of embeddings such that $\eta_I(d) = [p_d^I]$ where p_d^I is the constant polynomial d in $I[V_I]$.

Proposition 3.10 *The map* QF ∘ PR *via* η *is quasi-adjoint for* Θ_{acf}.

Proof Let I' be a model of Θ_{acf} and $h : I \to I'$ an embedding. Since $\mathrm{QF} \circ \mathrm{PR}(I)$ is an algebraic extension of $\eta_I(D)$, there is an embedding $\bar{h}$ from $\mathrm{QF} \circ \mathrm{PR}(I)$ into I' such that $\bar{h} \circ \eta_I = h$ by Theorem 2.8 in Chap. 5 of [8]. □

Proposition 3.11 *The map* QF ∘ PR *via* η *increments local satisfaction for* Θ_{acf}.

Proof Let I be a model of Θ_{f} and ρ an assignment over I. Let n in $\mathbb{N}^+$ and γ be the formula $\exists y(y^n + x_1 y^{n-1} + \cdots + x_n \cong 0)$. Assume that $I\rho \not\Vdash_{\Sigma_{\mathrm{f}}} \exists y(y^n + x_1 y^{n-1} + \cdots + x_n \cong 0)$. Let q be the polynomial in $I[x]$ of the form $x^n + \rho(x_1)x^{n-1} + \cdots + \rho(x_{n-1})x^1 + \rho(x_n)$. Then, as explained in the proof of Theorem 3.9 (Theorem 2.5 of [8]) in Chap. 5 of [8], $\eta_I \circ q$ has a root in $\mathrm{PR}(\mathrm{QF}(I))$. Consider an assignment ρ' over $\mathrm{PR}(\mathrm{QF}(I))$ such that $\rho'(z) = \eta_I \circ \rho(z)$ for every variable $z \neq y$, such that $\rho'(y)$ is that root. Then $\mathrm{PR}(\mathrm{QF}(I))\rho' \Vdash_{\Sigma_{\mathrm{f}}} y^n + x_1 y^{n-1} + \cdots + x_n \cong 0$ and so $\mathrm{PR}(\mathrm{QF}(I))\eta_I \circ \rho \Vdash_{\Sigma_{\mathrm{f}}} \exists y(y^n + x_1 y^{n-1} + \cdots + x_n \cong 0)$. □

Since QF ∘ PR via η is quasi-adjoint for Θ_{acf} and increments local satisfaction for Θ_{acf} by Propositions 3.10 and 3.11, respectively, we can use Theorem 3.2 to conclude that the following corollary holds.

Theorem 3.12 *Theory* Θ_{f} *has algebraically prime models with respect to* Θ_{acf}.

Observe that, although ACF admits quantifier elimination, it is not decidable. To obtain decidability, we must add an axiom to ACF specifying a fixed integral characteristic. Then, we can see that ACF_0 has QE and that the algebraic numbers (the algebraic closure of the rationals $\mathbb{Q}$) are a prime model for ACF_0.

4 Adjunction

An adjunction between two categories establishes a deep relationship between their objects and morphisms. Given first-order theories Υ and Δ with

$$\text{Mod}(\Upsilon) \subseteq \text{Mod}(\Delta),$$

we show in this section that every model of Δ has an algebraically prime model in Υ if the inclusion functor from the category of models of Υ and their embeddings into the category of models of Δ and their embeddings has a left adjoint.

We start by briefly recalling what is a natural transformation and a left adjoint. Given functors $F, H : \mathbf{C} \to \mathbf{D}$, a *natural transformation* $\alpha : F \to H$ is a family

$$\alpha = \left\{\alpha_c : F(c) \to H(c)\right\}_{c \in |\mathbf{C}|}$$

of morphisms in $\mathbf{D}$ such that

$$H(f) \circ \alpha_{c_1} = \alpha_{c_2} \circ F(f)$$

for every morphism $f : c_1 \to c_2$ in $\mathbf{C}$. Moreover, F is said to be *left adjoint* of functor H, denoted by

$$F \dashv H$$

if there is a natural transformation

$$\eta : \text{id}_{\mathbf{C}} \to H \circ F,$$

called the *unit* of the adjunction satisfying the following universal property: Given any morphism $h : c \to H(d)$ in $\mathbf{C}$, there is a unique morphism $\bar{h} : F(c) \to d$ in $\mathbf{D}$ such that

$$H(\bar{h}) \circ \eta_c = h.$$

Theorem 4.1 *Let Υ and Δ be first-order theories with* $\text{Mod}(\Upsilon) \subseteq \text{Mod}(\Delta)$. *If the inclusion functor from* $\text{Mod}(\Upsilon)$ *to* $\text{Mod}(\Delta)$ *has a left adjoint $\bar{E}$ with unit η, then, $\bar{E}(I)$ is algebraically prime with respect to I via η_I.*

Proof Denote the inclusion functor from $\text{Mod}(\Upsilon)$ to $\text{Mod}(\Delta)$ by $J_{\Upsilon\Delta}$. Let I be a model of Δ, I' a model of Υ, and $h : I \to I'$ an embedding. Let h' be the unique embedding from $\bar{E}(I)$ to I' such that $h = J_{\Upsilon\Delta}(h') \circ \eta_I$, which exists since E is a left adjoint of $J_{\Upsilon\Delta}$ with unit η. The thesis follows immediately since $J_{\Upsilon\Delta}(h') = h'$. □

Algebraically Closed Fields—from $\Theta^{\forall}_{\text{acf}}$ to Θ_{f} Observe that the models of $\Theta^{\forall}_{\text{acf}}$ are the integral domains since

$$\Theta_{\text{acf}} \vDash_{\Sigma_\text{f}} \forall x_1 \forall x_2 \left(\left(\left(\neg(x_1 \cong 0)\right) \wedge \left(\neg(x_2 \cong 0)\right)\right) \Rightarrow \left(\neg(x_1 \times x_2) \cong 0\right)\right),$$

and that every field is an integral domain, i.e. $\text{Mod}(\Theta_\text{f}) \subseteq \text{Mod}(\Theta^{\forall}_{\text{acf}})$.

Denote by J the inclusion functor from the category of models of Θ_f and their embeddings into the category of models of $\Theta^{\forall}_{\text{acf}}$ and their embeddings, and by FF the functor

$$\mathrm{Mod}(\Theta^{\forall}_{\mathrm{acf}}) \xrightarrow{\mathrm{FF}} \mathrm{Mod}(\Theta_{\mathrm{f}}) \xrightarrow{(\mathrm{QF}\circ\mathrm{PR})^{\omega}} \mathrm{Mod}(\Theta_{\mathrm{acf}})$$

Fig. 4 $\Theta^{\forall}_{\mathrm{acf}}$ has algebraically prime models with respect to Θ_{acf}

that associates to each integral domain in $\mathrm{Mod}(\Theta^{\forall}_{\mathrm{acf}})$ its field of fractions (in $\mathrm{Mod}(\Theta_{\mathrm{f}})$) (see [8]). Then, it is not difficult to show that FF is a left adjoint of J, as is stated in the next proposition.

Proposition 4.2 *Functor* FF *is left adjoint of the inclusion functor* J.

So, we can use Theorem 4.1 to conclude the following corollary.

Theorem 4.3 *Theory* $\Theta^{\forall}_{\mathrm{acf}}$ *has algebraically prime models with respect to* Θ_{f}.

It remains to show that the two forms of proving that a theory has algebraically prime models with respect to another theory can be composed.

Proposition 4.4 *Let* Δ, Ω *and* Υ *be first-order theories such that*:

- Δ *has algebraically prime models with respect to* Ω;
- Ω *has algebraically prime models with respect to* Υ.

Then Δ *has algebraically prime models with respect to* Υ.

Proof Assume that every model of Δ has an algebraically prime model in Ω and that every model of Ω has an algebraically prime model in Υ. Let I be a model of Δ. Denote by I° a model of Ω and by $\eta^\circ_I : I \to I^\circ$ an embedding such that I° is algebraically prime with respect to I via η°_I. Since Ω has algebraically prime models with respect to Υ, I° has an algebraically prime model in Υ. Denote by $\bar{I}$ a model of Υ and by $\bar{\eta}_{I^\circ} : I^\circ \to \bar{I}$ an embedding such that $\bar{I}$ is algebraically prime with respect to I° via $\bar{\eta}_{I^\circ}$. We now show that $\bar{I}$ is algebraically prime with respect to I via $\bar{\eta}_{I^\circ} \circ \eta^\circ_I$. Indeed: let I' be a model of Υ and $g : I \to I'$ an embedding. Since $\mathrm{Mod}(\Upsilon) \subseteq \mathrm{Mod}(\Omega)$ we have that I' is a model of Ω and so, taking into account that I° is algebraically prime with respect to I via η°_I, let $g^\circ : I^\circ \to I'$ be an embedding such that $g = g^\circ \circ \eta^\circ_I$. Similarly, since $\bar{I}$ is algebraically prime with respect to I° via $\bar{\eta}_{I^\circ}$, let $\bar{g}^\circ : \bar{I} \to I'$ be an embedding such that $g^\circ = \bar{g}^\circ \circ \bar{\eta}_{I^\circ}$. Hence $g = (\bar{g}^\circ \circ \bar{\eta}_{I^\circ}) \circ \eta^\circ_I$, that is, $g = \bar{g}^\circ \circ (\bar{\eta}_{I^\circ} \circ \eta^\circ_I)$. Therefore, there is an embedding from $\bar{I}$ to I' whose composition with $\bar{\eta}_{I^\circ} \circ \eta^\circ_I$ is g. □

Algebraically Closed Fields—From $\Theta^{\forall}_{\mathrm{acf}}$ to Θ_{acf} We omit the proof of the following theorem since it follows immediately by Theorem 4.3, Theorem 3.12, and by Proposition 4.4 (see Fig. 4).

Theorem 4.5 *The first-order theory* Θ_{acf} *has algebraically prime models.*

5 Adequacy for ∃

Local satisfaction of $\exists x \varphi$ formulas, where φ is a quantifier-free formula, is preserved by embeddings. However, reflection of local satisfaction of those formulas by embeddings does not hold, in general. A theory is *1-model-complete* whenever this property holds, that is,

$$I'h \circ \rho \Vdash_\Sigma \exists x \varphi \quad \text{implies } I\rho \Vdash_\Sigma \exists x \varphi$$

given an embedding $h : I \to I'$ and an assignment $\rho : X \to |I|$ over I.[5] We now provide a condition ("adequacy for ∃") sufficient for that reflection to hold. This condition tries to provide an answer to the following problem: What can be done when one wants to prove that a theory is 1-model-complete? The idea was to abstract the common aspects of proofs of reflection for several theories. As we detail below, the proposal consists in investigating the literals that really matter for the theory and variables at hand and trying to show that their satisfaction is reflected. As we will see below, those literals are the ones that, when satisfied by a model of the theory, are not equivalent to literals without those variables.

Care must be taken with the variables that may appear existentially quantified when proving the reflection of satisfaction. So, we consider not a set of literals, but a family Ω^e of sets of literals indexed by the finite sets of variables that can be existentially quantified (those are the sets in the family X^e below).

In the sequel, we denote by X the set of all variables and by L the set of all finite non-empty sets of literals. Given a theory Θ and a variable x, a pair of families

- $X^e = \{X^{\bar{\Lambda}}\}_{\Lambda \in L}$ where $X^{\bar{\Lambda}}$ is finite and $\{x\} \subseteq X^{\bar{\Lambda}} \subseteq (X \setminus \text{Vars}(\Lambda)) \cup \{x\}$;
- $\{\Omega(X^{\bar{\Lambda}}) : \Omega(X^{\bar{\Lambda}}) \text{ is a set of literals}\}_{X^{\bar{\Lambda}} \text{ in } X^e}$

is said to be Θ *exhaustive* for ∃ and x, whenever for every finite set Λ of literals there are finite sets $\Lambda_1, \ldots, \Lambda_n$ with literals in $\Omega(X^{\bar{\Lambda}})$ such that

$$\Theta \vDash_\Sigma \left(\exists x \bigwedge \Lambda\right) \Leftrightarrow \left(\bigvee_{i=1}^{n} \exists x_1 \ldots \exists x_m \bigwedge \Lambda_i\right)$$

where $\{x_1, \ldots, x_m\} = X^{\bar{\Lambda}}$. Moreover, we say that a set of variables X_1 is Θ *essential in* a literal ν *with respect to* a set of literals Ω_1 whenever for every μ in Ω_1 if $\Theta \vDash_\Sigma \mu \Leftrightarrow \nu$ then a variable of X_1 occurs in μ.

A theory Θ is *adequate for* ∃ whenever there are a variable x, a Θ exhaustive pair (X^e, Ω^e) for ∃ and x, and a family $A^e = \{A(X^{\bar{\Lambda}}) : A(X^{\bar{\Lambda}}) \subseteq \Omega(X^{\bar{\Lambda}})\}_{X^{\bar{\Lambda}} \text{ in } X^e}$, such that, given

- An embedding $h : I \to I'$ in $\text{Mod}(\Theta)$;
- Finite and non-empty sets Λ and C of literals with $C \subseteq \Omega(X^{\bar{\Lambda}})$;
- Assignment ρ' over I' such that $\rho'(z)$ is in $h(D)$ for every variable z not in $X^{\bar{\Lambda}}$, and $\rho'(z)$ is in $D' \setminus h(D)$ for some variable z in $X^{\bar{\Lambda}}$,

the following holds:

- For every ν in $C \setminus A(X^{\bar{\Lambda}})$, if $I'\rho' \Vdash \nu$ then $X^{\bar{\Lambda}}$ is not Θ essential in ν with respect to $\Omega(X^{\bar{\Lambda}})$;

[5] Given an embedding $h : I \to I'$ and an assignment $\rho : X \to |I|$ over I, observe that $h \circ \rho : X \to |I'|$ is an assignment of values of I' to the variables of X. So $I'h \circ \rho \Vdash \exists x \varphi$ means that formula $\exists x \varphi$ is satisfied by model I' and assignment $h \circ \rho$ over I'.

- There is an assignment σ over I with $h \circ \sigma \equiv_{X^{\bar{\Lambda}}} \rho'$ such that for every ν in $C \cap A(X^{\bar{\Lambda}})$, if $I'\rho' \Vdash_{\Sigma} \nu$ then $I\sigma \Vdash_{\Sigma} \nu$.

In this case, Θ is said to be *adequate for* $\exists$ *and* x *with respect to* (X^e, Ω^e) *and* A^e.

Observe that the family A^e identifies for each finite set $X^{\bar{\Lambda}}$ of variables in X^e, a subset of the exhaustive set $\Omega(X^{\bar{\Lambda}})$ of literals associated with those variables containing the literals in which there is one such variable that is really essential. So, A^e contains the literals whose satisfaction really need to be preserved.

Proposition 5.1 *Every theory adequate for* $\exists$ *is* 1*-model-complete.*

Proof Let Θ be a theory over a signature Σ adequate for $\exists$, x a variable and X^e, Ω^e and A^e families such that Θ is adequate for $\exists$ and x with respect to (X^e, Ω^e) and A^e. Moreover, let I and I' be models of Θ, $h : I \to I'$ an embedding, ρ an assignment over I, and Λ a finite set of literals. Suppose that $I'h \circ \rho \Vdash_{\Sigma} \exists x \bigwedge \Lambda$.

Since (X^e, Ω^e) is Θ exhaustive for x, let $\Lambda_1, \dots, \Lambda_n$ be finite sets of literals in $\Omega(X_{\bar{\Lambda}})$ where a variable of $X_{\bar{\Lambda}}$ only occurs if $X_{\bar{\Lambda}}$ is Θ essential in the literal in question for $\Omega(X_{\bar{\Lambda}})$, such that

$$I' h \circ \rho \Vdash_{\Sigma} \left(\bigvee_{i=1}^{n} \exists x_1 \dots \exists x_m \bigwedge \Lambda_i \right)$$

where $\{x_1, \dots, x_m\} = X_{\bar{\Lambda}}$. Let i be such that

$$I' h \circ \rho \Vdash_{\Sigma} \exists x_1 \dots \exists x_m \bigwedge \Lambda_i .$$

We want to show that $I\rho \Vdash_{\Sigma} \exists x_1 \dots \exists x_m \bigwedge \Lambda_i$. Let ρ' be an assignment over I' such that $\rho' \equiv_{X_{\bar{\Lambda}}} h \circ \rho$ and $I'\rho' \Vdash_{\Sigma} \bigwedge \Lambda_i$. Consider two cases:

1. $\rho'(x_j) \in h(D)$ for every $j = 1, \dots, m$. Let σ be such that $\sigma \equiv_{X_{\bar{\Lambda}}} \rho$ and $h \circ \sigma = \rho'$. Hence $I'h \circ \sigma \Vdash_{\Sigma} \bigwedge \Lambda_i$. Thus $I\sigma \Vdash_{\Sigma} \bigwedge \Lambda_i$. Therefore, $I\rho \Vdash_{\Sigma} \exists x_1 \dots \exists x_m \bigwedge \Lambda_i$ as we wanted to show.
2. $\rho'(x_j)$ is in $D' \setminus h(D)$ for some j in $\{1, \dots, m\}$. Observe that the image by ρ' of every variable in Λ_i not in $X^{\bar{\Lambda}}$ is in $h(D)$ since $\rho' \equiv_{X_{\bar{\Lambda}}} h \circ \rho$. Hence, since Θ is adequate for $\exists$ and x with respect to (X^e, Ω^e) and A^e, let σ be an assignment over I such that $h \circ \sigma \equiv_{X_{\bar{\Lambda}}} \rho'$ and $I\sigma \Vdash_{\Sigma} \nu$ whenever $I'\rho' \Vdash_{\Sigma} \nu$ for every $\nu \in \Lambda_i \cap A(X^{\bar{\Lambda}})$. We now show that $I\sigma \Vdash_{\Sigma} \bigwedge \Lambda_i$. So let ν be a literal of Λ_i. Recall that $I'\rho' \Vdash_{\Sigma} \bigwedge \Lambda_i$. Consider two cases:
   (a) $\nu \in A(X^{\bar{\Lambda}})$. Then $\nu \in \Lambda_i \cap A(X^{\bar{\Lambda}})$ and so $I\sigma \Vdash_{\Sigma} \nu$ by hypothesis since $I'\rho' \Vdash_{\Sigma} \bigwedge \Lambda_i$ and ν is in Λ_i.
   (b) $\nu \notin A(X^{\bar{\Lambda}})$. Then, $X^{\bar{\Lambda}}$ is not Θ essential in ν for $\Omega(X^{\bar{\Lambda}})$ since $I'\rho' \Vdash_{\Sigma} \nu$. Hence no variable of $X^{\bar{\Lambda}}$ occurs in ν since by hypothesis a variable of $X^{\bar{\Lambda}}$ only occurs in literals of Λ_i when $X^{\bar{\Lambda}}$ is essential. So $I'h \circ \sigma \Vdash_{\Sigma} \nu$ since $I'\rho' \Vdash_{\Sigma} \nu$ and $(h \circ \sigma)(z) = \rho'(z)$ for every variable z not in $X^{\bar{\Lambda}}$ and no variable of $X^{\bar{\Lambda}}$ occurs in ν. Hence $I\sigma \Vdash_{\Sigma} \nu$ taking into account that ν is a literal.

So $I\sigma \Vdash_\Sigma \bigwedge \Lambda_i$. Observe that $h \circ \sigma \equiv_{X^{\bar{\Lambda}}} h \circ \rho$ and so $\sigma \equiv_{X^{\bar{\Lambda}}} \rho$. Therefore,

$$I\rho \Vdash_\Sigma \exists x_1 \dots \exists x_m \bigwedge \Lambda_i .$$

Thus

$$I\rho \Vdash_\Sigma \left(\bigvee_{i=1}^{n} \exists x_1 \dots \exists x_m \bigwedge \Lambda_i \right)$$

and so

$$I\rho \Vdash_\Sigma \exists x \varphi$$

as we wanted to show. □

Natural Numbers with Successor The theory of natural numbers with successor is adequate for ∃, as we show below, and so, by Proposition 5.1, also 1-model-complete. Recall the first-order theory Θ_{S} for term algebras described in Sect. 3. Given a variable x and a finite and non-empty set Λ of literals over Σ_{S}, let

- $X_{\mathrm{S}}^{\bar{\Lambda}}$ be the set $\{x\}$;
- $\Omega_{\mathrm{S}}(\{x\})$ be the set of all literals over the atomic formulas of the form $(\mathbf{S}^m\, t) \cong t'$ where t and t' are either 0 or a variable, and m is a natural number,

and denote by X_{S} the family $\{X_{\mathrm{S}}^{\bar{\Lambda}}\}_{\Lambda \in L_{\mathrm{S}}}$ and by Ω_{S} the family $\{\Omega_{\mathrm{S}}(X_{\mathrm{S}}^{\bar{\Lambda}})\}_{X_{\mathrm{S}}^{\bar{\Lambda}} \in X_{\mathrm{S}}}$. Then the following proposition immediately follows:

Proposition 5.2 *The pair $(X_{\mathrm{S}}, \Omega_{\mathrm{S}})$ is Θ_{S} exhaustive for ∃ and x.*

Denote by A_{S} the family

$$\left\{A_{\mathrm{S}}\left(X_{\mathrm{S}}^{\bar{\Lambda}}\right)\right\}_{X_{\mathrm{S}}^{\bar{\Lambda}} \text{ in } X_{\mathrm{S}}}$$

where $A_{\mathrm{S}}(X_{\mathrm{S}}^{\bar{\Lambda}})$ is the set of literals of $\Omega_{\mathrm{S}}(X_{\mathrm{S}}^{\bar{\Lambda}})$ of the form $\neg((\mathbf{S}^m\, t) \cong t')$ where m is a natural number greater than zero, t and t' are either 0 or a variable, and either t or t' is x but not both.

Proposition 5.3 *The theory Θ_{S} is adequate for ∃ and x with respect to $(X_{\mathrm{S}}, \Omega_{\mathrm{S}})$ and A_{S}.*

Proof Let I and I' be models of Θ_{S}, $h : I \to I'$ an embedding, Λ a finite non-empty set of literals, C a finite non-empty set of literals in $\Omega_{\mathrm{S}}(X_{\mathrm{S}}^{\bar{\Lambda}})$ and ρ' an assignment over I' such that $\rho'(z)$ is in $h(D)$ for every variable z not in $X_{\mathrm{S}}^{\bar{\Lambda}} = \{x\}$, and $\rho'(x)$ is in $D' \setminus h(D)$.

1. Let ν be a literal in $C \setminus A_{\mathrm{S}}(X_{\mathrm{S}}^{\bar{\Lambda}})$ such that $I'\rho' \Vdash \nu$. Then ν is either of the form $(\mathbf{S}^m\, t) \cong t'$ or $\neg(t \cong t')$ where t and t' are either 0 or a variable and m is a natural number, or of the form $\neg((\mathbf{S}^m\, t) \cong t')$ where m is a natural number greater than 0, t and t' are either 0 or a variable, and either both t and t' are x or both are not x. The proof proceeds by case analysis:

(a) m is 0 or ν is $\neg(t \cong t')$. Assume that ν is of the form $t \cong t'$. We have two cases:

  (i) t is t'. Then ν is equivalent to $0 \cong 0$ and so $X_S^{\bar{\Lambda}}$ is not Θ_S essential in ν for $\Omega_S(X_S^{\bar{\Lambda}})$.

  (ii) t is not t'. Then both t and t' are not x since $\rho'(x) \in D' \setminus h(D)$ and $0^{F'}$ and $\rho'(z) \in h(D)$ for every z different from x. Then $X_S^{\bar{\Lambda}}$ is not Θ_S essential in ν for $\Omega_S(X_S^{\bar{\Lambda}})$. We omit the proof when ν is $\neg(t \cong t')$ since it is similar to the case just proved.

(b) m is not 0 and t is t'. Assume that ν is of the form $(\mathbf{S}^m\, t) \cong t$. This case is not possible due to axiom S4. Assume now that ν is of the form $\neg((\mathbf{S}^m\, t) \cong t)$. Then ν is equivalent to $0 \cong 0$ and so $X_S^{\bar{\Lambda}}$ is not Θ_S essential in ν for $\Omega_S(X_S^{\bar{\Lambda}})$.

(c) m is not 0, t and t' are distinct and both are not x. Then $X_S^{\bar{\Lambda}}$ is not Θ_S essential in ν for $\Omega_S(X_S^{\bar{\Lambda}})$.

(d) ν is of the form $(\mathbf{S}^m\, t) \cong t'$, m is not 0, t and t' are distinct and either t or t' is x. This case is not possible since it implies that $\rho'(x) \in h(D)$, and we are assuming that $\rho'(x) \in D' \setminus h(D)$. We now detail the proof when $m = 1$. Assume, by contradiction that $\mathbf{S}^{F'}(\rho'(x)) = h(d)$ for some $d \in D$. Consider two cases:

  (i) $d = 0^F$. This case contradicts axiom S1.

  (ii) d is not 0^F. Let d_1 be such that $\mathbf{S}^F(d_1) = d$. Then $\mathbf{S}^{F'}(h(d_1)) = h(\mathbf{S}^F(d_1)) = h(d) = \mathbf{S}^{F'}(\rho'(x))$. Thus, by axiom S2, $h(d_1) = \rho'(x)$ which contradicts the hypothesis.

2. Without loss of generality, denote the elements of $C \cap A_S(X_S^{\bar{\Lambda}})$ by $\neg((\mathbf{S}^{m_1}\, t_1) \cong t_1'), \ldots, \neg((\mathbf{S}^{m_n}\, t_n) \cong t_n')$. Assume that $I'\rho' \Vdash_{\Sigma_S} \nu$ for every $\nu \in C \cap A_S(X_S^{\bar{\Lambda}})$. Let σ be an assignment over I such that $h \circ \sigma \equiv_x \rho'$ and such that $\sigma(x)$ satisfies the constraints imposed by those inequations. There is such a value in D since D is infinite and the set of values that $\sigma(x)$ should not take according to those inequations is finite. Then $I\sigma \Vdash_{\Sigma_S} \nu$ for every $\nu \in C \cap A_S(X_S^{\bar{\Lambda}})$ by definition of σ. ☐

So, capitalizing on the previous proposition and on Proposition 5.1, we can establish the following theorem.

Theorem 5.4 *Theory Θ_S is 1-model-complete.*

Term Algebras Recall the first-order theory Θ_{ta} for term algebras described in Sect. 3 over the signature Σ_{ta} induced by a signature Σ with no predicate symbols besides equality. Given a variable x and a finite and non-empty set Λ of literals over Σ_{ta},

- Let $X_{ta}^{\bar{\Lambda}}$ be a set $\{x, x_1, \ldots, x_m\}$ where $x_1, \ldots, x_m$ are variables not occurring in Λ and m in $\mathbb{N}$ is greater than the product of the maximum depth of a term in Λ containing x with the greatest arity of a function in Σ;
- Let $\Omega_{ta}(X_{ta}^{\bar{\Lambda}})$ be the set of all literals of Σ_{ta} with no variables in $X_{ta}^{\bar{\Lambda}}$ together with the literals of the following form:
  - $\neg(z \cong t)$ where z is in $X_{ta}^{\bar{\Lambda}}$ and t is either a variable of $X_{ta}^{\bar{\Lambda}}$ distinct of z or has no variables of $X_{ta}^{\bar{\Lambda}}$;

- $\mathbf{Is}_c(z)$, $\neg\,\mathbf{Is}_c(z)$ for every z in $X_{\mathrm{ta}}^{\bar\Lambda}$ and constant c in Σ;
- $\mathbf{Is}_f(z)$, $\neg\,\mathbf{Is}_f(z)$ for every z in $X_{\mathrm{ta}}^{\bar\Lambda}$ and function symbol f in Σ.

Denote by X_{ta} the family $\{X_{\mathrm{ta}}^{\bar\Lambda}\}_{\Lambda\in L_{\mathrm{ta}}}$ and by Ω_{ta} the family $\{\Omega(X_{\mathrm{ta}}^{\bar\Lambda})\}_{X_{\mathrm{ta}}^{\bar\Lambda}\text{ in }X_{\mathrm{ta}}}$. Then the following proposition, Proposition 5.5, follows immediately from the first steps of the proof of Theorem 2.7.5 of [7].

Proposition 5.5 *The pair* $(X_{\mathrm{ta}}, \Omega_{\mathrm{ta}})$ *is* Θ_{ta} *exhaustive for* $\exists$ *and* x.

Denote by A_{ta} the family $\{A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})\}_{X_{\mathrm{ta}}^{\bar\Lambda}\text{ in }X_{\mathrm{ta}}}$ where $A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ is the set of literals of $\Omega_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ with a variable of $X_{\mathrm{ta}}^{\bar\Lambda}$.

Proposition 5.6 *The theory* Θ_{ta} *is adequate for* $\exists$ *and* x *with respect to* $(X_{\mathrm{ta}}, \Omega_{\mathrm{ta}})$ *and* A_{ta}, *provided that* Σ *has an infinite number of constants.*

Proof Let I and I' be models of Θ_{ta}, $h : I \to I'$ an embedding, Λ a finite non-empty set of literals, C a finite non-empty set of literals in $\Omega_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ and ρ' an assignment over I' such that $\rho'(z)$ is in $h(D)$ for every variable z not in $X^{\bar\Lambda}$, and $\rho'(z)$ is in $D' \setminus h(D)$ for some variable z in $X^{\bar\Lambda}$.

1. Let ν be a literal in $C \setminus A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ such that $I'\rho' \Vdash \nu$. Then, since $C \subseteq \Omega_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$, by definition of $A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ no variable of $X_{\mathrm{ta}}^{\bar\Lambda}$ occurs in ν. So $X_{\mathrm{ta}}^{\bar\Lambda}$ is not Θ_{ta} essential in ν for $\Omega_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$.
2. Assume that $I'\rho' \Vdash_{\Sigma_{\mathrm{ta}}} \nu$ for every $\nu \in C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$. Let σ be an assignment over I such that $(h \circ \sigma)(z) = \rho'(z)$ for every variable z with $\rho'(z) \in h(D)$ and such that if $\rho'(z) \notin h(D)$ then

   - If $\mathbf{Is}_f(z)$ is in $C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ for some function symbol f in Σ then $\sigma(z)$ is an element in D satisfying the predicate $\mathbf{Is}_f$ and different of $[\![t]\!]^{I\sigma}$ for every $\neg(z \cong t)$ in $C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$. Observe that there is such an element in D since Σ has an infinite number of constants by assumption and so by axioms T2, T3 and T4 there is an infinite number of elements of D satisfying the predicate $\mathbf{Is}_f$, and since the number of elements of the domain to which $\sigma(z)$ must be different is finite because $C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ has a finite number of literals;
   - If $\mathbf{Is}_f(z)$ is not in $C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ for every function symbol f in Σ then let $\sigma(z)$ be $[\![c']\!]^I$ for some constant c' in Σ, different of $[\![c]\!]^I$ for every $\neg\,\mathbf{Is}_c(z)$ in $C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$, and different of $[\![t]\!]^{I\sigma}$ for every $\neg(z \cong t)$ in $C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$. Observe that there is such an element in D since Σ has an infinite number of constants by assumption, which are interpreted as different elements of the domain by axiom T4, and since $C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$ has a finite number of literals.

Hence $I\sigma \Vdash_{\Sigma_{\mathrm{ta}}} \nu$ for every literal ν of $C \cap A_{\mathrm{ta}}(X_{\mathrm{ta}}^{\bar\Lambda})$, by definition of σ. □

So, capitalizing on the previous proposition and on Proposition 5.1, we can now establish the following theorem.

Theorem 5.7 *Theory Θ_{ta} is 1-model-complete.*

Algebraically Closed Fields We now prove that Θ_{acf} is adequate for $\exists$, and so, by Proposition 5.1, that Θ_{acf} is 1-model-complete. Given a variable x and a finite and non-empty set Λ of literals over Σ_{f}, let

- $X^{\bar{\Lambda}}_{\text{acf}}$ be the set $\{x\}$;
- $\Omega_{\text{acf}}(\{x\})$ be the set of all literals over the atomic formulas of the form

$$q(x_1, \ldots, x_n, x) \cong 0$$

for some natural number n and distinct variables $x_1, \ldots, x_n, x$.

Denote by X_{acf} and Ω_{acf} the families of $\{X^{\bar{\Lambda}}_{\text{acf}}\}_{\Lambda \in L_{\text{acf}}}$ and $\{\Omega_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})\}_{X^{\bar{\Lambda}}_{\text{acf}} \text{ in } X_{\text{acf}}}$, respectively. Then the following proposition, Proposition 5.8, follows immediately.

Proposition 5.8 *The pair $(X_{\text{acf}}, \Omega_{\text{acf}})$ is Θ_{acf} exhaustive for $\exists$ and x.*

Denote by A_{acf} the family $\{A_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})\}_{X^{\bar{\Lambda}}_{\text{acf}} \text{ in } X_{\text{acf}}}$ where $A_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})$ is the set of literals of $\Omega_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})$ whose main connective is negation.

Proposition 5.9 *The theory Θ_{acf} is adequate for $\exists$ and x with respect to $(X_{\text{acf}}, \Omega_{\text{acf}})$ and A_{acf}.*

Proof Let I and I' be models of Θ_{acf}, $h : I \to I'$ an embedding, Λ a finite non-empty set of literals, C a finite non-empty set of literals in $\Omega_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})$ and ρ' an assignment over I' such that $\rho'(z)$ is in $h(D)$ for every variable z not in $X^{\bar{\Lambda}}_{\text{acf}} = \{x\}$, and $\rho'(x)$ is in $D' \setminus h(D)$. Then:

1. Let ν be a literal in $C \setminus A_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})$. Assume that $I'\rho' \Vdash \nu$. Then, for some natural number n and distinct variables $x_1, \ldots, x_n, x$, ν is of the form $q(x_1, \ldots, x_n, x) \cong 0$. Assume by contradiction that $X^{\bar{\Lambda}}_{\text{acf}}$ is Θ_{acf} essential in ν for $\Omega_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})$. Observe that $\rho'(x)$ is a solution of the polynomial equation
$$q^{F'}\big(\rho'(x_1), \ldots, \rho'(x_n), x\big) = 0^{F'}.$$
Let ρ be an assignment over I such that $\rho(x_i) = h^{-1}(\rho'(x_i))$ for $i = 1, \ldots, n$ and let m be the number of roots of the equation $q^F(\rho(x_1), \ldots, \rho(x_n), x) = 0^F$ in D and $d_1, \ldots, d_m$ those roots. Note that $h(d_1), \ldots, h(d_m)$ are also the m roots in D' of the equation $q^{F'}(h(\rho(x_1)), \ldots, h(\rho(x_n)), x) = 0^{F'}$ that is of the equation $q^{F'}(\rho'(x_1), \ldots, \rho'(x_n), x) = 0^{F'}$. So $\rho'(x) = h(d_j)$ for some j in $\{1, \ldots, m\}$ which contradicts the fact that $\rho'(x) \in D' \setminus h(D)$.
2. Assume with no loss of generality that $C \cap A_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})$ contains the literals $\neg(q_1(x_{11}, \ldots, x_{1n_1}, x) \cong 0), \ldots, \neg(q_k(x_{k1}, \ldots, x_{kn_k}, x) \cong 0)$. Let σ be an assignment over I such that $h \circ \sigma \equiv_x \rho'$ and $\sigma(x)$ is not a root of the polynomial equation $q_j^F(h^{-1}(\rho'(x_{j1})), \ldots, h^{-1}(\rho'(x_{jn_j})), x) = 0^F$ for all $j = 1, \ldots, k$. There is such a value in D since D is infinite and the number of such roots is finite. Then $I\sigma \Vdash_{\Sigma_{\text{f}}} \nu$ for every $\nu \in C \cap A_{\text{acf}}(X^{\bar{\Lambda}}_{\text{acf}})$ by definition of σ. □

So, capitalizing on the previous proposition and on Proposition 5.1, we can establish the following theorem.

Theorem 5.10 *Theory* Θ_{acf} *is* 1*-model-complete.*

6 Concluding Remarks

The importance of decidability of first-order theories is well recognized in computer science applications. In many cases, proving decidability involves proving that the theory at hand has quantifier elimination. Quantifier elimination can be proved constructively using symbolic methods or non-constructively using semantic methods. The objective of this paper is to contribute to a toolbox making easier to prove, in a semantic way, that a theory has quantifier elimination. We believe that this step is useful when we still do not know if a theory has quantifier elimination. The work capitalizes on a well known result stating that a theory with algebraically prime models and 1-complete has quantifier elimination.

Proving that a theory Θ has algebraically prime models with respect to another theory Υ involves providing a map F from the models of Θ to the models of Υ, an embedding from each model I of Θ to $F(I)$ and involves verifying that a certain universal property holds. In this paper, we state that such a proof can be obtained either by iteration or by adjunction or by a combination of both. We also provide a sufficient condition for a theory to be 1-model-complete. This sufficient condition is targeted to identify the key literals for the theory at hand and to prove the reflection of satisfaction of these literals through embeddings. We illustrate the techniques on the theories of natural numbers with successor, term algebras and algebraically closed fields.

When applying our results to show that a theory has algebraically prime models with respect to another theory, one may wonder which method to use: the iterative or the category-theoretic. In order to answer this question, observe that the iterative approach is more general than the category-theoretic approach in the sense that the relationship considered between the categories of models of the theories and their embeddings does not need to be a functor, as it should be in the category-theoretic approach. For example, the relationship used in the proof that Θ_{f} has algebraically prime models with respect to Θ_{acf}, defined by the map $\text{QF} \circ \text{PR} : \text{Mod}(\Theta_{\text{f}}) \to \text{Mod}(\Theta_{\text{f}})$, is not a functor, which means that in this case the category-theoretic approach could not be used.

There are several ways to go on with this work. The most interesting and perhaps important aspect is to obtain results of preservation of quantifier elimination when adding axioms to a given theory.

Acknowledgements This work was partially supported, under the MCL (Meet-Combination of Logics) and PQDR (Probabilistic, Quantum and Differential Reasoning) initiatives of SQIG at IT, by FCT and EU FEDER, namely via the projects FCT PEst-OE/EEI/LA0008/2013 and AMDSC UTAustin/MAT/0057/2008, as well as by the European Union's Seventh Framework Programme for Research (FP7), namely through project LANDAUER (GA 318287).

References

1. Bradley, A.R., Manna, Z.: The Calculus of Computation. Springer, Berlin (2007)

2. Delahaye, D., Mayero, M.: Quantifier elimination over algebraically closed fields in a proof assistant using a computer algebra system. Electron. Notes Theor. Comput. Sci. **151**(1) (2006)
3. Enderton, H.B.: A Mathematical Introduction to Logic, 2nd edn. Harcourt/Academic Press, New York (2001)
4. Fagin, R., Halpern, J.Y., Megiddo, N.: A logic for reasoning about probabilities. Inf. Comput. **87**(1–2), 78–128 (1990)
5. Gabbay, D.M., Szałas, A.: Second-order quantifier elimination in higher-order contexts with applications to the semantical analysis of conditionals. Stud. Log. **87**(1), 37–50 (2007)
6. Gödel, K.: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatshefte Math. Phys. **38**(1), 173–198 (1931)
7. Hodges, W.: Model Theory. Encyclopedia of Mathematics and Its Applications, vol. 42. Cambridge University Press, Cambridge (1993)
8. Lang, S.: Algebra, 3rd edn. Graduate Texts in Mathematics, vol. 211. Springer, Berlin (2002)
9. Marker, D.: Model Theory: An Introduction. Graduate Texts in Mathematics, vol. 217. Springer, Berlin (2002)
10. Monniaux, D.: Quantifier elimination by lazy model enumeration. In: Computer Aided Verification. Lecture Notes in Computer Science, vol. 6174, pp. 585–599. Springer, Berlin (2010)
11. Nipkow, T.: Linear quantifier elimination. J. Autom. Reason. **45**(2), 189–212 (2010)
12. Platzer, A.: Differential-algebraic dynamic logic for differential-algebraic programs. J. Log. Comput. **20**(1), 309–352 (2010)
13. Sturm, T., Tiwari, A.: Verification and synthesis using real quantifier elimination. In: ISSAC 2011—Proceedings of the 36th International Symposium on Symbolic and Algebraic Computation, pp. 329–336. ACM, New York (2011)
14. Tarski, A.: A Decision Method for Elementary Algebra and Geometry, 2nd edn. University of California Press, Berkeley (1951)
15. Weispfenning, V.: Simulation and optimization by quantifier elimination. J. Symb. Comput. **24**(2), 189–208 (1997)

J. Rasga · C. Sernadas (✉) · A. Sernadas
Dep. Mathematics, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal
e-mail: cristina.sernadas@tecnico.ulisboa.pt

J. Rasga
e-mail: joao.rasga@tecnico.ulisboa.pt

A. Sernadas
e-mail: amilcar.sernadas@tecnico.ulisboa.pt

J. Rasga · C. Sernadas · A. Sernadas
SQIG, Instituto de Telecomunicações, Lisboa, Portugal

John Buridan on Non-contingency Syllogisms

Stephen Read

Abstract Whereas most of his predecessors attempted to make sense of, and if necessary correct, Aristotle's theory of the modal syllogism, John Buridan starts afresh in his *Treatise on Consequences*, treating separately of composite and divided modals, then of syllogisms of necessity, possibility, and with mixed premises. Finally, he comes in the penultimate chapter of the treatise, Book IV Chap. 3, to present a concise treatment of syllogisms with premises of contingency, that is, two-sided possibility. The previous modal syllogisms had all been taken with an affirmed mode only, since modal conversion equates negated necessity and possibility with affirmed possibility and necessity, respectively. But in his Conclusions concerning syllogisms of contingency, he also treats those with negated mode. These are the non-contingency syllogisms.

Keywords Necessity · Possibility · Contingency · Syllogism · Buridan

Mathematics Subject Classification (2000) 03A05 · 01A35

1 Necessity and Possibility

As is well known, much of the work on modal logic in the 1600 years after Aristotle's death was not only determined by the great strides he had made in its creation but also by the attempt to make sense of Aristotle's to some extent puzzling analysis and to find a coherent account of modality consistent with it. John Buridan seems to have been one of the first logicians to move on from this attempt and to create his own system of modalities and its foundation. In particular, he abandoned the attempt to provide an account which validates Barbara LXL while invalidating Barbara XLL.[1] For Buridan, both are invalid. But this is not a rejection of Aristotle's view, since arguably Buridan's interpretation of propositions of necessity is different from Aristotle's.

Where Aristotle provides only a brief analysis of modal propositions (in *De Interpretatione* 12–14 and *Prior Analytics* I 3), Buridan devotes the whole of Book II of his

[1]Barbara is a mnemonic for the AAA syllogism in the first figure. See, e.g., Aristotle [1, pp. 67–71], Buridan [2, §5.2.2, pp. 320–1]. L, M, X and Q stand respectively for the modalities necessity, possibility, assertoric and contingency. For the full mnemonic, and discussion of the problem of the two Barbaras, see, e.g., Thom [8].

I am very happy to dedicate this paper to Jean-Yves Béziau.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_21

Treatise on Consequences to a general analysis of modality before embarking in Book IV on his own account of the modal syllogism. He follows Aristotle in restricting attention to modalisations of assertoric syllogisms. He also uses two ideas which Aristotle introduced but which play little part in Aristotle's own analysis of modality. The first is the distinction between compounded and divided senses.[2] In applying this to modal propositions, Buridan seems to be following Abelard.[3] Most of Buridan's attention, and my own here, is on divided modal propositions.

At first glance, Aristotle seems to discuss only two modes, necessity and possibility. But this is to overlook a distinction Aristotle makes between two senses of possibility, one-sided and two-sided:

> "After these explanations, let us add that 'being possible' is said in two ways: in one way of what happens for the most part, when the necessity has gaps, such as that a man turns grey or grows or ages, or generally what belongs by nature. For this has no continuous necessity because a man does not exist forever, but while a man exists, it happens either of necessity or for the most part. In another way 'being possible' is said of what is indeterminate, that is, what is possible both this way and not this way, such as that an animal walks or that an earthquake happens while it walks, or, generally, what comes about by chance, for this is by nature no more this way than the opposite way." (*Prior Analytics* I 13, 32b4–14, tr. Gisela Striker in [1].)

That is, 'possible' may be simply the contradictory of 'impossible', in which case, what is necessary is also possible; but 'possible' may also be opposed to 'necessary', so that 'possible' describes not simply what is not impossible but what is also not necessary. The first sense of 'possible' is sometimes called 'one-sided possibility', the latter 'two-sided possibility'. Buridan himself in general reserves 'possible' for the first sense and 'contingent', or sometimes 'contingent each way' (*contingens ad utrumlibet*), for the latter, for if a proposition is contingent then things can be as it signifies and can also fail to be as it signifies.[4] What is contingent is not necessary, and so may possibly fail to obtain. Unfortunately, even though Aristotle has two words for 'possible' (*dunaton* and *endechesthai*), he uses both in both senses, often but not always noting whether he means possible in the weaker sense (not impossible) or stronger (neither impossible nor necessary). This equivocation runs right through Aristotle's discussion of the modal syllogism. As a matter of fact, he only takes 'possibility' premises in the stronger sense (contingency), but often considers possibility conclusions in the weaker sense.

The other idea which Buridan takes from Aristotle but unlike Aristotle extends throughout his account of modal propositions is that of ampliation of the subject. At *Prior Analytics* I 13, Aristotle suggests that the subject of propositions of contingency is ampliated to the contingent:

> "Given that 'this possibly belongs to that' may be understood in two ways—either of what that belongs to, or of what that may belong to (for '*A* possibly belongs to what *B* belongs to' signifies

---

[2] See *Sophistici Elenchi* 4, 166a24–26. "Amphiboly and ambiguity depend on these modes of speech. Upon the combination of words there depend instances such as the following: 'A man can walk while sitting, and can write while not writing'. For the meaning is not the same if one divides the words and if one combines them in saying that 'it is possible to walk-while-sitting'."

[3] See, e.g., Thom [9, p. 169].

[4] Buridan, *Summulae de Dialectica* [2, §1.8.5, p. 76]: "But as far as 'contingent' is concerned, we should realise that sometimes 'contingent' is taken broadly, and then it is synonymous with 'possible', and sometimes it is taken strictly, and this we call 'contingent both ways', and then it is a species of 'possible' distinguished from 'necessary'.".

> one or the other of these, either that *A* may belong to what *B* is said of, or that it may belong to what *B* may be said of)—and that there is no difference between '*A* possibly belongs to what *B* is said of' and '*A* possibly belongs to every *B*', it is evident that '*A* possibly belongs to every *B*' would be said in two ways." (32b25–30)

It seems clear that Aristotle means 'contingent' by 'possible' in this passage, so he is suggesting that the subject of contingency propositions is ampliated to the contingent. Buridan notes in the *Treatise on Consequences* [4, IV 3] [5, p. 130] that this is mistaken:

> "In this connection, moreover, it should be realised that in a proposition of contingency the subject is ampliated to supposit for those which are and for those which can be, and it is not required that the subject supposit for those which are contingently. For God is contingently creating, but nothing which is contingently God is contingently creating, because nothing is contingently God, indeed, everything is necessarily God or necessarily fails to be God."

Rather, the subject of propositions of contingency is ampliated to the possible, just as are the subjects of propositions of necessity and possibility.

That the subject of propositions of necessity is ampliated to the possible is crucial to Buridan's account, and forms the basis of his octagons of opposition.[5] Buridan's main discussion of modal syllogisms concerns only modal propositions of necessity and (one-sided) possibility (possibly mixed with assertorics). Applying these two modes to the four types of assertoric proposition yields eight modal types, in symbols: *La*, *Le*, *Li*, *Lo*, *Ma*, *Me*, *Mi* and *Mo*. Buridan's idea is that the duality of necessity and possibility, that is, that 'not necessary' is equivalent to 'possibly not' and so on, means that, e.g., *La* should be the contradictory of *Mo*, *Le* of *Mi*, and so on. But given that in possibility-propositions, the subject is naturally ampliated to the possible, that is, for example, 'Some *B* is possibly *A*' means that something which is or can be *B* can be *A*, the subject of the necessity-proposition which is its contradictory must similarly be ampliated to the possible. Note also that Buridan, like most other medievals, and arguably like Aristotle himself, took affirmatives to have existential import, negatives to be true if the subject is empty.[6]

I intend in what follows to make use of a concise formalisation introduced by Paul Thom, which used carefully allows both compact expression of each of the eight proposition types and also a compact method of proof. See Fig. 1.

Key

'Every *B* is *A*',	$\underline{b} \rightarrow a$ (inclusion);	$\underline{b}$,	indicates existential import;
'No *B* is *A*',	$b \vert a$ (exclusion);	$b^{\dagger}$,	what is *B* or possibly *B*;
'Some *B* is *A*',	$b \smile a$ (overlap);	b^{*},	what is necessarily *B*;
'Not every *B* is *A*',	$b \nrightarrow a$ (non-inclusion).		

[5] So-called by George Hughes [6]; Buridan himself calls them 'great figures'—*magnae figurae*.

[6] Buridan expands the universal affirmative to contain an explicit I-proposition in the 14th Conclusion of Book I of the *Treatise on Consequences* [5, p. 52] : " 'homo est animal et nullus homo est aliud ab animali' ... haec copulatiua aequiualent isti universali affirmatiuae 'omnis homo est animal'." So it is false if the subject is empty, just as its contradictory, the particular negative, is true in the same circumstances: "non sequitur 'chimaera non est homo; ergo non homo non est non chimaera', quia prima est uera et secunda falsa." Hubien [5, p. 53].

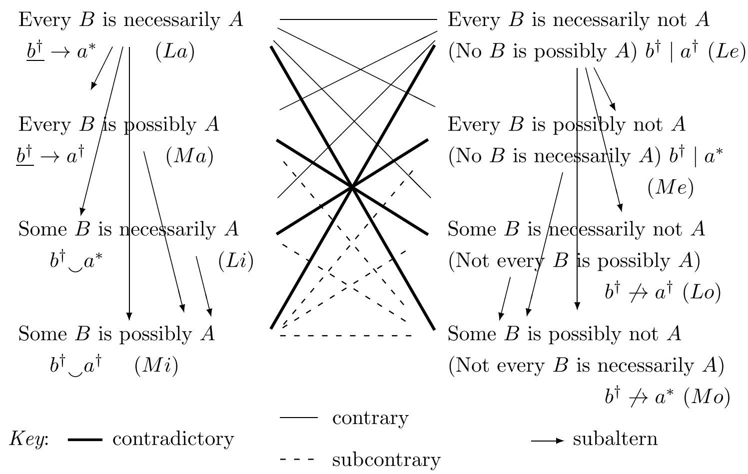


Fig. 1 Buridan's octagon in Thom's notation

Particular care is needed with Thom's use of underline to capture existential import.[7] Thom presents the following eight axioms,[8] which capture the idea that $\to$ means inclusion, $\smile$ indicates overlap and $\mid$ exclusion. Axioms 1.3 and 1.4 capture the *dictum de omni et nullo*:

1.1	If $\underline{b} \to a$ then $b \smile a$;	1.2	If $b \mid a$ then $a \mid b$;
1.3	If $c \to b \to a$ then $c \to a$;	1.4	If $c \to b \mid a$ then $c \mid a$;
1.5	$b \not\to a$ iff $\Sigma d, b \leftarrow d \mid a$;	1.6	$b \smile a$ iff $\Sigma d, b \leftarrow \underline{d} \to a$;
1.7	$a^* \to a$;	1.8	$a \to a^\dagger$.

Note that the terms a, b here may be empty terms, so we have substitutional quantification in 1.5 and 1.6, indicated by the unusual style of quantifier.[9] Only in 1.6 is d taken actually to exist, as indicated by the underline; the counter-instance in 1.5 may be an empty term. We need an additional principle which appears not to follow from Thom's axioms, viz:[10]

2.1 if $\underline{b} \to a$ then $b \to a$.

The natural expression of the negative modal propositions in Fig. 1 deserves comment. Take, e.g., *Le*-propositions: these apply the mode 'necessary' to the E-proposition 'No B is A', or equivalently, 'Every B fails to be A'. But since we are taking the divided sense of the modality, the mode has to be applied to the predicate, yielding 'Every B necessarily fails to be A', that is, 'Every B is not possibly A' (given the duality of 'necessary' and

[7]Note that, where $\underline{b} \to a$ adds to $b \to a$ the requirement that b is non-empty, $b \not\to a$ not only denies $b \to a$ but disjoins the possibility of emptiness of b, as shown by axiom 1.5 below.

[8]I've amended axioms 1.5 and 1.6 somewhat from Thom's formulation.

[9]Thom [9, p. 17] writes, e.g., "for some 'd', $b \leftarrow d \mid a$".

[10]Note that $\underline{b} \to a$ does not say that every existing B is A, but that every B is A *and* there are Bs.

'possible'), i.e., 'No B is possibly A'. Similarly, *Me*-propositions apply 'possibly' divisively to 'Every B fails to be A' yielding 'Every B possibly fails to be A', that is, 'Every B is not necessarily A', i.e., 'No B is necessarily A'. A similar somewhat unintuitive duality affects *Lo*- and *Mo*-propositions, so that the mode 'possibly' appears in the natural expression of *Lo*-propositions, 'Not every B is possibly A' and 'necessarily' in the natural expression of *Mo*-propositions, 'Not every B is necessarily A'. O-propositions are expressed here, following Aristotle, as the explicit contradictory of A-propositions, 'Not every B is A', rather than in the perhaps more familiar existential form, 'Some B is not A', to emphasise their lack of existential import.

2 Contingency and Non-contingency

In Chap. 3 of Book IV of his *Treatise on Consequences*, Buridan discusses syllogisms with premises and conclusions of contingency. Recall that 'is contingent' means 'is neither necessary nor impossible', that is, 'is possibly and possibly fails to be'. Thom represents the contingency of A by $a^{\ddagger}$, and it is immediate that

1.9 $a^{\ddagger} \rightarrow a^{\dagger}$,

as Buridan observes in Conclusion 8 of Book II, and that

1.10 $a^{\ddagger} \mid a^{*}$,

as he says in IV 3.

Aristotle claimed at *Prior Analytics* I 13 that

> "all possible [that is, contingent][11] premises convert to one another. I do not mean that affirmative ones convert to negatives, but that those that are affirmative in form convert with respect to opposites. So, for example, 'possibly belonging' converts to 'possibly not belonging', 'possibly belonging to all' converts to 'possibly belonging to none' or 'not to all', and 'possibly belonging to some' converts to 'possibly not belonging to some'." (32a30–35)

What Aristotle seems to mean is that *Qa*-propositions are equivalent to *Qe* and *Qi* to *Qo*, as Buridan observes in Conclusion 7 of Book II.

Thom (p. 171) represents divided A- and E-propositions of contingency as $b^{\dagger} \rightarrow a^{\ddagger}$, I- and O-propositions as $b^{\dagger} \smile a^{\ddagger}$. Note, however, that *Qa*-propositions have existential import. To be sure, according to the *Summulae* [2, §1.8.3, p. 73], 'B is contingently A' is "hypothetical", so neither affirmative nor negative, but with affirmative and negative parts. Rather, *Qa*-propositions have a conjunctive predicate, and what makes a proposition affirmative or negative is the quality of its predicate, which in this case is neither simply one nor the other. That is, 'Every B is contingently A' is equivalent to 'Every B is possibly A and possibly fails to be A', and so entails the *Ma*-proposition 'Every B is possibly A'. But *Ma*-propositions, being affirmative, have existential import, and so *Qa*-propositions must have existential import too, and indeed imply *Qi*-propositions. This is shown in Fig. 2.

In Book II (on modal propositions in general),[12] Buridan said that he will concentrate on premises and conclusions concerning possibility and necessity with an affirmed mode,

[11] The word Aristotle uses here is *endechesthai*.

[12] Hubien [5, II 5, p. 62].

Every B is contingently A	(Qa) $\underline{b^\dagger} \to a^\ddagger$	(Qe)	Every B contingently fails to be A
↓			↓
Some B is contingently A	(Qi) $b^\dagger {}_\smile a^\ddagger$	(Qo)	Some B contingently fails to be A

Fig. 2 Representation of contingency propositions in Thom's notation

since, as we noted, 'not possibly' is equivalent to 'necessarily not' (that is, 'necessarily fails to be'), and 'not necessarily' to 'possibly not' (that is, 'possibly fails to be'). However, he speaks in Conclusions 22–25 of Book IV of syllogisms with premises with a negated mode. He writes, earlier in IV 3 [5, p. 130]:

> "From every proposition of necessity with an affirmed mode, whether affirmative or negative, there follows a proposition of contingency with a negated mode. For it follows, 'B is necessarily A', or also 'B necessarily fails to be A', 'so B is not contingently A', and 'so B does not contingently fail to be A'. So 'No B is contingently A' is equivalent to 'Every B is necessarily A or necessarily fails to be A'."[13]

What Buridan needs here is a notion dual to 'contingently', just as 'possibly' is dual to 'necessarily'. The obvious choice is 'non-contingently', for 'non-contingently' means 'not contingently', and just as 'contingently' is equivalent to 'contingently not', so too 'non-contingently' is equivalent to 'non-contingently not'. Hence 'not contingently not' is equivalent to 'non-contingently not' and so to 'non-contingently', and hence 'contingently' and 'non-contingently' are dual.

Accordingly, (Qe) is correctly represented as $\underline{b^\dagger} \to a^\ddagger$, that is, 'Every B is contingently A', which is equivalent to 'Every B contingently fails to be A', as noted in Fig. 2. In contrast, ($\overline{Q}a$) has the negated mode to which Buridan refers in the above quotation, and which interprets 'No B is contingently A' in parallel to Buridan's interpretation of 'No B is necessarily A' and 'No B is possibly A'. To represent ($\overline{Q}a$), it helps (but is not essential) to supplement Thom's notation with a further symbol for non-contingency: let us represent the non-contingency of A by a°. It is axiomatic that

2.2 $\quad a^\ddagger \mid a^\circ$

and that

2.3 $\quad a^* \to a^\circ$.

Then ($\overline{Q}a$) 'Every B is not contingently A' is $b^\dagger \to a^\circ$, that is, every B is necessarily A or necessarily fails to be A, equivalently, $b^\dagger \mid a^\ddagger$, since $a^\ddagger$ and a° exhaust the possibilities. Note that the $\overline{Q}a$-proposition 'Every B is not contingently A' is equivalent to the $\overline{Q}e$-proposition 'Every B does not contingently fail to be A', just as Qa and Qe were equivalent.

Buridan emphasises in IV 3 that 'B is necessarily A or necessarily fails to be A' is not equivalent to a disjunctive proposition:

> "But we should not accept that a universal of contingency with a negated mode is equivalent to a disjunction made up of an affirmative and a negative of necessity with an affirmed mode. For

[13] Cf. *Summulae de Dialectica* [2, §1.8.3, p. 73]: "But a proposition about contingency with a negated mode can be said to be about necessity and impossibility disjunctively, for 'B is not contingently A' is equipollent to 'B has to be A or B cannot be A'. (Buridan [3, pp. 87–8]: *Sed illa de contingenti modo negato potest dici de necessario et impossibili disiunctive; nam istae aequipollent 'B non contingit esse A' et 'B necesse est esse A vel impossibile est esse A'.*) See also §1.8.10, [2, p. 99].

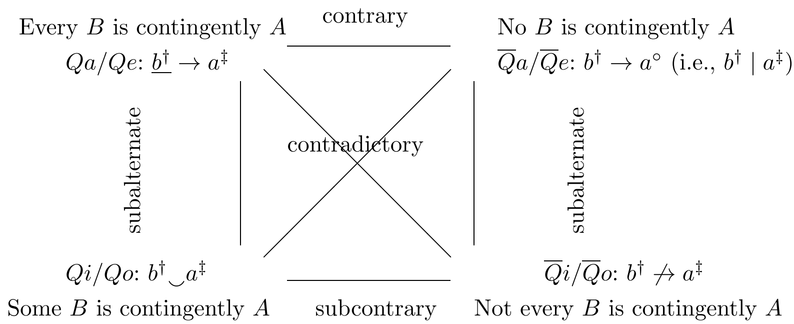


Fig. 3 Square of opposition for contingency propositions

this is true, 'No planet is contingently the moon', but this is false, 'Every planet is necessarily the moon or every planet necessarily fails to be the moon'." Hubien [5, p. 130]

In fact, we read in *Summulae* §1.8.3 that 'B is not contingently A' is "hypothetical", in this case, disjunctive, so again neither affirmative nor negative, but with affirmative and negative parts. In fact, the $\overline{Q}a$-proposition lacks existential import, for it is entailed by the clearly negative *Le*-proposition 'Every B necessarily fails to be A': if $b^\dagger \mid a^\dagger$ then $b^\dagger \mid a^\ddagger$, so $b^\dagger \rightarrow a^\circ$. Hence its contradictory is the *Qi*-proposition $\underset{\smile}{b^\dagger} a^\ddagger$. *Qi*-propositions (equivalent as we know to *Qo*-propositions) are in turn entailed by *Qa*-propositions, whose contradictory will be the $\overline{Q}i$-proposition 'Not every B is contingently A': $b^\dagger \not\rightarrow a^\ddagger$. Thus we have a new Square of Opposition for affirmed and negated modes of contingency. See Fig. 3.

3 Non-contingency Syllogisms

In Conclusions 22–25 of this chapter, Buridan discusses contingency propositions with negated mode. Paul Thom [9] and others [6, 7] have discussed the Conclusions of IV 3 as far as they concern affirmed modes.[14] However, only in the *Treatise on Consequences* does Buridan treat of contingency syllogisms with negated mode, and he clearly means $\overline{Q}$-propositions when he speaks of contingency-propositions with negated mode in Conclusions 22–25:

Conclusion 22 "From whatever premises there follows a conclusion of necessity with an affirmed mode there [also] follows a conclusion of contingency with a negated mode.

[14]Thom [9, p. 183 ff.] seems to miss the fact that these four conclusions concern non-contingency rather than contingency.

This Conclusion is proved by the fourth Conclusion of Book I, just as the previous one,[15] for [propositions] of necessity imply [propositions] of contingency [with negated mode]."[16]

The fourth Conclusion of Book I says *inter alia* that what follows from the conclusion of a syllogism follows from its premises. Suppose first that a universal necessity conclusion of the form *La* follows, i.e., 'Every B is necessarily A': $\underline{b^\dagger} \to a^*$. Then $b^\dagger \to a^* \to a^\circ$, by axioms 2.1 and 2.3, so $b^\dagger \to a^\circ$ by 1.3, that is, 'No B is contingently A', a contingency-conclusion of the form $\overline{Q}a$ with negated mode.

Now suppose we have an *Le*-conclusion, 'No B is possibly A': $b^\dagger \mid a^\dagger$. Then every B necessarily fails to be A, so no B contingently fails to be A, i.e., $b^\dagger \mid a^\ddagger$, by 1.9 and 1.4, and we have a contingency-conclusion of the form $\overline{Q}e$ with negated mode.

Thirdly, suppose we have an *Li*-conclusion, 'Some B is necessarily A': $b^\dagger \smile a^*$. Then $\Sigma d, b^\dagger \leftarrow \underline{d} \to a^* \to a^\circ$, by 1.6 and 2.3, i.e., $b^\dagger \smile a^\circ$, that is, some B is non-contingently A. Thus we can conclude by 2.2, 1.4 and 1.5 that not every B is contingently A: $b^\dagger \not\to a^\ddagger$, a contingency-conclusion of the form $\overline{Q}i$ with negated mode.

Finally, suppose we have an *Lo*-conclusion, 'Not every B is possibly A': $b^\dagger \not\to a^\dagger$. Then $\Sigma d, b^\dagger \leftarrow d \mid a^\dagger$. But $a^\ddagger \to a^\dagger \mid d$ so $d \mid a^\ddagger$ by 1.4, whence $\Sigma d, b^\dagger \leftarrow d \mid a^\ddagger$. So by 1.5, $b^\dagger \not\to a^\ddagger$, that is, not every B contingently fails to be A, a $\overline{Q}o$-proposition with negated mode.

Conclusion 23 "In the first and third figures from a major [premise] of contingency, whether with an affirmed mode or a negated mode, there follows a similar conclusion of contingency if the minor [premise] is of necessity or of possibility or of contingency.

This Conclusion, in the case of the first figure, is shown by the *dictum de omni et nullo*, just as the fourth Conclusion of this book was shown.

But as regards the third figure, it can be shown by expository syllogisms and per impossibile, just as the sixth Conclusion of this book was shown."[17]

---

[15]The fourth Conclusion of I 8 reads: "In any good consequence whatever follows from the consequent follows from the antecedent, and the consequent follows from whatever the antecedent follows from, and similarly, put in the negative, whatever does not follow from the antecedent does not follow from the consequent, and the antecedent does not follow from whatever the consequent does not follow from." (*Omnis bonae consequentiae quidquid sequitur ad consequens sequitur ad antecedens et ad quodcumque sequitur antecedens ad illud sequitur consequens; et similiter, negatiue, quidquid non sequitur ad antecedens non sequitur ad consequens et ad quodcumque non sequitur consequens ad illud non sequitur antecedens* [5, pp. 33–4].)

[16]Hubien [5, p. 131]: *Ad quascumque praemissas sequitur conclusio de necessario de modo affirmato ad easdem sequitur conclusio de contingenti de modo negato.*

Haec conclusio probatur per quartam conclusionem primi libri, sicut praecedens. Quia tales de necessario antecedunt talibus de contingenti.

[17]Hubien [5, p. 131]: *In prima figura et in tertia ex maiore de contingenti, siue de modo affirmato siue de modo negato, sequitur conclusio similiter de contingenti si minor sit de necessario uel de possibili uel de contingenti.*

Haec conclusio, quantum ad primam figuram, manifesta est per dici de omni uel de nullo, sicut manifesta erat quarta conclusio huius libri.

Sed quantum ad tertiam figuram manifestari potest per syllogismos expositorios et per impossibile, sicut manifestabatur sexta conclusio huius libri.

For example, take a first-figure syllogism with a major premise with negated mode, ‘No B is contingently A’ using ($\overline{Q}e$). That is, we have $b^\dagger \rightarrow a^\circ$ and $\underline{c^\dagger} \rightarrow b^\dagger$, so $\underline{c^\dagger} \rightarrow a^\circ$, by 1.3 (that is, the *dictum de omni*), and hence Barbara $\overline{\mathrm{Q}}\mathrm{M}\overline{\mathrm{Q}}$ is indeed valid.

Similarly, in the third figure, consider Disamis $\overline{\mathrm{Q}}\mathrm{M}\overline{\mathrm{Q}}$, say. Then we have $b^\dagger \not\twoheadrightarrow a^\ddagger$ and $\underline{b^\dagger} \rightarrow c^\dagger$ as premises. Thus $\Sigma d, b^\dagger \leftarrow d \mid a^\ddagger$ by 1.5, so $c^\dagger \leftarrow d \mid a^\ddagger$ by 1.3, whence $c^\dagger \not\twoheadrightarrow a^\ddagger$ by 1.5 again as required, showing that Disamis $\overline{\mathrm{Q}}\mathrm{M}\overline{\mathrm{Q}}$ is valid.

Conclusion 24 “From a major [premise] of contingency and an assertoric minor the first figure is valid to a particular conclusion of contingency, but not a universal.

This Conclusion is demonstrated just as the second part of the tenth Conclusion of this book was demonstrated. For that a universal conclusion does not follow is seen because every human contingently laughs and everything running is a human (assuming it is so), [but] then the universal conclusion would be false. And if the major [premise] has a negated mode there is a counter-instance because no horse contingently laughs, everything running is a horse (assuming it is so), [but] the universal conclusion would also be false.”[18]

Barbara QXQ is invalid, as is Barbara $\overline{\mathrm{Q}}\mathrm{X}\overline{\mathrm{Q}}$, for the same reason. Just because everything running is human, or a horse (we suppose), it does not follow that everything which might be running contingently laughs, since some things are not capable of laughter.

Conclusion 25 “From a universal major [premise] of contingency in the third figure and an assertoric minor there follows a conclusion also of contingency, but if the major is particular a conclusion of contingency does not follow.

The first part of the Conclusion is proved because in all moods of the third figure having a universal major [premise], if the minor, which is taken to be assertoric, is converted, the first figure results, which was said to be valid in the previous Conclusion.

But the second part is clear because while someone running is contingently laughing and everything running is a horse, nonetheless no horse is contingently laughing. It’s the same if the major [premise] is taken to be negative, because it is equivalent to an affirmative.

But if we speak of a negated mode then there is a counter-instance, for someone thinking is not contingently creating and everyone thinking is God (let us suppose), but it does not follow, ‘so God is not contingently creating’.”[19]

The first counterexample in the proof of the second part shows the invalidity of Disamis QXQ: with ‘laughs’, ‘runs’ and ‘horse’ for A, B and C, we have $b^\dagger \smile a^\ddagger$, $\underline{b} \rightarrow c$ and ‘No C

---

[18]Hubien [5, p. 131]: *Ex maiore de contingenti et minore de inesse ualet prima figura ad conclusionem de contingenti particularem, non ad uniuersalem.*

Haec conclusio declaratur sicut declarabatur secunda pars decimae conclusionis huius libri. Quod enim non sequatur conclusio uniuersalis patet. Quia omnem hominem contingit ridere et omne currens est homo (ponatur ita esse), tunc conclusio uniuersalis esset falsa. Et si maior sit de modo negato instantia est. Quia nullum equum contingit ridere, omne currens est equus (ponatur ita esse), conclusio etiam uniuersalis esset falsa.

[19]Hubien [5, p. 132]: *Ex maiore uniuersali de contingenti in tertia figura et minore de inesse sequitur conclusio similiter de contingenti, sed si maior sit particularis non sequitur conclusio de contingenti.*

Prima pars conclusionis probatur. Quia in omnibus modis tertiae figurae habentibus maiorem uniuersalem si minor quae ponitur de inesse conuertatur fiet prima figura, quae ualebat, ut dictum est in conclusione praecedente.

is contingently A'. But 'No C is contingently A' is the contradictory of a Qi-proposition, *viz* the $\overline{Q}e$-proposition $c^{\dagger} \mid a^{\ddagger}$, showing that $c^{\dagger} \smile a^{\ddagger}$ does not follow.

To appreciate the final counterexample, recall the remark cited earlier from IV 3, where Buridan affirms that God is indeed contingently creating:

> For God is contingently creating, but nothing which is contingently God [is] contingently creating, because nothing is contingently God, indeed, everything is necessarily God or necessarily fails to be God."[20]

Thus the counterexample has premises $b^{\dagger} \not\rightarrow a^{\ddagger}$, $\underline{b} \rightarrow c$, which could be true so interpreted. But $c^{\dagger} \not\rightarrow a^{\ddagger}$ does not follow. It is false on the given interpretation. Hence $c^{\dagger} \rightarrow a^{\ddagger}$ is compatible with the premises, so Disamis $\overline{\text{Q}}\text{X}\overline{\text{Q}}$ is invalid.

References

1. Aristotle: Prior Analytics Book I. Clarendon Aristotle Series. Clarendon Press, Oxford (2009). Translated with an Introduction and Commentary by Gisela Striker
2. Buridan, J.: Summulae de Dialectica Yale U.P., New Haven (2001). An annotated translation, with a philosophical introduction by Gyula Klima
3. Buridan, J.: Summulae de Propositionibus. Brépols, Turnhout (2005). Ria van der Lecq, ed.
4. Buridan, J.: Treatise on Consequences. Fordham University Press, New York (2014). S. Read, Eng. Tr.
5. Hubien, H.: Iohanni Buridani: Tractatus de Consequentiis. Publications Universitaires, Louvain (1976). English translation in [4]
6. Hughes, G.: The modal logic of John Buridan. In: Atti del Congresso Internazionale di Storia della Logica: La Teorie delle Modalità, pp. 93–111. CLUEB, Bologna (1989)
7. Lagerlund, H.: Modal Syllogistics in the Middle Ages. Brill, Leiden (2000)
8. Thom, P.: The two Barbaras. Hist. Philos. Logic **12**(2), 135–149 (1991)
9. Thom, P.: Medieval Modal Systems. Ashgate, Farnham (2003)

S. Read (✉)
Arché Research Centre, University of St Andrews, 17-19 College St., St Andrews, Scotland, UK
e-mail: slr@st-andrews.ac.uk

Secunda pars patet. Quia quamuis quendam currentem contingat ridere et omne currens sit equus, tamen nullum equum contingit ridere. Similiter est si maior ponatur negatiua, quia aequiualet affirmatiuae.

Si autem loquamur de modo negato adhuc instatur. Quia quendam intelligentem non contingit creare et omnis intelligens est deus (ponatur hoc), non sequitur "ergo deum non contingit creare".

[20] Hubien [5, p. 130]: *Quoniam deum contingit creare, et tamen nihil quod contingit esse deum et contingit creare, quia nihil contingit esse deum, immo omne necesse est esse deum uel necesse est non esse deum.*

Symbolic Existence in Hugh MacColl: A Dialogical Approach

Juan Redmond

Abstract In this paper, we present a dynamic dialogical interpretation of the notion of symbolic existence proposed by Hugh MacColl in 1906 in his book *Symbolic Logic and its Applications*. We begin by analyzing how MacColl presented the notion in the framework of his symbolic logic, which includes nonexistent objects. Then, from the perspective of dialogical logic, we propose a logical interpretation of the notion of symbolic existence implementing a dynamic quantifier and a dependency predicate inspired by the artifactual theory of Amie Thomasson. This allows us to show how the ontologically symbolic character of an entity may be understood as it relates to its possible existence and in a pragmatic context.

Keywords MacColl · Existence · Dialogical logic

Mathematics Subject Classification Primary 03A05 · Secondary 01A60 · 03B99

1 Historical Introduction

Hugh MacColl (1837–1909) was a mathematician and logician who was born and raised in Scotland. After working in various regions of Great Britain for a few years, he moved to Boulogne-sur-Mer (France), where the bulk of his work was carried out. Although his work does not fully satisfy the canons or rigorous standards of Frege's philosophy of mathematics, it does represent an original and challenging approach to logic. His primary contribution to 19th century algebraic logic was that his calculus admitted not only an interpretation of classes (as in Boolean algebra), but also an interpretation based on propositions.

In effect, MacColl gave priority to the propositional interpretation because of its generality and he called it *pure logic*. In his book *Symbolic Logic and its Applications* [10] of 1906, MacColl published the final version of his logic, where the propositions are classified as either certain, impossible, contingent, true, or false. After his death, MacColl's work was largely neglected and received little of the attention it deserved. Moreover, many of his ideas were attributed to other authors. The most notorious instances are those of strict implication, the first formal approach to modal logic, and the discussions on the paradoxes of material implication, for which C.I. Lewis generally gets credit. The same can be said for his contributions to probabilistic logic (conditional probability), many-valued logic, relevant logics, and connexive logic. Even less known is the fact that he

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_22

explored the possibilities of constructing a formal system capable of sustaining reasoning with fictions.

There were two principal factors that accounted for his work falling into neglect. One of them concerns technical questions and the other one is related to his philosophical position.

With respect to the technical questions, the algebraic method used by MacColl was replaced by the logistic method proposed by Frege and followed by Peano, Russell, and others. It consisted in presenting a logical system as a closed set of axioms operated by a relation of logical consequence.

The second factor deals with MacColl's philosophy of logic. His philosophical ideas were based on a certain type of instrumentalism that went beyond the canons of formal 19th century logic, that is to say, mathematics as logic (logicism) and logic as algebra (the Boolean algebraic approach). In a recent paper, Shahid Rahman [12] remarked that the philosophical approach of MacColl, before being associated with empiricism or logicism, can be linked with conventionalism and French instrumentalism, as developed by his younger contemporaries, Henri Poincaré, Pierre Duhem, and the American pragmatism of Charles Saunders Peirce.

An interesting aspect of his work, which is often scorned, is that, in addition to his scientific publications, MacColl had literary interests. He published two novels, *Mr. Stranger's Sealed Packet* [6] *in* 1889 and *Ednor Whitlock* [7] *in* 1891, and the essay, *Man's Origin, Duty and Destiny* [11] in 1909. The first is the story of a voyage to the planet Mars. In fact, it was the third novel in English written about this subject. The other two works discuss conflicts between science and religion, as well as the problems of faith, doubt, and disbelief.

It is almost impossible to resist the temptation to compare the contributions of MacColl in these two fields, the literary and the scientific. From our point of view, such a comparison yields a rather curious result. On the one hand, MacColl, especially in his novel about the voyage to Mars, follows the conservative Victorian spirit of his era, but, on the other hand, his scientific work contains the most thought-provoking, innovative ideas that anyone could have proposed in the 19th century logic.

2 MacColl and a Logic of Nonexistent Objects

The most influential approach to the logic of nonexistent objects is the tradition that begins with Frege and Russell. The main idea is simple but, at the same time, a bit disappointing: to reason with fictions is to reason with propositions that are always false, with the exception of affirming the nonexistence of such fictions (Russell [16]), or that they possess no truth value (Frege).

MacColl adamantly disagreed with these perspectives.

MacColl's work on nonexistent objects came about as a result of his reaction to one of the issues of discussion in vogue at that time: *the ontological commitment of propositions*. This is related to a classical topic: the ontological commitment of the copula that links the subject and the predicate in a reasoned judgement.

MacColl's logic of nonexistence is based on a bipartite ontology and a domain of quantification. He makes the following distinctions:

1. The class of the existent or real:

> "Let e1, e2, e3, etc. (up to any number of individuals mentioned in our argument or investigation) denote our universe of real existences" MacColl [9] (page 74). "[T]hese are the class of individuals which, in the given circumstances, have a real existence" (MacColl [10] (page 42).

2. The class of nonexistents:

> "Let 01, 02, 03, etc., denote our universe of nonexistences, that is to say, of unrealities, such as centaurs, nectar, ambrosia, fairies, with self-contradictions, such as round squares, square circles, flat spheres, etc., including, I fear, the non-Euclidean geometry of four dimensions and other hyperspatial geometries" MacColl [9] (page 74); [10] (page 42).
>
> "[T]he class of individuals which, in the given circumstances, have not a real existence... It does not exist really, though (like everything else named) it exists symbolically" MacColl [10] (page 42). "[I]n no case, however, in fixing the limits of the class e, must the context or given circumstances be overlooked" MacColl [10] (page 43).

3. The domain of quantification, the universe of discourse, which contains the two preceding classes:

> "Finally, let S1, S2, S3, etc., denote our Symbolic Universe or 'Universe of Discourse,' composed of all things real or unreal that are named or expressed by words or other symbols in our argument or investigation" MacColl [9] (page 74). The individuals who are members of this class, are also members of one of the first two classes. "We may sum up briefly as follows: firstly, when any symbol A denotes an individual; then any intelligible statement f(A), containing the symbol A, implies that the individual represented by A has a symbolic existence; but whether the statement f(A) implies that the individual represented by A has real existence depends on the context" MacColl [9] (page 77).

Based on this, the predicates can be interpreted by means of classes containing real, unreal, or both: "secondly, when any symbol A denotes a class, then, any intelligible statement f(A) containing the symbol A implies that the whole class A has a symbolic existence; but whether the statement f(A) implies that the class A is wholly real, or wholly unreal, or partly read and partly unreal, depends upon the context."

> "When the members A1, A, & c., of any class A consist wholly of realities or wholly of unrealities, the class A is said to be a pure class; when A contains at least one reality and also at least one unreality, it is called a mixed class" MacColl [10] (page 43).

MacColl speaks of the existence of a class. In order to avoid misunderstanding, we believe that this should be understood as the existence of the elements of the class. It is worth noting that, with respect to the clarification that we just made, many imprecisions of this type are found in his work.

How should we understand these ideas? How do we formally reconstruct them in a way that does not betray their basic intuitions? It is not an easy task and it can be dismissed as being anachronistic. We will submit two possible interpretations: one is the previously mentioned publication of Shahid Rahman about the notion of symbolic existence in MacColl; the other is the formal approach from the perspective of dialogical logic developed principally in Redmond [14].

3 MacColl and Meinong: Rahman's Proposal on Utterances and Ontological Commitment

According to Rahman there are two types of existential utterances. In his analysis of the two utterances, he suggests that the interpretations of the notions of existent and nonexistent objects in MacColl, Meinong, and Russell are related to each another.

Rahman also defends a certain type of Meinongianism in MacColl—under one of the types of existential utterances, and would allow for a reconstruction with quantifiers of this type. However, we believe that Rahman's assessment does not do justice to a certain dynamic notion present in MacColl's ideas.

The two types of existential utterances Rahman finds in MacColl are the following:

1. A type of utterance whose scope is the symbolic universe. That is, it deals with the quantification of a totality of objects of domain, realities, and unrealities. Hence, everything in the universe of discourse exists, at least symbolically. This type of quantification of an entire domain is not compatible or concurrent with Meinongianism. It seems to be closer to the notion of existence of the early Russell.

Meinong, you will recall, possesses three ontological domains: (i) the domain of existent objects, concrete objects in space and time, indicated by the verb *existieren*; (ii) the domain of subsistent objects, such abstract objects, events, propositions, that do not exist, but rather possess a certain type of being that Meinong calls *bestehen*; and (iii) the domain of nonexistent objects that do not possess any kind of being, as in the case of chimeras and other fictional entities. In accordance with this basic scheme, then, Meinong's notion of nonexistence *excludes* any kind of being, including subsistence. In contrast, MacColl's notion of symbolic existence and Russell's version of subsistence include existent and nonexistent objects. In effect, MacColl and the early Russell think that even nonexistent objects possess a certain kind of being that Russell called (inappropriately, according to Rahman) *subsistence* and MacColl called *symbolic existence*. Comparing the two following texts could help to illustrate this idea:

(i) One of MacColl's texts from 1902:

> "Take, for example, the proposition "Non-existences are non-existent." This is a self-evident truism; can we affirm that it implies the existence of its subject nonexistences?
> (...) In pure logic the subject being always a statement, must exist—that is, it must exist as a statement" MacColl [8] (page 356).

(ii) One of Russell's texts in *Principia Mathematica*:

> "Whatever may be an object of thought, or may occur in any true or false proposition, or can be counted as one, I call a term... [E]very term has being, i.e. is in some sense. A man, a moment, a number, a class, a relation, a chimaera, or anything else that can be mentioned, is sure to be a term" (Russell [15] (page 43).

In this respect, Rahman suggests, it can be understood that Russell credits Meinong with uniting subsistent and nonexistent objects in one single domain.

It is worth noting that, following Rahman, MacColl's notion of symbolic existence allows us to find inferences that follow the rule of particularization.

$$\phi[x/ki] \rightarrow \exists x\phi \quad \text{(Particularization)}$$

For example, if we derive:

"Something is (symbolically) nonexistent"

from

"Don Quijote is (symbolically) nonexistent"

Perhaps the following derivation approximates MacColl's analysis:

"The nonexistent Don Quijote exists symbolically"

Therefore,

"Something nonexistent exists symbolically"

Another relevant point, which will be further developed below, is that there seems to be a certain space to think dynamically about the interaction between real and symbolic existence. Real existence should take place once the constitution of the universe of discourse has been specified. In this regard, symbolic existence would be assumed as long as we do not know the ontological constitution of our universe of discourse.

2. The second type of existential utterance ranges over the domain of realities.

If this idea, Rahman posits, is thought of in the context of two different types of particular and universal statements, with the implementation of two types of quantifiers, and we free this symbolic pair from any type of ontological commitment (quantified expressions free of any ontological commitment), then MacColl's proposal approximates the Meinongian interpretation of positive free logic, although not completely.

In accordance with MacColl's idea, consequently, the following use of the particularization leads us to a non-valid inference:

"Don Quijote is (symbolically) nonexistent"

Therefore,

"Something nonexistent really exists"

Of course, formulations such as "the nonexistent Don Quijote exists symbolically" sound strange and provoke the rejection of Russell from that perspective.

4 Predicates and Ontological Commitment

MacColl's two notions of existence (the real and the symbolic) seem to have been conceived of as predicates. MacColl's notation of existence, when it is applied to an individual or the members of a class, is indicated as an exponential. Exponentials are used, in principle, to express a predicative role. MacColl makes the following elemental presentation:

H^B where H is the domain and B is the predicate.

Example:

H the domain of horses.
B brown.
H^B The horse is brown: all the elements of H (horses) are brown.

With respect to the predicates of existence, we have:

H^e the horse is real or possesses real existence: all the elements of H (horses) are real existent objects.
H^0 the horse is an unreality: all the elements of H (horses) do not really exist.
H^S the horse possesses a symbolic existence: all the elements of H (horses) are symbolically existent.

5 MacColl and the Existential Dynamic: The Proposal of Dialogical Logic

Another proposal consists in reading MacColl's notion of existence from a dialogical approach. We believe that this perspective has significant advantages for a favorable formal implementation of its intuitions concerning existential utterances and their ontological commitments.

One of the advantages has to do with the notion of quantification. The dialogical perspective understands quantifiers as functions of choice and this seems to approximate the way in which MacColl attributed existence and other ontological conditions to objects. In effect, MacColl indicated the existent or nonexistent condition of a chosen object A, attributing the respective predicate as an exponential. Existence does not speak of individuals or objects as pertaining to a concept (the way Frege understood quantifiers, that is, as second order predicates). Existence attributes properties to chosen objects or individuals, which have been chosen from among a larger set originally without ontological specification.

Let us remember that MacColl also attributed truth values to utterances as exponentials. These values were going beyond true and false, as this chart illustrates:

Truth	τ	$\boldsymbol{A}^{\tau}$	A is true in a particular case or instance
False	ι	$\boldsymbol{A}^{\iota}$	A is false in a particular case
Certain	ε	$\boldsymbol{A}^{\varepsilon}$	A is always true (true in every case) within the limits of our data, so that its probability is 1
Impossible	η	$\boldsymbol{A}^{\eta}$	A contradicts certain datum or definition, so that its probability is 0
Variable	θ	$\boldsymbol{A}^{\theta}$	A is possible but uncertain (A is contingent) It is equivalent $A^{\neg\eta} : A^{\neg\varepsilon}$. A is neither impossible nor certain; A is possible but uncertain. The probability is neither 0 nor 1, but some proper fraction between the two

An expression of type $(\boldsymbol{A}_G)^{\tau}$ be read as: "It is true that this grey horse exists."[1]

Another advantage, perhaps the most important advantage for this paper, is that the dialogical reconstruction of the idea of symbolic existence in MacColl allows one to become aware—from a logical point of view—of the ontological dynamic displayed in the symbolic-real interface. In order to develop this idea, we shall first present the dialogical approach of logic.

[1]For more on the syntax of MacColl, see Rahman & Redmond [13].

6 A New Semantic Point of View: Dialogical Logic

The dialogic approach deals with logic as a pragmatic notion and is presented as an argument in the form of a dialogue. This dialogue is constructed of two parts: a *proponent* that defends a thesis, and an *opponent* that attacks the thesis. The thesis is valid if and only if the proponent manages to defend it against all the possible attacks that the opponent can carry out. The dialogues are organized according to two types of rules: (i) rules of particles that give the local signification of the game and (ii) structural rules that give the global signification.

For our purposes of proposing a logic that captures the idea of symbolic existence (in its dynamic relation with real existence), a free logic (*free logic*[2]) was designed from a dialogical perspective. To do this, our analysis incorporates the idea of ontological dependence developed by Amie Thomasson [16]. Closely following Ingarden [4] and Husserl [3] she proposes a realist perspective of abstract objects, which she calls *artifacts*. Understanding abstract objects as artifacts means conceiving of them as entities that are dependent on an act of singular creation or birth that transforms them into existent objects (this normally comes accompanied by a "baptism"). They are dependents of an author with whom they maintain a historical link, and dependents at the same time of the backup copies where they are materialized (for example, the copies of a book in which they are described). These copies allow the artifacts to be disseminated and, in turn, the existence of these copies sustains their existence. By these same means, they retain the capability of being referred to through chains of dependence, identifiable through these same relations and, finally, become susceptible to dying (disappearing) if these relations are irreversibly dissolved. Thomasson's distinctions allow us to think in a bipolar domain: on the one hand, the *dependents*, and on the other hand, the entities on which they depend (which correspond to the classical distinction between existent and nonexistent objects).

Using this as a base, it is possible for us to develop a dynamic perspective incorporating MacColl's notion of symbolic existence. It will entail deploying the predicate of dependence as it relates to the notion of dialogical choice, for a first-order logic.

To deploy the notion of dependence in a dialogical approach, we introduce the predicate of relation of dependence $\boldsymbol{D}$, to which we will give the specific semantic, $\boldsymbol{D}k_ik_j$, which is read k_i *depends ontologically on* k_j. The idea is that if a constant k_i is considered to be a fiction, it must form a relationship of ontological dependence with regard to a constant k_j that corresponds to an existent object (in an independent manner):

– $\boldsymbol{D}k_ik_j$ and $k_i = k_j$ if and only if k_i (k_j) denotes an existent object (in an independent manner).
– $\boldsymbol{D}k_ik_j$ and $k_i \neq k_j$ if and only if k_i is a fiction that depends ontologically on k_j such that $\boldsymbol{D}k_ik_j$.

7 Subdialogues and the Dynamic Quantifier: The Symbolic Level of Existence in MacColl

We will consider two levels of dialogues or subdialogues: *symbolic subdialogue* and *actualist subdialogue*. This distinction aims to formally capture the idea of a dynamic—

[2]See Lambert [5] and Bencivenga [1].

inspired by MacColl—that is, a function of the interpretation of the quantifiers. In effect, we are going to consider just one type of quantifier, but a quantifier that changes ontological importance as the proof advances. In this sense, we represent the quantifiers as possessing a dynamic ontological commitment: they pass from a stage of ontological indifferentiation to a more committed stage. In the symbolic subdialogue, then, the quantifiers behave as possibilist quantifiers to the extent that the variables of substitution correspond to the set of ontologically dependent and independent objects. Next, in the actualist subdialogue, the constants that substitute the linked variables will correspond only to ontologically independent entities.

Below we present the structural rules that define a dynamic dialogical logic that integrates relations of dependence, subdialogues, and a dynamic quantifier, all together provide an account of the symbolic-real interphase suggested by Hugh MacColl.

8 Rules

(RS-FLF) The proponent defends an existential quantifier or attacks a universal quantifier with symbolic constants or constants already introduced by the opponent.

This means that the proponent can use constants that appear in the thesis.

For relationships of ontological dependence, we have:

(R*D*-0) X can only attack the relationship of ontological dependence by applying **(RD-1)**.

(R*D*-5) Once the symbolic subdialogue has finished and with regard to the last atomic formula played by Y.

Definition 1 We say that a dialogue is symbolically finished if and only if the dialogue is closed and finished in accordance with the classical rules.

In order to be able to attack the relationship of dependence, the proponent must win the first part of the game. Then the opponent interrogates the proponent about the ontological commitment of the constants played in the atomic formulas in order to specify the objects to which they correspond. The rules have been conceived in such a manner that if the constants come from a rule of quantification, the independent status of the constant must be justified. This procedure integrates a logic that corresponds to the logic of supervaluation. In effect, in the two logics, the set of valid formulas that contains linked variables coincides with the set that corresponds to classical logic.

Definition 2 A *symbolic subdialogue* is defined as a subdialogue in which the symbolic status of the objects, corresponding to the constants played, have still not been specified by the application of the rules **(R*D*-1)**–**(R*D*-5)**. We call the subdialogue in which we apply the rules **(R*D*-1)**–**(R*D*-5)** *actualist subdialogue*.

(R*D*-1) When X plays an atomic formula that contains a k_i, Y can ask X the following question "?-$\boldsymbol{D}k_ik_j$", that is, what is the k_j on which k_i depends ontologically (k_j can be identical or different than k_i). X must, then, defend itself by justifying a relationship of dependence $\boldsymbol{D}k_ik_j$.

We have

Formula	Attack	Defense
X-!-Ak_i	Y-?-$\boldsymbol{D}k_ik_j$	X-!-$\boldsymbol{D}k_ik_j$

Let

O			P		
n	Ak_1				
$n+2$	$\boldsymbol{D}k_1k_j$		n	?-$\boldsymbol{D}k_1k_j$	$n+1$

Or let

O			P		
				Ak_1	n
$n+1$	?-$\boldsymbol{D}k_1k_j$	n		$\boldsymbol{D}k_1k_j$	$n+2$

(R*D*-2) When X plays an atomic formula containing a k_i and that same k_i has been used by X to defend an existential quantifier or attack a universal quantifier, Y can ask it "?-$\boldsymbol{D}k_ik_i$": Y can demand that he justify that k_i is found to be in a relationship of reflexive dependence, that is to say, that k_i exists independently X defends itself by justifying the relationship of reflexive dependence, $\boldsymbol{D}k_ik_i$.

(R*D*-3) When X concedes a relationship of ontological dependence $\boldsymbol{D}k_ik_j$ with $k_i \neq k_j$, it is saying that k_i is a fiction that depends on k_j, and at the same time concedes $\boldsymbol{D}k_jk_j$.

Corollary to the formal rule (RS-3) P does not have the right to introduce a relationship of ontological dependence.

If the proponent has won the first part and there are ramifications, the opponent has the right to interrogate about the status of the objects that correspond to the constants played at the end of each ramification. For example,

O						P					
							$(Ak_1 \wedge \exists x Ax \rightarrow \exists x(Ax \wedge \exists x Ax))$				0
1	$Ak_1 \wedge \exists x Ax$				0		$\exists x(Ax \wedge \exists x Ax)$				2
3	?-$\exists$				2		$Ak_1 \wedge \exists x Ax$				4
5	?$\wedge_1$	4	5′	?$\wedge_2$	4		Ak_1	8		$\exists x Ax$	8′
7	Ak_1		7′	$\exists x Ax$		1	?$\wedge_1$	6	1	?$\wedge_2$	6′
9	?-$\boldsymbol{D}k_2k_j$	8	9′	?-$\exists$	8′					Ak_2 ☺	12′
			11′	Ak_2					7′	?-$\exists$	10′
			13′	?-$\boldsymbol{D}k_2k_2$	12′(8′)						

(R*D*-4) X can update a constant (repeat the defense of an existential quantifier or the attack on a universal quantifier) if and only if Y has introduced a new constant which may serve the purpose of X or if Y has conceded this constant in a relationship of reflexive ontological dependence ($\boldsymbol{D}k_i k_i$).

Definition 3 We say that X *symbolically concedes an atomic formula* when in the symbolic subdialogue is defended from an attack ?-k_i from Y (about a universal quantifier) affirming an atomic formula $\phi[x/k_i]$.

(R*D*-5) When X has symbolically conceded an atomic formula $\phi[x/k_i]$ and k_i is a dependent object in the actualist subdialogue, then X can annul the concession of the atomic formula $\phi[x/k_i]$; that is, the concession does not have value except to the extent that the quantifiers were not interpreted in an actualist manner.

Thus, a symbolic constant is a constant for which the ontological status of an entity to which it corresponds continues to be indeterminate (as dependent or independent).

Definition 4 We call symbolic the following constants: totally new constants played by P, constants which appear in the initial thesis, or constants that are introduced by the opponent, and which continue indeterminate according to the rules **(R*D*-1)**–**(R*D*-5)**.

Thus, the dialogue is displayed initially with symbolic constants (they can be either real or fictions) and once this stage is finalized, the interrogations begin with respect to the constants that intervene in the atomic statements.

As examples we look at cases of specification and particularization:

	O			P	
				$Ak_1 \to \exists x Ax$	0
1	Ak_1	0		$\exists x Ax$	2
3	?-$\exists$	2		Ak_1	4
5	?-$\boldsymbol{D}k_1k_1$	4(3)			
7	$\boldsymbol{D}k_1k_2$ ☺		1	?-$\boldsymbol{D}k_1k_i$	6

	O			P	
				$\forall x Ax \to Ak_1$	0
1	$\forall x Ax$	0		Ak_1	4
3	[~~Ak_1~~]		1	?k_1	2
5	?-$\boldsymbol{D}k_1k_1$	4(2)			
7	$\boldsymbol{D}k_1k_2$		3	?-$\boldsymbol{D}k_1k_i$	6
9	Ak_2 ☺		1	?k_2	8

Note that in both proofs the individual or object k_1 maintains a *symbolic* status of existence until it has come to the second part of the dialogue in which it is required to respond to the questions (attacks) about the type of dependence that characterizes this individual or object. In this case, the triumph of the opponent proves that this formula, one of the principles of classical logic, is not valid for an extension of the universe of things beyond the causally existent entities.

For the case of Smullyan:

O			P		
				$\exists x(Ax \to \forall x Ax)$	0
1	?-$\exists$	0		$Ak_1 \to \forall x Ax$	2
3	Ak_1	2		$\forall x Ax$	4
5	? k_2	4		Ak_2	8
				$Ak_2 \to \forall x Ax$	6
7	Ak_2	6			
9	?-$\boldsymbol{D}k_2k_2$	8(1)		$\boldsymbol{D}k_2k_2$ ☺	12
11	$\boldsymbol{D}k_2k_2$		5(7)	?-$\boldsymbol{D}k_2k_2$	10

An immediate consequence of understanding abstract entities as dependent artifacts is that the domain cannot be empty. Whereas for Meinongians there are inexistent objects independent of the literary creators, the abstract entities as artifacts always presuppose (historically and consistently) independent objects (existent objects).

To conclude, in the following dialogue we show that for any object there is another object upon which it depends, whether it be a different object (as in the case of a fiction) or the same object, in the case of existent objects.

O			P		
				$Ak_1 \to \exists x \boldsymbol{D}k_1x$	0
1	Ak_1	0		$\exists x \boldsymbol{D}k_1x$	2
3	?-$\exists$	2		$\boldsymbol{D}k_1k_2$	6
5	$\boldsymbol{D}k_1k_2$		1	?-$\boldsymbol{D}k_1k_i$	4
Σ	$\boldsymbol{D}k_2k_2$				
7	?-$\boldsymbol{D}k_2k_2$	(3)6		$\boldsymbol{D}k_2k_2$ ☺	8

9 Conclusion

We have explored the notion of symbolic existence in Hugh MacColl. We analyzed how he presented it in the framework of his symbolic logic, which included nonexistent objects. We pointed out its advantages and its limitations. After that, we examined Shahid Rahman's interpretation of the notion of symbolic existence in Hugh MacColl, as it relates to the utterances of existence and the ontological commitments that can follow from them. We point out that this approach does not capture the dynamic relationship between symbolic and real existence as suggested by MacColl. With regard to this, we proposed an interpretation from the perspective of dialogical logic that is displayed in subdialogues (possibilist and actualist) that have a dynamic quantifier. In effect, we provided the rules that define this playful perspective, integrating the notion of artifact of Amie Thomasson in the form of a dependent predicate, which allowed us to show how the ontologically symbolic character of an entity can be understood relative to its possible existence and in a pragmatic context.

Appendix 1

Universal Quantifier

$\forall$		
Formula uttered	Attack	Defense
$\forall x\varphi$	A question	A formula
dialogical expressions		
X-!-$\forall x\varphi$	Y-?-k_i	X-!-$\varphi[x/k_i]$[6]

Explanation 1
X utters a universal quantified formula ($\forall x\varphi$) that must be defended (!). Upon uttering $\forall x\varphi$, the player X commits himself to uttering $\varphi[x/k_i]$ for any k_i. How does Y challenge the utterance "$\forall x\varphi$"? By demanding the utterance of $\varphi[x/k_i]$. However, since X commits itself to the utterance of $\varphi[x/k_i]$ for any k_i, the challenger Y has the right to decide for which one the player X should perform the utterance (Y-?-k_i). The defense consists of uttering $\varphi[x/k_i]$, when possible, for the k_i chosen by Y.

Existential Quantifier

$\exists$		
Formula uttered	Attack	Defense
$\exists x\varphi$	A question	A formula
dialogical expressions		
X-!-$\exists x\varphi$	Y-?	X-!-$\varphi[x/k_i]$

Explanation 2
X utters a universal quantified formula ($\exists x\varphi$) that must be defended (!). Upon uttering $\exists x\varphi$, the player X commits himself to uttering $\varphi[x/k_i]$ for at least one k_i. How does Y challenge the utterance "$\exists x\varphi$"? By demanding the utterance of $\varphi[x/k_i]$. However, since X commits itself to the utterance of $\varphi[x/k_i]$ for at least one k_i, X has the right to decide which one. The defense consists of uttering $\varphi[x/k_i]$, when possible, for the k_i chosen by X (X-!-$\varphi[x/k_i]$).

Appendix 2[3]

Let L be a first-order language built as usual upon the propositional connectives, the quantifiers, a denumerable set of individual variables, a denumerable set of individual constants and a denumerable set of predicate symbols (each with a fixed arity).

[3]Extracted from Clerbout [2].

We extend the language L with two labels O and P, standing for the players of the game, and the question mark '?'. When the identity of the player does not matter, we use variables X or Y (with $X \neq Y$). A *move* is an expression of the form 'X-e', where e is either a formula φ of L or the form '$?[\varphi_1, \ldots, \varphi_n]$'.

We now present the rules of dialogical games. There are two distinct kinds of rules named particle (or local) rules and structural rules. We start with the particle rules.

Previous move	X-$\varphi \wedge \psi$	X-$\varphi \vee \psi$	X-$\varphi \rightarrow \psi$	X-$\neg\varphi$
Challenge	Y-$?[\varphi]$ or Y-$?[\psi]$	Y-$?[\varphi, \psi]$	Y-φ	Y-φ
Defense	X-φ resp. X-ψ	X-φ or X-ψ	X-ψ	–

Previous move	X-$\forall x \varphi$	X-$\exists x \varphi$
Challenge	Y-$?[\varphi(a/x)]$	Y-$?[\varphi(a_1/x), \ldots, \varphi(a_n/x)]$
Defense	X-$\varphi(a/x)$	X-$\varphi(a_i/x)$ with $1 \leq i \leq n$

In this table, the $a_i s$ are individual constants and $\varphi(a_i/x)$ denotes the formula obtained by replacing every occurrence of x in φ by a_i. When a move consists in a question of the form '$?[\varphi_1, \ldots, \varphi_n]$', the other player chooses one formula among $\varphi_1, \ldots, \varphi_n$ and plays it. We can thus distinguish between conjunction and disjunction, on the one hand, and universal and existential quantification, on the other hand, in terms of which player has a choice. In the cases of conjunction and universal quantification, the challenger chooses which formula he asks for. Conversely, in the cases of disjunction and existential quantification, the defender is the one who can choose between various formulas. Notice that there is no defense in the particle rule for negation.

Particle rules provide an abstract description of how the game can proceed locally: they specify the way a formula can be challenged and defended according to its main logical constant. In this way, we say that these rules govern the local level of meaning. Strictly speaking, the expressions occurring in the table above are not actual moves because they feature formulas schemata and the players are not specified. Moreover, these rules are indifferent to any particular situations that might occur during the game. For these reasons we say that the description provided by the particle rules is abstract.

Since the players' identities are not specified in these rules, we say that particle rules are symmetric: that is, the rules are the same for the two players. The fact that the local meaning is symmetric (in this sense) is one of the biggest strengths of the dialogical approach to meaning. In particular, it is the reason why the dialogical approach is immune to a wide range of trivializing connectives such as Prior's *tonk*.

The expressions occurring in particle rules are all move schematas. The words "challenge" and "defence" are convenient to name certain moves according to their relationship with other moves. Such relationships can be precisely defined in the following way. Σa sequence of moves. The function p_Σ assigns a position to each move in Σ, starting with 0. The function F_Σ assigns a pair $[m, Z]$ to certain moves N in Σ, where m denotes a position smaller than $p_\Sigma(N)$ and Z is either C or D, standing respectively for "challenge" and "defense". That is, the function F_Σ keeps track of the relations of challenge and defense as they are given by the particle rules. Consider for example the following sequence Σ:

$$\text{P-}\varphi \wedge \psi, \quad \text{P-}\chi \vee \psi, \quad \text{O-}?[\varphi], \quad \text{P-}\varphi.$$

In this sequence, we have, for example, $p_\Sigma(\text{P-}\chi \vee \psi) = 1$.

A *play* (or dialogue) is a legal sequence of moves, i.e. a sequence of moves which observes the game rules. The rules of the second kind that we mentioned, the structural rules, give the precise conditions under which a given sentence is a play. The *dialogical game* for φ, written $D(\varphi)$, is the set of all plays with φ as the thesis (see the Starting rule below). The structural rules are the following:

SR0 (Starting rule) Let φ be a complex formula of L. For every $\pi \in D(\varphi)$ we have:

- $p_\pi(\text{P-}\varphi) = 0$,
- $p_\pi(\text{O-}n := i) = 1$,
- $p_\pi(\text{P-}m := j) = 2$.

In other words, any play π in $D(\varphi)$ starts with P-φ. We call φ the *thesis* of the play and of the dialogical game. After that, the Opponent and the Proponent successively choose a positive integer called *repetition rank*. The role of these integers is to ensure that every play ends after finitely many moves, in a way specified by the next structural rule.

SR1 (Classical game-playing rule)

- Let $\pi \in D(\varphi)$. For every M in π with $p_\pi(M) > 2$ we have $F_\pi(M) = [m', Z]$ with $m' < p_\pi(M)$ and $Z \in \{C, D\}$
- Let r be the repetition rank of player X and $\pi \in D(\varphi)$ such that
  - The last member of π is a Y move,
  - M_0 is a Y move of position m_0 in π,
  - $M_1, \ldots, M_n$ are X moves in π such that $F_\pi(M_1) = \cdots = F_\pi(M_n) = [m_0, Z]$.

Consider the sequence[4] $\pi' = \pi * N$ where N is an X move such that $F_{\pi'}(N) = [m_0, Z]$. We have $\pi' \in D(\varphi)$ only if $n < r$.

The first part of the rule states that every move after the choice of repetition ranks is either a challenge or a defense. The second part ensures finiteness of plays by setting the player's repetition rank as the maximum number of times he can challenge or defend against a given move of the other player.

SR2 (Formal rule) Let ψ be an elementary sentence, N be the move P-ψ and M be the move O-ψ. A sequence π of moves is a play only if we have: if $N \in \pi$ then $M \in \pi$ and $p_\pi(M) < p_\pi(N)$.

That is, the Proponent can play an elementary sentence only if the Opponent played it previously. The formal rule is one of the characteristic features of the dialogical approach: other game-based approaches do not have it (see comments below).

One way to understand the formal rule is that it establishes a kind of game where one of the players must play without knowing meaning of the elementary sentences involved. Now, if the ultimate grounds of a dialogical thesis are elementary sentences and if this is implemented by the use of a formal rule, then the dialogues are in this sense necessarily asymmetric. Indeed, if both contenders were restricted by the formal rule no elementary

[4] We use $\pi * N$ to denote the sequence obtained by adding move N to the play π.

sentence can ever be posited. Thus, we implement the formal rule by designing one player, called the *proponent*, whose declarative utterances of elementary sentences are at least, at the start of the dialogue, restricted by this rule. Moreover, the formal rule triggers a novel notion of validity. Validity is not being understood as being true in every model, but as *having a winning strategy independently of any model* or more generally independently of any *material* grounding claim (such as truth or justification). The copy-cat strategy implicit in the formal rule is not copy-cat of groundings but copy-cat of declarative utterances involving elementary sentences. The copy-cat of groundings or contents corresponds rather to the modified formal rule for material analytic dialogues.

A play is called *terminal* when it cannot be extended by further moves in compliance with the rules. We say it is X terminal when the last move in the play is an X move.

SR3 (Winning rule) Player X wins the play π only if it is X terminal.

Consider, for example, the following sequences of moves:

$$\text{P-}Qa \wedge Qb, \quad \text{O-}n := 1, \quad \text{P-}m := 6, \quad \text{O-?}[Qa], \quad \text{P-}Qa$$

$$\text{P-}Qa \rightarrow Qa, \quad \text{O-}n := 1, \quad \text{P-}m := 12, \quad \text{O-}Qa, \quad \text{P-}Qa$$

The first one is not a play because it contravenes the Formal rule: with his last move, the Proponent plays an atomic sentence although the Opponent did not play it beforehand. By contrast, the second sequence is a play in $D(\text{P-}Qa \rightarrow Qa)$. We often use a convenient table notation for plays. For example, we can write this play as follows:

	O			P	
				$Qa \rightarrow Qa$	0
1	$n := 1$			$m := 12$	2
3	Qa	(0)		Qa	4

The numbers in the external columns are the positions of the moves in the play. When a move is a challenge, the position of the challenged move is indicated in the internal columns, as with move 3 in this example. Notice that such tables carry the information given by the functions p and F in addition to represent the play itself.

However, when we want to consider several plays together—or example when building a strategy—such tables are not that perspicuous. So we do not use them to deal with dialogical games for which we prefer another perspective. The *extensive form* of the dialogical game $D(\varphi)$ is simply the tree representation of it, also often called the game-tree. More precisely, the extensive form E_φ of $D(\varphi)$ is the tree (T, l, S) such that:

(i) Every node t in T is labelled with a move occurring in $D(\varphi)$
(ii) $l : T \rightarrow N$
(iii) $S \subseteq T^2$ with:

- There is a unique t_0 (the root) in T such that $l(t_0) = 0$, and t_0 is labeled with the thesis of the game.
- For every $t \neq t_0$ there is a unique t' such that $t'St$.
- For every t and t' in T, if tSt' then $l(t') = l(t) + 1$.
- Given a play π in $D(\varphi)$ such that $p_\pi(M') = p_\pi(M) + 1$ and t, t' respectively labelled with M and M', then tSt'.

Many metalogical results concerning dialogical games are obtained by considering them by leaving the level of rules and plays and moving to the level of strategies. Among these results, significant ones are given in terms of the existence of winning strategies for a player. We now define these notions and give examples of results.

A *strategy* for Player X in $D(\varphi)$ is a function which assigns an X move M to every non terminal play π with a Y move as last member such that extending π with M results in a play. An X strategy is *winning* if playing according to it leads to X's victory no matter how Y plays.

A strategy can be considered from the viewpoint of extensive forms: the extensive form of an X strategy σ in $D(\varphi)$ is the tree-fragment $E_{\varphi,\sigma} = (T_\sigma, l_\sigma, S_\sigma)$ of E_φ such that:

(i) The root of $E_{\varphi,\sigma}$ is the root of E_φ.
(ii) Given a node t in E_φ labelled with an X move, we have that $tS_\sigma t'$ whenever tSt'.
(iii) Given a node t in E_φ labelled with a Y move and with at least one t' such that tSt', then there is a unique $\sigma(t)$ in T_σ where $tS_\sigma\sigma(t)$ and $\sigma(t)$ is labelled with the X move prescribed by σ.

Here are some examples of results which pertain to the level of strategies.[5]

– Winning P strategies and leaves. *Let w be a winning* P *strategy in $D(\varphi)$. Then every leaf in $E_{\varphi,w}$ is labelled with a* P *signed atomic sentence*.
– Determinacy. *There is a winning X strategy in $D(\varphi)$ if and only if there is no winning Y strategy in $D(\varphi)$*.
– Soundness and Completeness of Tableaux. *Consider first-order tableaux and first-order dialogical games. There is a tableau proof for φ if and only if there is a winning* P *strategy in $D(\varphi)$*.

By soundness and completeness of the tableau method with respect to model-theoretical semantics, it follows that existence of a winning P strategy coincides with validity: *There is a winning* P *strategy in $D(\varphi)$ if and only if φ is valid*.

Examples of Extensive Forms Extensive forms of dialogical games and of strategies are infinitely generated trees (trees with infinitely many branches). Thus it is not possible to actually write them down. But an illustration remains helpful, so we add Figs. 1 and 2 below.

Figure 1 partially represents the extensive form of the dialogical game for the formula $\forall x(Qx \to Qx)$. Every play in this game is represented as a branch in the extensive form: we have given an example with the leftmost branch which represents one of the simplest and shortest plays in the game. The root of the extensive form is labelled with the thesis. After that, the Opponent has infinitely many possible choices for her repetition rank: this is represented by the root having infinitely many immediate successors in the extensive form. The same goes for the Proponent's repetition rank, and every time a player is to choose an individual constant.

Figure 2 partially represents the extensive form of a strategy for the Proponent in this game. It is a fragment of the tree of Fig. 1 where each node labelled with an O move has at most one successor. We do not keep track of all the possible choices for P any more: every time the Proponent has a choice in the game, the strategy selects exactly one

[5] These results are proven, together with others, in Clerbout [2].

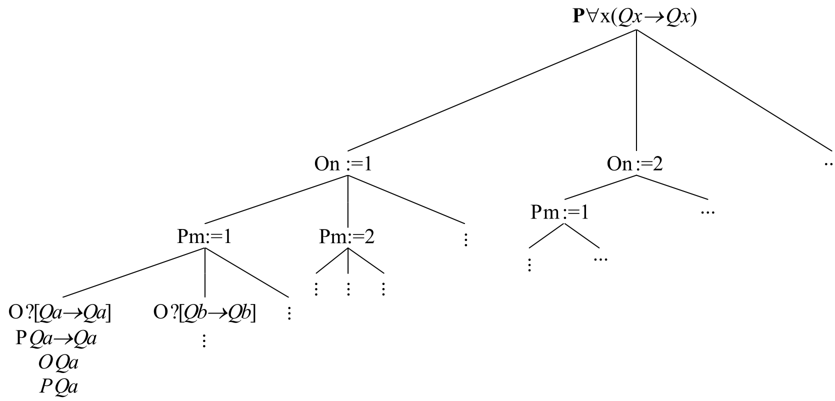


Fig. 1 Extensive form of the dialogical game for the formula $\forall x(Qx \rightarrow Qx)$

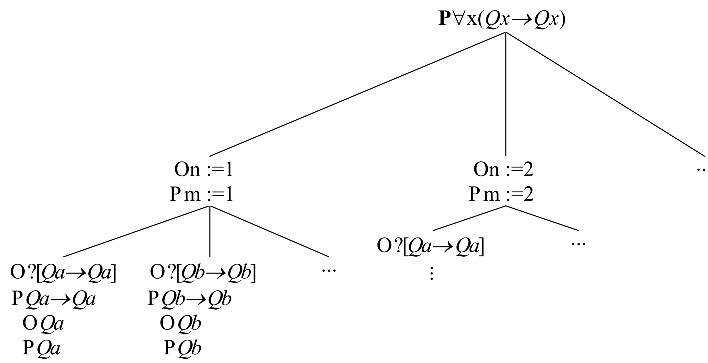


Fig. 2 It is a fragment of the tree of Fig. 1 where each node labelled with an O move has at most one successor

of the possible moves. But since all the possible ways for the Opponent to play must be taken into account by a strategy, the other ramifications are kept. In our example, the strategy prescribes to choose the same repetition rank as the Opponent. Of course, there are infinitely many other strategies available for P.

References

1. Bencivenga, E.: Free logics. In: Gabbay, D.M., Guenther, F. (eds.) Handbook of Philosophical Logic. Vol. 3: Alternatives to Classical Logic. Synthese Library, vol. 166, pp. 373–426. Reidel, Dordrecht (1983).
2. Clerbout, N.: First-order dialogical games and tableaux. J. Philos. Log. **42**(3) (2013)
3. Husserl, E.: Logical Investigations. Routledge and Kegan Paul, London (1970) (J.N. Findlay, Trans.) 1900
4. Ingarden, R.: The Literary Work of Art. Northwestern University Press, Evanston (1931) (George Grabowicz, Trans.)
5. Lambert, K.: Free Logics: Their Foundations, Character, and Some Application Thereof. ProPhil: Projekte zur Philosophie, vol. 1. Academia, Sankt Augustin (1997)

6. MacColl, H.: Mr. Stranger's Sealed Packet. Chatto and Windus, London (1889)
7. MacColl, H.: Ednor Whitlock. Chatto and Windus, London (1891)
8. MacColl, H.: Symbolic reasoning (IV). Mind **11**(43), 352–368 (1902)
9. MacColl, H.: Symbolic reasoning (VI). Mind **14**(53), 74–81 (1905)
10. MacColl, H.: Symbolic Logic and Its Applications. Longman, Green & Co, London (1906)
11. MacColl, H.: Man's Origin, Destiny and Duty. Williams and Norgate, London (1909)
12. Rahman, S.: Some remarks on Hugh MacColl's notion of symbolic existence. Philos. Sci. **15**(1), 149–162 (2011)
13. Rahman, S., Redmond, J.: Hugh MacColl: An Overview of His Logical Work with Anthology. College Publications, London (2007)
14. Redmond, J.: Logique Dynamique de la Fiction: Pour une Approche Dialogique. College Publications, London (2010)
15. Russell, B.: The Principles of Mathematics. Cambridge University, Cambridge (1903)
16. Thomasson, A.: Fiction and Metaphysics. Cambridge University Press, Cambridge (1999)

J. Redmond (✉)
Instituto de Filosofía, Universidad de Valparaíso and Conicyt, Valparaíso, Chile
e-mail: juan.redmond@uv.cl

Béziau's Contributions to the Logical Geometry of Modalities and Quantifiers

Hans Smessaert and Lorenz Demey

Abstract The aim of this paper is to discuss and extend some of Béziau's (published and unpublished) results on the logical geometry of the modal logic S5 and the subjective quantifiers *many* and *few*. After reviewing some of the basic notions of logical geometry, we discuss Béziau's work on visualising the Aristotelian relations in S5 by means of two- and three-dimensional diagrams, such as hexagons and a stellar rhombic dodecahedron. We then argue that Béziau's analysis is incomplete, and show that it can be completed by considering another three-dimensional Aristotelian diagram, viz. a rhombic dodecahedron. Next, we discuss Béziau's proposal to transpose his results on the logical geometry of the modal logic S5 to that of the subjective quantifiers *many* and *few*. Finally, we propose an alternative analysis of *many* and *few*, and compare it with that of Béziau's. While the two analyses seem to fare equally well from a strictly *logical* perspective, we argue that the new analysis is more in line with certain *linguistic* desiderata.

Keywords Logical geometry · Modal logic · S5 · Subjective quantifiers · Many/few · Aristotelian diagram · Stellated rhombic dodecahedron · Cuboctahedron · Rhombic dodecahedron

Mathematics Subject Classification (2000) Primary 03B45 · 03B65 · 52B10 · Secondary 03C80 · 52B05 · 52B15 · 68T30

1 Introduction

In recent years, Jean-Yves Béziau has been the driving force behind the renewed interest in the square of oppositions and related Aristotelian diagrams. On a practical level, he is the main organiser of a number of highly successful conference series and the editor-in-chief of a journal and a book series, all of which have functioned as a platform for discussion of recent discoveries about Aristotelian diagrams. On a theoretical level, Béziau has made significant contributions to the study of Aristotelian diagrams for a number of logical systems, such as the modal logic S5 and the subjective quantifiers *many* and *few*. The main aims of this paper are to provide a detailed presentation of some of these (published and unpublished) results, to evaluate them from a logico-linguistic perspective, and finally, to show how they relate to the framework of logical geometry that we have recently been developing.

The paper is organised as follows. Section 2 provides a brief overview of some of the basic notions of logical geometry. Next, in Sect. 3, we discuss Béziau's work on

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_23

visualising the Aristotelian relations in S5 by means of two- and three-dimensional diagrams, such as hexagons and a stellar rhombic dodecahedron. We then argue in Sect. 4 that Béziau's analysis is incomplete, and show that it can be completed by considering another three-dimensional Aristotelian diagram, viz. a rhombic dodecahedron. In Sect. 5, we discuss Béziau's proposal to transpose his results on the logical geometry of the modal logic S5 to that of the subjective quantifiers *many* and *few*. Next, in Sect. 6, we propose an alternative analysis of *many* and *few*. While the two analyses seem to fare equally well from a strictly *logical* perspective, we argue that the new analysis is more in line with certain *linguistic* desiderata. Section 7 provides a comparison between both analyses from the perspective of logical geometry, i.e. in terms of the various Aristotelian diagrams that they give rise to. Section 8, finally, wraps things up and mentions some questions for further research.

2 The Basics of Logical Geometry

The logical geometry of a certain logical system or lexical field consists in the visual representation of the logical behaviour of its members. This behaviour can be classified according to a number of families of logical relations, such as the family of Aristotelian relations, the family of duality relations, etc. In the present paper, we will focus exclusively on the Aristotelian relations, i.e. contradiction (*CD*), contrariety (*C*), subcontrariety (*SC*) and subalternation (*SA*). Intuitively, the first three relations are defined in terms of whether the formulas can be true together (the $\varphi \wedge \psi$ part in the formal definition below) and whether they can be false together (the $\neg\varphi \wedge \neg\psi$ part in the formal definition below);[1] the fourth relation, *SA*, is defined in terms of truth propagation. Formally, the Aristotelian relations are defined relative to a logical system S:[2] two formulas φ and ψ are said to be

S-*contradictory*	iff	$\mathsf{S} \models \neg(\varphi \wedge \psi)$	and	$\mathsf{S} \models \neg(\neg\varphi \wedge \neg\psi)$,
S-*contrary*	iff	$\mathsf{S} \models \neg(\varphi \wedge \psi)$	and	$\mathsf{S} \not\models \neg(\neg\varphi \wedge \neg\psi)$,
S-*subcontrary*	iff	$\mathsf{S} \not\models \neg(\varphi \wedge \psi)$	and	$\mathsf{S} \models \neg(\neg\varphi \wedge \neg\psi)$,
in S-*subalternation*	iff	$\mathsf{S} \models \varphi \rightarrow \psi$	and	$\mathsf{S} \not\models \psi \rightarrow \varphi$.

These relations are abbreviated and visualised according to the code in Fig. 1. When the system S is clear from the context, we will often leave it implicit, and simply talk about 'contrary' instead of 'S-contrary', etc.

In logical geometry, we often use *bitstrings* as a compact way to represent the denotations of formulas [25, 28]. Bitstrings are, quite simply, sequences of bits (1 or 0), such as 101, 1100, 10110, etc. In the present paper, we will exclusively work with bitstrings of length 4. Furthermore, we distinguish between bitstrings of *level 1* (L1), *level 2* (L2), and *level 3* (L3), which are defined as having a value 1 in one, two or three of their

[1]It is well-known that $\neg(\neg\varphi \wedge \neg\psi)$ is equivalent to $\varphi \vee \psi$, but we choose to stick with the former notation, because it more clearly expresses the idea of φ and ψ being false together.

[2]The system S is assumed to have connectives expressing classical negation ($\neg$), conjunction ($\wedge$) and implication ($\rightarrow$), and a model-theoretic semantics ($\models$).

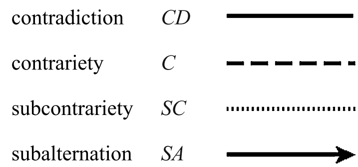


Fig. 1 Code for visually representing the Aristotelian relations

bit positions, respectively.[3] The Boolean operations on bitstrings are defined bitwise, for example, $1100 \wedge 1010 = 1000$, $1100 \vee 1010 = 1110$ and $\neg 1100 = 0011$. Given the availability of these Boolean operations, we can reformulate the definitions of the Aristotelian relations in terms of bitstrings: two bitstrings φ and ψ are said to be

$$\begin{array}{llll} \textit{contradictory} & \text{iff} & \varphi \wedge \psi = 0000 & \text{and} \quad \varphi \vee \psi = 1111, \\ \textit{contrary} & \text{iff} & \varphi \wedge \psi = 0000 & \text{and} \quad \varphi \vee \psi \neq 1111, \\ \textit{subcontrary} & \text{iff} & \varphi \wedge \psi \neq 0000 & \text{and} \quad \varphi \vee \psi = 1111, \\ \textit{in subalternation} & \text{iff} & \varphi \wedge \psi = \varphi & \text{and} \quad \varphi \vee \psi \neq \psi. \end{array}$$

The squares in Fig. 2 show the Aristotelian relations between three sets of formulas/bitstrings. Square (a) is the oldest and most widely known square, which is decorated with the quantified formulas of Aristotelian syllogistics, whereas square (b) represents the Aristotelian relations between four formulas from the modal logic S5. Finally, square (c) represents the Aristotelian relations between four bitstrings 'in abstracto', i.e. without referring to any particular logical system or lexical field. Note that square (b) for the modal logic S5 can be seen as a specific instance of square (c), if we take the bitstrings in the latter to be the denotations of the modal formulas in the former. Similar remarks apply to square (a).

The squares in Fig. 2 have been generalised to larger and more complex Aristotelian diagrams. One classical example is the hexagon first studied by Jacoby, Sesmat and Blanché [7, 8, 15, 23]. Note that the square is not closed under the Boolean operators; for example, the square for S5 contains the formulas $\Box p$ and $\neg\Diamond p$, but it does not contain their disjunction $\Box p \vee \neg\Diamond p$ (or any formula that is logically equivalent to it). By adding two extra vertices to the square, we obtain a hexagon that *is* Boolean closed.[4] Furthermore, the newly added formulas stand in Aristotelian relations to all formulas that were already present in the square and to each other, e.g. $\Box p \vee \neg\Diamond p$ is subcontrary to $\Diamond p$ and to $\neg\Box p$, and contradictory to $\Diamond p \wedge \neg\Box p$. The resulting configuration is an Aristotelian diagram called the *Jacoby–Sesmat–Blanché hexagon* (JSB); see Fig. 3(a). The abstract representation of this hexagon is the bitstring JSB hexagon in Fig. 3(b). Equivalently, the latter can be seen as the Boolean closure of the square in Fig. 2(c). For example, the bitstring representation of the disjunction $\Box p \vee \neg\Diamond p$ is 1001, which is the join of the bitstrings 1000 and 0001, representing the formulas $\Box p$ and $\neg\Diamond p$, respectively.

[3] As limiting cases, the non-contingent bitstrings 0000 and 1111 can be called *level 0* (L0) and *level 4* (L4), respectively.

[4] Formally, a diagram or set of formulas is said to be *Boolean closed* iff whenever it contains formulas φ, ψ, it also contains their contingent Boolean combinations ($\neg\varphi, \varphi \wedge \psi, \varphi \vee \psi$). Note that this definition is restricted to *contingent* Boolean combinations, so if $\Box p$ and $\neg\Box p$ occur in a diagram, then it is *not* required for the diagram to be Boolean closed that it also contains the contradiction $\Box p \wedge \neg\Box p$ and the tautology $\Box p \vee \neg\Box p$.

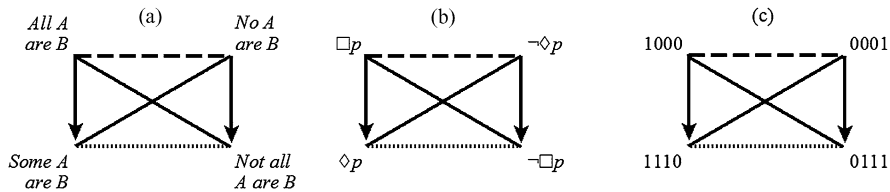


Fig. 2 Aristotelian squares for (**a**) syllogistics, (**b**) the modal logic S5, and (**c**) bitstrings

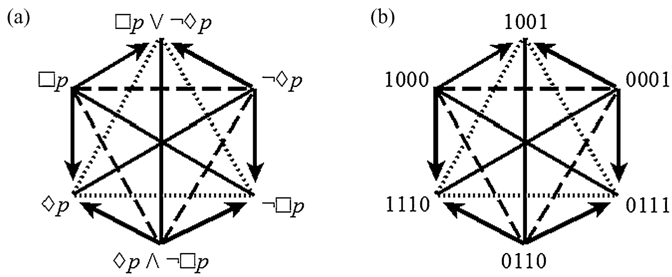


Fig. 3 Aristotelian JSB hexagons for (**a**) the modal logic S5 and (**b**) bitstrings

3 Béziau on the Logical Geometry of S5

We now turn to Béziau's results on the logical geometry of S5 [1, 2, 5]. Starting from the Aristotelian square in Fig. 2(b), he constructs three JSB hexagons for S5. The first hexagon is simply the Boolean closure of the original square, which was already discussed above (see Fig. 3(a)), and which is repeated here as Fig. 4(a). The two other hexagons are the Boolean closures of two new squares, which are obtained from the original square by replacing one of its diagonals with p–$\neg p$. The first new square is obtained by replacing the $\Box p$–$\neg\Box p$ diagonal, and gives rise to the JSB hexagon in Fig. 4(b). Similarly, the second new square is obtained by replacing the $\Diamond p$–$\neg\Diamond p$ diagonal, and gives rise to the JSB hexagon in Fig. 4(c).[5]

Béziau used these JSB hexagons to show that the modal logic S5 can model classical as well as non-classical modes of reasoning. The hexagon in Fig. 4(b) shows the relation between classical negation ($\neg$) and paracomplete negation ($\neg\Diamond$). Classical negation is a contradictory-forming operator: p and $\neg p$ are contradictory, and thus can be neither true together nor false together. By contrast, paracomplete negation is 'merely' a contrary-forming operator: p and $\neg\Diamond p$ are contrary, and thus can be false together. Similarly, the hexagon in Fig. 4(c) shows the relation between classical negation ($\neg$) and paraconsistent

[5]The process described above works not only for the S5 square (Fig. 2(b)), but also for the bitstring square (Fig. 2(c)). We thus obtain three bitstring JSB hexagons, the first one of which was already shown in Fig. 3(b). The two new ones are displayed simultaneously with their S5 counterparts in Fig. 4. For reasons of space, we will in the remainder of this paper no longer distinguish between S5 diagrams and bitstring diagrams, and decorate the Aristotelian diagrams sometimes with concrete formulas, sometimes with bitstrings, and sometimes with both simultaneously.

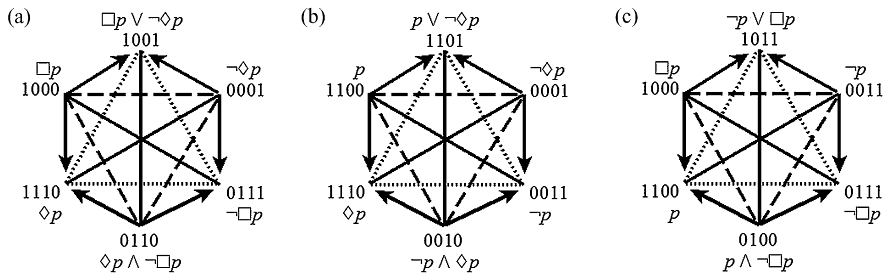
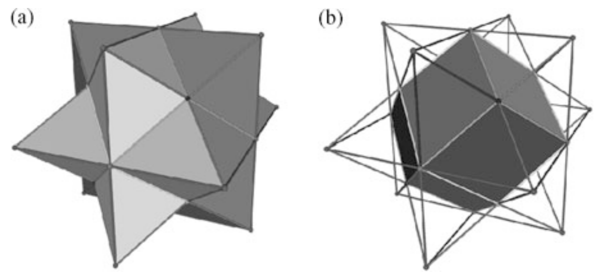


Fig. 4 Béziau's three JSB hexagons for S5: (**a**) the classical, (**b**) the paracomplete, and (**c**) the paraconsistent hexagon

Fig. 5 (**a**) Béziau's stellar rhombic dodecahedron, (**b**) the stellar rhombic dodecahedron as the first stellation of the rhombic dodecahedron



negation (¬□). Paraconsistent negation is 'merely' a subcontrary-forming operator: p and $\neg\Box p$ are subcontrary, and thus can be true together.[6]

The three JSB hexagons in Fig. 4 contain a total number of 12 formulas.[7] Béziau suggested that by moving from two-dimensional to three-dimensional diagrams, these 12 formulas can be visualised quite elegantly by means of a single Aristotelian diagram, viz. a *stellar rhombic dodecahedron* (Fig. 5(a)). This polyhedron is also known as 'Escher's solid'; geometrically speaking, it is the first stellation of the rhombic dodecahedron [17] (Fig. 5(b)).[8] Furthermore, Béziau remarked that if we ignore the subalternations on the outside edges of the three JSB hexagons in Fig. 4, we obtain three JSB *stars* (each star

[6]These hexagons also show that, from the perspective of subalternation, classical negation occupies a position that is intermediate between paracomplete and paraconsistent negation. In Fig. 4(b), the classical negation ($\neg p$) is *entailed by* the paracomplete negation ($\neg\Diamond p$), while in Fig. 4(c), it *entails* the paraconsistent negation ($\neg\Box p$).

[7]Since there are 3 hexagons and each hexagon contains 6 formulas, one might expect the total number to be $3 \times 6 = 18$ formulas. However, this calculation ignores the fact that certain formulas occur in two distinct hexagons. In particular, the formulas $\Diamond p$ and $\neg\Diamond p$ occur in hexagons (a) and (b), the formulas $\Box p$ and $\neg\Box p$ occur in hexagons (a) and (c), and the formulas p and $\neg p$ occur in hexagons (b) and (c).

[8]Somewhat confusingly, Béziau [1] talks about the 'stellar dodecahedron' instead of the 'stellar *rhombic* dodecahedron', thereby suggesting that the solid he had in mind (but never actually drew) is the stellation of the 'ordinary' *pentagonal* (i.e. Platonic) dodecahedron, rather than that of a *rhombic* dodecahedron. This confusion resurfaces in Moretti's remarks that Béziau's solid is "obtained by constructing a *pentagonal* pyramid or spike over each of the 12 *pentagonal* faces of a dodecahedron" [19, p. 75, our emphases]. Accordingly, the figure given by Moretti [19, p. 76] shows the stellation of a pentagonal dodecahedron, rather than that of a rhombic dodecahedron (although he still calls it 'Escher's solid' and attributes it to Béziau). In a more recent paper, Béziau does provide a figure of the stellar rhombic dodecahedron (without its S5-decoration) [5, p. 13], but still calls it the 'stellar dodecahedron'.

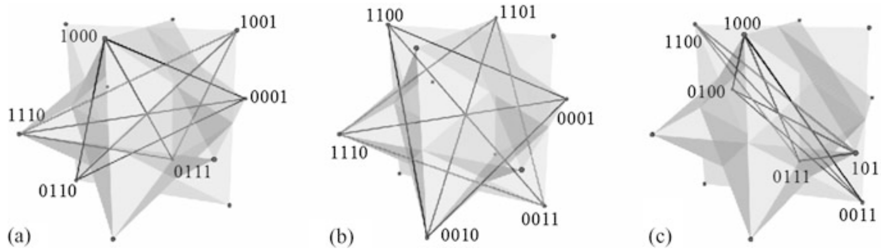


Fig. 6 Béziau's three JSB stars embedded inside the stellar rhombic dodecahedron

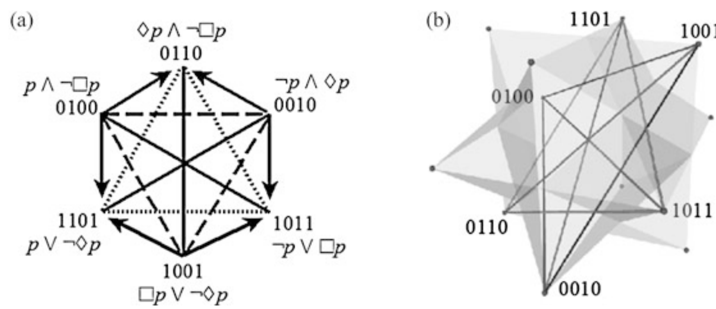


Fig. 7 (**a**) The fourth JSB hexagon; (**b**) the corresponding JSB star embedded inside the stellar rhombic dodecahedron

consists of a triangle of contraries interlocked with a triangle of subcontraries). Each of these three JSB stars can be embedded inside Béziau's stellar rhombic dodecahedron; see Fig. 6.

Soon after Béziau's discoveries, H. Smessaert and A. Moretti (then a PhD student of Béziau's) independently observed that a fourth JSB star can be constructed using the 12 formulas appearing in Béziau's stellar rhombic dodecahedron. Just as before, this JSB star can be seen as a JSB hexagon with the subalternations left out; see Fig. 7(a). This fourth JSB star can also be embedded inside the stellar rhombic dodecahedron (Fig. 7(b)), and therefore each of its six formulas already appears in one of the first three JSB stars/hexagons in Fig. 4. In other words, the novelty of the fourth JSB star does not consist in its formulas (these were already present in the previous stars), but rather in the fact that the pattern of Aristotelian relations between these formulas is again that of a JSB star/hexagon. Furthermore, by taking into account this fourth JSB star, we see that the stellar rhombic dodecahedron achieves a certain 'equilibrium':[9] on the one hand, each of the 12 formulas of the stellar rhombic dodecahedron appears in 2 JSB stars ($12 \times 2 = 24$); on the other hand, there are 4 JSB stars embedded inside the stellar rhombic dodecahedron, each of which contains 6 formulas ($4 \times 6 = 24$).

Dissatisfied with Béziau's decision to ignore the subalternation relations (i.e. to move from the JSB *hexagons* to the corresponding JSB *stars*), Moretti [18] went on to look for alternative polyhedra to represent the 12 formulas and their Aristotelian relations

[9] Also recall Footnote 7.

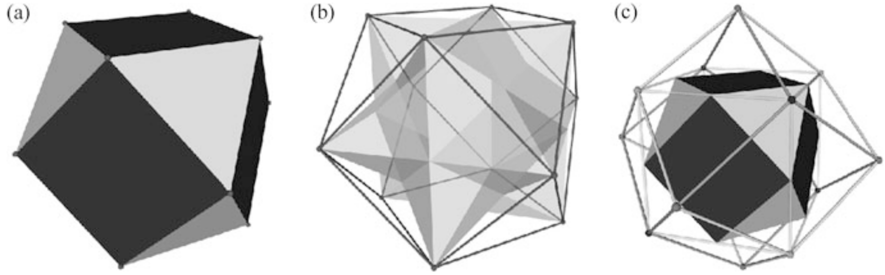


Fig. 8 (**a**) Moretti's cuboctahedron, (**b**) the cuboctahedron as the convex hull of Béziau's stellar rhombic dodecahedron, (**c**) the cuboctahedron as the dual polyhedron of the rhombic dodecahedron

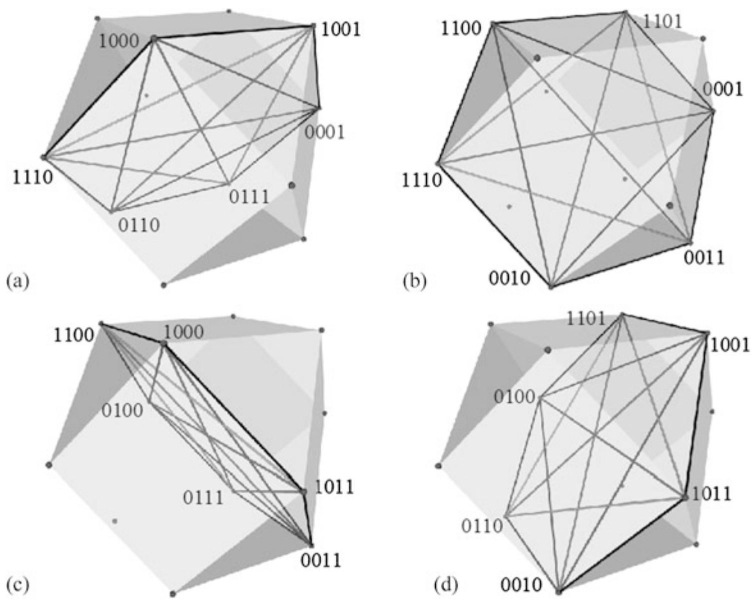


Fig. 9 Four JSB hexagons embedded inside Moretti's cuboctahedron

(including the subalternations). He proposed the *cuboctahedron* (Fig. 8(a)), and showed that the four JSB hexagons can be embedded inside it (Fig. 9)—just like the four corresponding JSB stars can be embedded inside Béziau's stellar rhombic dodecahedron (recall Figs. 6 and 7(b)). Interestingly, Moretti's cuboctahedron turns out to be the convex hull of Béziau's stellar rhombic dodecahedron (Fig. 8(b))—just like the JSB hexagons are the convex hulls of the corresponding stars. This geometric observation has a direct logical analogue: Moretti's cuboctahedron can be seen as the result of adding the subalternation relations to Béziau's stellar rhombic dodecahedron, and the four JSB hexagons are embedded in the former in exactly the same way as the corresponding JSB stars are embedded in the latter (compare Figs. 6 and 7(b) with Fig. 9). Finally, it should be noted that Moretti's cuboctahedron is the dual polyhedron of a rhombic dodecahedron (Fig. 8(c)).

Table 1 The 14 formulas of S5 and their bitstring representations

S5-formula	Bitstring	S5-formula	Bitstring
$\Box p$	1000	$\neg\Box p$	0111
$\neg\Diamond p$	0001	$\Diamond p$	1110
$\Box p \vee \neg\Diamond p$	1001	$\neg\Box p \wedge \Diamond p$	0110
p	1100	$\neg p$	0011
$p \wedge \neg\Box p$	0100	$\neg p \vee \Box p$	1011
$\neg p \wedge \Diamond p$	0010	$p \vee \neg\Diamond p$	1101
$\Box p \vee (\neg p \wedge \Diamond p)$	1010	$\neg\Box p \wedge (p \vee \neg\Diamond p)$	0101

4 Extending Béziau's Results on the Logical Geometry of S5

We have seen above that the 12 formulas of S5 that were considered by Béziau give rise to four distinct JSB hexagons. Each of these hexagons is, by itself, closed under the Boolean operators. However, when the 12 formulas are taken together (as is done in Béziau's stellar rhombic dodecahedron and Moretti's cuboctahedron), the resulting diagram is *not* Boolean closed. For example, it contains the formulas[10] $\Box p$ and $\neg p \wedge \Diamond p$, but it does not contain their disjunction $\Box p \vee (\neg p \wedge \Diamond p)$ (or any formula that is logically equivalent to it). Reformulating the example in terms of bitstrings: Béziau's analysis deals with the bitstrings 1000 and 0010, but not with their join 1010. This shows that Béziau's analysis is incomplete.

In order to obtain the Boolean closure of Béziau's analysis, we have to consider two additional formulas, viz. $\Box p \vee (\neg p \wedge \Diamond p)$ (which was already mentioned above) and its negation, $\neg\Box p \wedge (p \vee \neg\Diamond p)$. Note that these two formulas are syntactically more complex than any of the 12 formulas considered by Béziau, which might explain why these are exactly the two that were not included in his original analysis. In terms of bitstrings, the two new formulas correspond to the two new bitstrings 1010 and 0101, respectively. It is easy to see that by adding these 2 formulas/bitstrings to the 12 considered by Béziau, we obtain a set of 14 formulas/bitstrings that *is* Boolean closed; see Table 1.

Of course, the two new formulas enter into a variety of Aristotelian relations with the 12 old ones. In particular, Smessaert [25] noted that they yield two new JSB hexagons, which are shown in Fig. 10. Furthermore, it should be stressed that because of combinatorial reasons, the list of six JSB hexagons that we have now obtained (comprising the four that were already present in Béziau's analysis, together with the two new ones) is *exhaustive*, i.e. there are no additional JSB hexagons that can be constructed with the 14 S5-formulas under consideration.[11]

The question now arises as to how this set of 14 formulas (and the Aristotelian relations between them) can be visualised by means of a three-dimensional Aristotelian dia-

[10] Note, trivially perhaps, that these two formulas do not occur together in any of the four JSB hexagons considered above. After all, if a diagram contains these two formulas, then it cannot be Boolean closed.

[11] Pellissier [20] and Moretti [18] distinguish between a strong and a weak kind of JSB hexagon (the strong kind is the kind we have considered up till now), and note that the 14 formulas not only yield 6 strong JSB hexagons, but also 4 weak JSB hexagons.

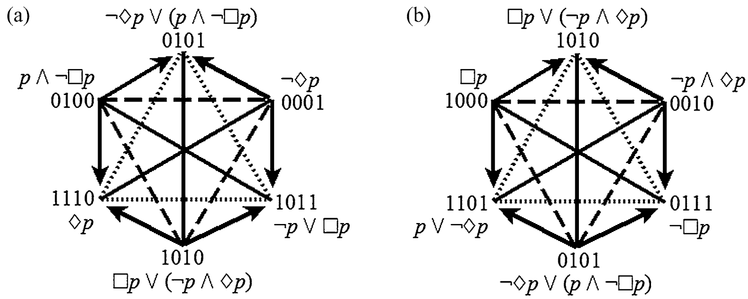


Fig. 10 Two new JSB hexagons for S5

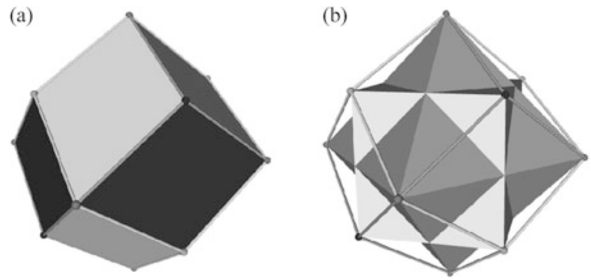


Fig. 11 (**a**) The rhombic dodecahedron, (**b**) the rhombic dodecahedron as the compound of a cube and an octahedron

gram. There have recently been a number of related—albeit subtly different—proposals. Smessaert [25, 26] and Demey [10, 11] make use of a *rhombic dodecahedron*. Moretti [18] and Pellissier [20] make use of a so-called *tetraicosahedron*. Finally, it has recently been discovered that this visualisation issue was already discussed in full detail in the 1960s by P. Sauriol, who made use of a so-called *tetrahexahedron* [22].[12] Of these three, only the rhombic dodecahedron is canonically discussed in the mathematical literature on polyhedra [9] and does equal justice to its cube and octahedron components [27]; see Fig. 11. Furthermore, as was already noted above, the rhombic dodecahedron is geometrically related to both of the three-dimensional Aristotelian diagrams that were discussed in the previous section: Béziau's stellar rhombic dodecahedron is its first *stellation* (recall Fig. 5(b)), while Moretti's cuboctahedron is its *dual* polyhedron (recall Fig. 8(c)). Finally, the rhombic dodecahedron fits naturally in a unified perspective on Aristotelian diagrams and Hasse diagrams [12]. Because of these reasons, we will henceforth use the rhombic dodecahedron as the canonical representation of the logical geometry of S5.

Because its set of 14 formulas is Boolean closed, the rhombic dodecahedron constitutes a natural endpoint in the analysis of the logical geometry of S5. Putting it in terms of bitstrings, the rhombic dodecahedron provides a complete account of the logical geometry of bitstrings *of length 4*: every Aristotelian diagram that can be constructed with bitstrings of length 4, can be embedded inside the rhombic dodecahedron.[13] We are currently devel-

---

[12]The differences between these three visualisations are discussed in more detail in [27, Sect. 2].

[13]Going beyond the rhombic dodecahedron would thus require us to introduce bitstrings of length 5. This can certainly be done, but the Aristotelian diagrams will become exponentially larger. For example, as far as Boolean closed diagrams are concerned, we move from the rhombic dodecahedron (which has $2^4 - 2 = 14$ vertices) to a diagram that has $2^5 - 2 = 30$ vertices. (Recall Footnotes 3 and 4.)

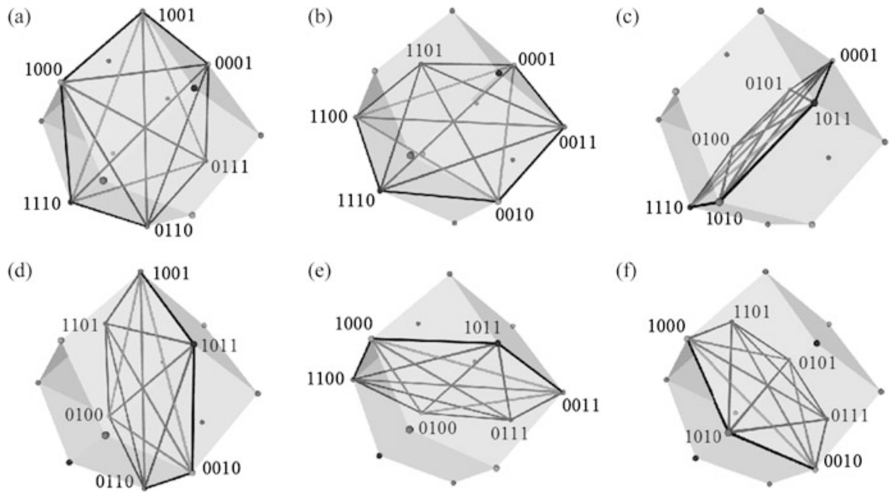


Fig. 12 Six JSB hexagons embedded inside the rhombic dodecahedron

oping a systematic typology of all these diagrams [30]; however, for our current purposes, it suffices to note that each of the six JSB hexagons (recall Figs. 4, 7(b) and 10) can be embedded inside the rhombic dodecahedron; see Fig. 12.

5 Béziau on the Logical Geometry of the Subjective Quantifiers

Over the years, Béziau has transposed his analysis of the logical geometry of S5 to a number of other fields. For example, he has recently shown how one can also construct JSB hexagons for metalogical notions such as theoremhood and consistency [4, 6]. In the second part of this paper, we will focus on another application of his analysis, viz. the subjective quantifiers, which he presented in 2008 at the LNAT 1 conference in Brussels [3], but which has remained unpublished so far.

Béziau's starting point is the observation that the logical geometry of S5 seems analogous to that of a certain set of quantifiers, including the subjective quantifiers *many* and *few*. In more abstract terms, both the formulas from S5 and the lexical field of subjective quantification can be given a semantics in terms of bitstrings of length 4. The exact analogy proposed by Béziau is described in Table 2. The first six rows comprise the standard universal and existential quantifiers and their Boolean combinations; the next eight rows comprise the subjective quantifiers $many_1$ and few_1 and their Boolean combinations with each other and with the first six.[14]

The analogy between S5-formulas and quantifiers in the first six rows is very natural, given the familiar Kripke semantics of modal logic in terms of quantification over possible worlds. For example, the truth of $\Box p$ consists in p being true in *all* (accessible) possible

[14]We will henceforth add a '1' in subscript to the expressions *many* and *few* to refer to the semantic interpretation they receive in Béziau's analysis. Similarly, in the next sections, we will add a '2' in subscript to refer to our alternative analysis.

Table 2 The analogy between S5 and the (subjective) quantifiers

S5-formula	Bitstring	(Subjective) quantifiers
$\Box p$	1000	*all*
$\neg\Box p$	0111	*not all*
$\neg\Diamond p$	0001	*no*
$\Diamond p$	1110	*at least one*
$\Box p \vee \neg\Diamond p$	1001	*no or all*
$\neg\Box p \wedge \Diamond p$	0110	*some*
p	1100	$many_1$
$\neg p$	0011	few_1
$p \wedge \neg\Box p$	0100	$many_1$ *but not all*
$\neg p \wedge \Diamond p$	0010	*at least one but* few_1
$\neg p \vee \Box p$	1011	*all or* few_1
$p \vee \neg\Diamond p$	1101	*no or* $many_1$
$\Box p \vee (\neg p \wedge \Diamond p)$	1010	*all or (at least one but* few_1*)*
$\neg\Box p \wedge (p \vee \neg\Diamond p)$	0101	*no or (*$many_1$ *but not all)*

worlds, while the truth of $\Diamond p$ consists in p being true in *at least one* (accessible) possible world. Note that Béziau explicitly distinguishes between the two existential expressions *at least one* and *some*: for example, 'at least one A is B' does not exclude the possibility that *all* As are Bs, but 'some As are B' does exclude this possibility (i.e. it entails that at least one A is not B). This distinction corresponds to the linguistic distinction between the 'one-sided' and 'two-sided' readings of the existential quantifier [13].[15,16] In terms of the Boolean operators, we have the following equivalences:

$$\textit{some} \equiv \textit{at least one but not all} \qquad (0110 = 1110 \wedge 0111)$$
$$\textit{at least one} \equiv \textit{some or all} \qquad (1110 = 0110 \vee 1000)$$

There is disagreement among linguists whether the two-sided reading of the natural language expression *some* is a matter of semantics or pragmatics [13, 24]; Béziau thus sides with those who take it to be a matter of semantics. Finally, it should be noted that these six quantifiers yield an alternative decoration for the JSB hexagon shown in Fig. 4(a).

As far as the bottom eight rows of Table 2 are concerned, the core of Béziau's analysis consists in treating the subjective quantifiers $many_1$ and few_1 on a par with the so-called 'null-modalities' in S5, namely the formulas p and $\neg p$, which do not contain a modal operator. In S5, we start with a tripartition of logical space into 'necessity' (1000), 'contingency' (0110) and 'impossibility' (0001), and superimpose upon it a bipartition into 'actually true' (1100) and 'actually false' (0011). Keeping in mind the Kripke semantics of modal logic described above, the space of quantification can be tripartitioned by means

[15] In terms of bitstrings, the one-sided reading corresponds to a bitstring that has *one* transition in bit values, i.e. from 1 to 0 or vice versa (e.g. 1110), whereas the two-sided reading corresponds to a bitstring having *two* transitions in bit values (e.g. 0110).

[16] This distinction also applies to the modal operators, where one-sided possibility ($\Diamond p$) is compatible with necessity, but two-sided possibility ($\Diamond p \wedge \neg\Box p$, usually called 'contingency') is not.

of the expressions *all* (1000), *some* (0110) and *no* (0001). Béziau's analysis now superimposes a bipartition by means of the subjective quantifier expressions $many_1$ (1100) and few_1 (0011). The entailments in S5 from the level 1 (L1) notion of 'necessity' (1000) to the level 2 (L2) notion of 'actual truth' (1100) and from the L1 notion of 'impossibility' (0001) to the L2 notion of 'actual falsehood' (0011) get straightforward counterparts in the realm of subjective quantification. More in particular, $many_1$ and few_1 are L2 elements: $many_1$ (1100) is entailed by *all* (1000), whereas few_1 (0011) is entailed by *no* (0001). This accounts for the first two rows of the bottom part of Table 2, which constitute the core of Béziau's analogy between the modalities and the quantifiers.

The remaining six rows are then built by means of the Boolean operators of conjunction and disjunction. The purpose of the first pair of Boolean combinations, i.e. the conjunctions $many_1$ *but not all* and *at least one but* few_1, is to create the L1 elements 0100 and 0010 by excluding the extreme values of the tripartition, namely *all* (1000) and *no* (0001), respectively. The two disjunctions *all or* few_1 and *no or* $many_1$ yield the L3 elements 1011 and 1101. The final two quantifier expressions,[17] *all or (at least one but* few_1*)* and *no or (*$many_1$ *but not all)*, have a layered Boolean structure (in that the top level disjunction has a second disjunct which is itself a conjunction) and correspond to the L2 elements 1010 and 0101, respectively. Recall from Sect. 4 that also in S5, it is precisely these two L2 elements which correspond to the most complex formulas. More in general, the bottom part of Table 2 reveals a very strong parallelism between the formulas of S5 and the subjective quantifiers in terms of lexico-syntactic complexity, i.e. the minimal[18] number of binary connectives they require: the two basic elements (1100 and 0011) contain *no* binary connective, the next four elements (0100, 0010, 1011 and 1101) get *one* binary connective, and the final two (1010 and 0101) have *two* binary connectives.

6 An Alternative Analysis of the Subjective Quantifiers

In this section, we propose an alternative analysis for the logical geometry of the subjective quantifiers based on linguistic considerations.[19] Notice, by the way, that Béziau's own distinction between the two-sided *some* and the one-sided *at least one* provides the perfect starting point for such an alternative analysis. From the perspective of lexicalisation, i.e. the amount of lexical material an expression consists of, *some* is more primitive than *at least one*. The semantic complexity in terms of the levels of the corresponding

[17] Strictly speaking, these two quantifier expressions did not feature in Béziau's presentation [3], but as was argued in Sect. 4, they can straightforwardly be added by taking the Boolean closure of the original set of 12 expressions.

[18] We only look at the *minimal* number of binary connectives, because every semantic value (bitstring) can be expressed in a number of syntactically different ways. For example, the bitstring 0010 can be expressed as $\neg p \wedge \Diamond p$ [$0011 \wedge 1110$] with *one* binary connective, but also as $\neg p \wedge (\Box p \vee (\neg p \wedge \Diamond p))$ [$0011 \wedge 1010$] with *three* binary connectives.

[19] From a linguistic point of view, the semantics of *many* and *few* is notoriously complex. For example, Keenan writes: "we shall largely exclude *many* and *few* from the generalisations we propose since our judgements regarding their interpretations are variable and often unclear" [16, pp. 47–48].

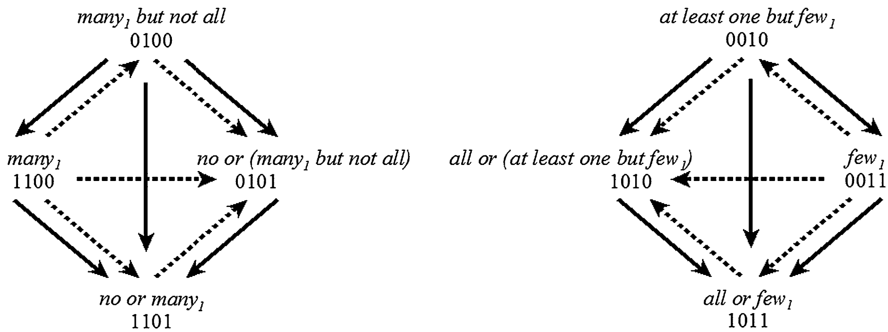


Fig. 13 Semantic and lexical complexity in Béziau's analysis

bitstrings then runs perfectly parallel to this lexical complexity: the L2 bitstring 0110 for *some* is less complex than the L3 bitstring 1110 for *at least one*.[20]

This correlation between lexical and semantic complexity no longer holds, however, with the subjective quantifier expressions in the bottom part of Table 2. First of all, *many*$_1$ is lexically more primitive than *many*$_1$ *but not all*, but the former's L2 bitstring (1100) is semantically more complex than the latter's L1 bitstring (0100). Completely analogously, *few*$_1$ is lexically more primitive than *at least one but few*$_1$, but the former's L2 bitstring (0011) is again semantically more complex than the latter's L1 bitstring (0010). Similar discrepancies can be observed the next level up: *no or many*$_1$ is lexically less complex than *no or (many*$_1$ *but not all)*, but nevertheless the former is L3 (1101) whereas the latter is only L2 (0101). The mismatches between semantic and lexical complexity are visually represented in Fig. 13. The full line arrows represent the semantic complexity increasing from L1 at the top to L3 at the bottom, thus reflecting the entailment or subalternation relation. The dashed line arrows, by contrast, reflect the increase in lexical complexity. In 8 out of the 12 cases, there is a mismatch between the increases in semantic and lexical complexity.[21] In more visual terms, the 'orientation' of the two lattices for semantic complexity in Fig. 13 is *from the top downwards*, whereas that of the two lattices for lexical complexity is *from the outside inwards* (i.e. with *many*$_1$ and *few*$_1$ as their respective starting points).

In view of these considerations, the key property of our alternative analysis is that *many* and *few* are characterised as the L1 bitstrings 0100 and 0010, respectively. As a consequence, *many*$_2$ is incompatible with *all*: the bitstrings 0100 and 1000 are contrary. In contrast, recall that Béziau's *many*$_1$ (1100) is entailed by *all* (1000). This entailment is due to the analogy Béziau draws between modalities and quantifiers: just like $\Box p$ (1000) entails p (1100), he takes *all* (1000) to entail *many*$_1$ (1100). Although the former entailment is beyond any doubt, the latter is more questionable. Consider a situation, for instance, in

[20]If we were to work with bitstrings of length 3, ignoring the subjective quantifiers and focussing on the six quantifier expressions in the top part of Table 2, *some* would be the L1 bitstring 010, whereas *at least one* would be the L2 bitstring 110.

[21]More specifically, three types of mismatches can be distinguished: (i) there is both a semantic and a lexical arrow but they point in opposite directions—e.g. between 1100 and 0100, (ii) there is a semantic arrow but no lexical arrow at all—e.g. between 0100 and 1101, and (iii) there is a lexical arrow but no semantic arrow at all—e.g. between 1100 and 0101.

which the universe of discourse contains three books, all three of which have been read by John. In this situation, the proposition *John has read all books* is obviously true, but the proposition *John has read many books* is very likely to be considered false, for the simple reason that 'three books' do not really count as 'many books'. In other words, *all* need not entail *many*, although on many occasions it will actually do so, of course. In order to reflect this possible absence of entailment, our alternative analysis of the subjective quantifiers assigns a two-sided reading to natural language *many*:[22] whereas Béziau's $many_1$ is the one-sided L2 element 1100 (which has a single transition in bit values), our $many_2$ is the two-sided L1 bitstring 0100 (which has two transitions in bit values).[23] This analysis is further supported by the fact that a lexically complex expression such as *many if not all* exactly allows us to turn the two-sided $many_2$ into a one-sided reading, by incorporating the *all* in the disjunction,[24] thus retrieving the L2 semantics 1100 of Béziau's $many_1$:

$$many_1 \equiv many_2 \textit{ or all}/many_2 \textit{ if not all} \quad (1100 = 0100 \vee 1000)$$

A largely analogous story can now be told for the negative subjective quantifier *few*. For Béziau, the validity of the modal entailment from $\neg\Diamond p$ (0001) to $\neg p$ (0011) carries over to that from *no* (0001) to few_1 (0011). Once again, however, the latter entailment is somewhat problematic. Concluding from the truth of *John has read no books* to that of *John has read few books* runs into conflict with the existential presupposition that seems to accompany the latter proposition: qualifying the amount of books read as 'few' requires that there exists 'at least one' book read. In order to do justice to this intuition, few_2 receives a two-sided L1 analysis which is incompatible with *no*: the respective bitstrings 0010 and 0001 are contrary. Here as well, the analysis is further supported by the existence of lexically complex expressions such as *few if any*. The two-sided few_2 changes into a one-sided reading, by incorporating the *no* in the disjunction,[25] thus recovering the L2 semantics 0011 of Béziau's few_1:

$$few_1 \equiv few_2 \textit{ or no}/few_2 \textit{ if any} \quad (0011 = 0010 \vee 0001)$$

Whereas Béziau's $many_1$ (1100) and few_1 (0011) are contradictories and thus yield a partition of the entire logical space, our own $many_2$ (0100) and few_2 (0010) are merely contraries and yield a more fine-grained partition of the two-sided quantifier *some* (0110):

$$some \equiv many_2 \textit{ or } few_2 \quad (0110 = 0100 \vee 0010)$$

[22]Notice, incidentally, that the difference in subscripts between Béziau's $many_1$ and our $many_2$ nicely reflects this contrast between the *one*-sided and the *two*-sided readings.

[23]Recall Footnote 15.

[24]Using simple propositional reasoning, the expression *many if not all* can be shown to be equivalent to the expression *many or all*. Intuitions differ as to whether an expression of the form *p if not q* should be read as the conditional $\neg p \rightarrow q$ or rather as $\neg q \rightarrow p$, but both readings are equivalent to the disjunction $p \vee q$.

[25]Since *any* can be seen as the negation of *no*, the expression *few if any* is semantically equivalent to *few if not no*, and can thus also be shown to boil down to the disjunction *few or no* (recall Footnote 24). Additional linguistic evidence for this equivalence comes from translational equivalents such as Dutch *weinig of geen* and French *peu ou pas*. Furthermore, even in English the disjunctive semantics of *few or no* is lexicalised, albeit only for abstract and mass nouns, viz. as *little or no*.

Table 3 An alternative to Béziau's analysis of the (subjective) quantifiers

Béziau's analysis	Bitstring	Alternative analysis
all	1000	*all*
not all	0111	*not all*
no	0001	*no*
at least one	1110	*at least one*
no or all	1001	*no or all*
some	0110	*some*
$many_1$	1100	$many_2$ *if not all*
few_1	0011	few_2 *if any*
$many_1$ *but not all*	0100	$many_2$
at least one but few_1	0010	few_2
all or few_1	1011	*all or (*few_2 *if any)*
no or $many_1$	1101	*no or (*$many_2$ *if not all)*
all or (at least one but few_1*)*	1010	*all or* few_2
no or ($many_1$ *but not all)*	0101	*no or* $many_2$

Continuing along these lines, we can calculate all the Boolean combinations of our $many_2$ and few_2, and thereby obtain a Boolean closed set of 14 quantifier expressions (just like in Béziau's analysis). Table 3 presents a comparative overview of Béziau's and our analyses. Note that both analyses make use of the same set of 14 bitstrings, and will thus have exactly the same logical properties (e.g. the same types and numbers of Aristotelian relations); their differences are thus purely a matter of how these bitstrings are mapped onto the concrete natural language expressions. More specifically, we see that both analyses agree on this mapping for the first six expressions, i.e. on the 'ordinary' universal and existential quantifiers (and their Boolean combinations). The differences in the two mappings are thus entirely situated in the final eight expressions, i.e. those involving the subjective quantifiers *many* and *few*.

We have argued for our alternative analysis by appealing to logical intuitions (e.g. concerning the entailments between *all* and *many*) as well as lexicalisation patterns in natural languages (e.g. *little or no*, *weinig of geen*, *peu ou pas*, etc.). Additionally, our alternative analysis turns out to avoid the mismatches between semantic and lexical complexity in Béziau's analysis, which were shown in Fig. 13. The corresponding diagram for our own analysis is shown in Fig. 14: full line arrows still represent increases in semantic complexity (entailment/subalternation), while dashed line arrows represent increases in lexical complexity. Recall that in Béziau's analysis, there is a mismatch between semantic and lexical complexity in 8 out of the 12 cases. In the alternative analysis, however, there are no mismatches whatsoever: the lattices for semantic and lexical complexity share a single 'orientation', viz. *from the top downwards*.[26]

[26]Note that there is no horizontal semantic arrow between the bitstrings 1100 and 0101, since neither of them entails the other one. There is no horizontal lexical arrow between them either, since their corresponding quantifier expressions have the same degree of lexical complexity (viz. a single Boolean operator). Because of this twofold absence, the correlation between semantic and lexical complexity is preserved in this case as well. Similar remarks apply to the case of 1010 and 0011.

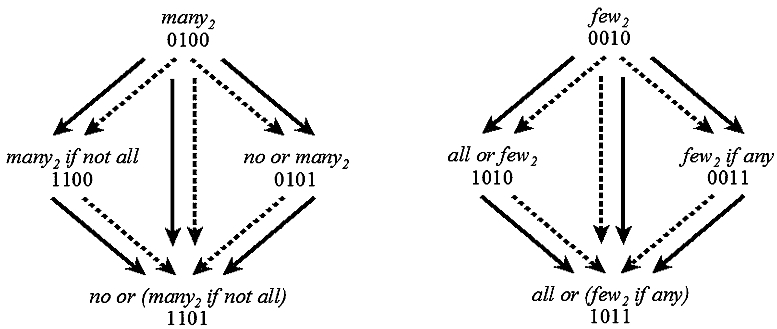


Fig. 14 Semantic and lexical complexity in the alternative analysis

7 Aristotelian Diagrams for the Subjective Quantifiers

In this section, we compare our alternative analysis of the subjective quantifiers with that of Béziau's from the perspective of logical geometry, i.e. in terms of the various Aristotelian diagrams they give rise to. Considering Table 3 from Sect. 6, we already noted that both analyses entirely agree with each other on the first six bitstrings. Furthermore, in Sect. 5 these six bitstrings were shown to yield a JSB hexagon, which was displayed in Fig. 4(a).

The differences between both analyses are thus entirely situated within the next eight bitstrings of Table 3, i.e. the subjective quantifiers *many* and *few* (and their Boolean combinations). The diagrams in Figs. 13 and 14 can be seen as partial Aristotelian diagrams for Béziau's and our analysis of the subjective quantifiers, respectively, in the sense that the full line arrows in these diagrams represent subalternation, which is one of the four Aristotelian relations. One might consider turning these diagrams into full-fledged Aristotelian diagrams by adding the other Aristotelian relations, but this leads to a suboptimal visualisation. For example, the bitstrings 1100 and 0011 are contradictory to each other, and similarly for 0101 and 1010, but since these four bitstrings are collinear—they lie on a single (horizontal) line—, the 'short' contradiction edge between 0101 and 1010 would not be visually distinguishable from the 'long' contradiction edge between 1100 and 0011.

We therefore turn to an alternative visualisation, which has the shape of a convex octagon. This diagram was first used as an Aristotelian diagram by the medieval logician John Buridan, and is therefore canonically called the *Buridan octagon* [14, 21, 27, 29]. Figure 15 shows Buridan octagons for Béziau's analysis and our own analysis of the subjective quantifiers.[27] These centrally symmetric Aristotelian diagrams each consist of 4 contradictions, 5 contrarieties, 5 subcontrarieties and 10 subalternations (the latter correspond to the 10 full line arrows in Fig. 14). Furthermore, note that there are no Aristotelian relations at all between the L2 bitstrings 1100, 1010, 0011 and 0101 (except for the diagonals of contradiction 1100—0011 and 1010—0101, of course). For Béziau's analysis (Fig. 15(a)), these four bitstrings correspond to the two lexically most primitive and the

[27]It should be emphasised that we take the term 'Buridan octagon' to refer to (the visualisation of) a certain constellation of Aristotelian relations, rather than to the particular form or content matter of the formulas involved. As a matter of fact, Buridan's own use of his octagon was in describing the logical geometry of first-order modal logic, rather than that of the subjective quantifiers.

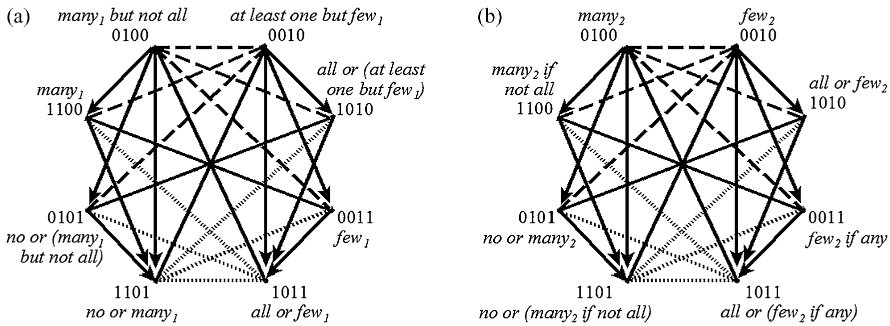


Fig. 15 Buridan octagons for the subjective quantifiers: (**a**) Béziau's analysis, (**b**) our alternative analysis

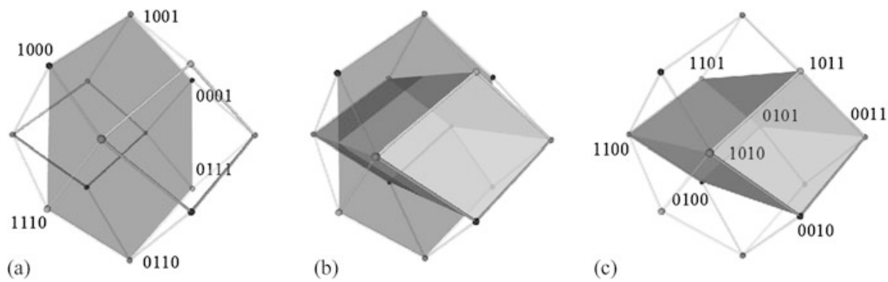


Fig. 16 JSB hexagon (**a**)–(**b**) and Buridan octagon (**b**)–(**c**) embedded in RDH

two lexically most complex subjective quantifier expressions. By contrast, for our own analysis (Fig. 15(b)) these four bitstrings correspond to the quantifier expressions which have intermediate lexical complexity. This is an immediate consequence of the fact that they are bitstrings of level 2, i.e. of intermediate semantic complexity, in combination with the strict correlation between semantic and lexical complexity in our analysis (recall Fig. 14).

As the starting point for a third visualisation, note that there exists a logical complementarity between the 6 formulas in the JSB hexagon and the 8 formulas in the Buridan octagon. In terms of Table 3, its upper and lower parts jointly constitute a set of $6 + 8 = 14$ formulas that is Boolean closed.[28] We have recently shown that this *logical complementarity* between sets of formulas corresponds to a visually appealing *geometrical complementarity* between Aristotelian diagrams [27, 29]. To see this, consider the three-dimensional Aristotelian diagram for the entire Boolean closed set of 14 formulas that was introduced in Sect. 4, i.e. the rhombic dodecahedron. Figure 16(a) shows how the JSB hexagon for the 6 'ordinary' quantifiers is embedded into the rhombic dodecahedron. The 8 remaining vertices constitute a 'squeezed' cube, which we have elsewhere called a *rhombicube* [27, 29]. Figure 16(c) shows how this rhombicube is embedded into the rhombic dodecahedron. Finally, Fig. 16(b) shows the geometrical complementarity between the JSB hexagon of 'ordinary' quantifiers and the rhombicube of subjective quantifiers.

[28] I.e. the set is essentially a Boolean algebra with its top and bottom elements left out (recall Footnotes 3 and 4).

It should be noted that from a strictly logical perspective, there is no difference between the Buridan octagon and the rhombicube: they are merely two different visualisations of a single configuration of 8 formulas and the Aristotelian relations between them. Each of these visualisations has its own advantages and disadvantages. On the one hand, the Buridan octagon is—just like the JSB hexagon—a well-known and canonical diagram, whose visual apprehension is probably facilitated by its two-dimensional nature. On the other hand, the rhombicube stands in a clear visual-geometrical relation of complementarity to the JSB hexagon, and thus better reflects the underlying logical complementarity.

In sum, then, our analysis agrees with that of Béziau's with respect to the JSB hexagon of 'ordinary' quantifiers, but disagrees with respect to the Buridan octagon/rhombicube of subjective quantifiers. In other words, the two analyses share a JSB hexagon, but they complement it with different rhombicubes.

8 Conclusion

In this paper, we have evaluated Béziau's Aristotelian diagrams for modalities and quantifiers from a logico-linguistic perspective, and shown how they relate to our framework of logical geometry. In the first main part, we have considered his Jacoby–Sesmat–Blanché stars and stellar rhombic dodecahedron for a set of 12 S5-formulas, and discussed the visualisation of its Boolean closure by means of a rhombic dodecahedron. In the second main part, we have discussed his proposal to transpose his results from S5 to the lexical field of subjective quantification with *many* and *few*, and proposed an alternative analysis based on a number of logical and linguistic considerations. In the final part, we have compared our own analysis with that of Béziau's, making use of a number of two- and three-dimensional Aristotelian diagrams, such as the JSB star/hexagon, the Buridan octagon, and the rhombicube, all of which can be embedded inside the rhombic dodecahedron.

In ongoing work, we are developing a systematic account of the rhombic dodecahedron and its various (families of) subdiagrams [30]. We already have a firm grasp of how the family of JSB hexagons is embedded inside the rhombic dodecahedron; however, in future research we also intend to explore in more detail the embeddings of larger diagrams, such as Béziau's stellar rhombic dodecahedron. Additionally, given the notorious complexity of the subjective quantifiers, it will be interesting to investigate how the two analyses discussed above will hold up under further linguistic scrutiny.

Acknowledgement We thank Dany Jaspers, Alessio Moretti, Fabien Schang and Margaux Smets for their comments on earlier versions of this paper. The second author gratefully acknowledges financial support from the Research Foundation—Flanders (FWO).

References

1. Béziau, J.-Y.: New light on the square of oppositions and its nameless corner. Log. Invest. **10**, 218–232 (2003)
2. Béziau, J.-Y.: Paraconsistent logic from a modal viewpoint. J. Appl. Log. **3**, 7–14 (2005)
3. Béziau, J.-Y.: Some—geometrical representations of the quantifiers. Talk presented at Logic Now and Then (Brussels) (2008)
4. Béziau, J.-Y.: The power of the hexagon. Log. Univers. **6**, 1–43 (2012)
5. Béziau, J.-Y.: The new rising of the square of opposition. In: Béziau, J.-Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition, pp. 6–23. Springer/Birkhäuser, Berlin (2012)

6. Béziau, J.-Y.: The metalogical hexagon of opposition. Log. Invest. **10**, 111–122 (2013)
7. Blanché, R.: Quantity, modality, and other kindred systems of categories. Mind **61**, 369–375 (1952)
8. Blanché, R.: Structures Intellectuelles. Essai sur l'Organisation Systématique des Concepts. Vrin, Paris (1969)
9. Coxeter, H.: Regular Polytopes. Courier Dover, New York (1973)
10. Demey, L.: Structures of oppositions for public announcement logic. In: Béziau, J.-Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition, pp. 313–339. Springer/Birkhäuser, Berlin (2012)
11. Demey, L.: Believing in logic and philosophy. PhD thesis, KU Leuven (2014)
12. Demey, L., Smessaert, H.: The relationship between Aristotelian and Hasse diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrammatic Representation and Inference. Lecture Notes in Computer Science, vol. 8578, pp. 213–227. Springer, Berlin (2014)
13. Horn, L.: A Natural History of Negation. University of Chicago Press, Chicago (1989)
14. Hughes, G.: The modal logic of John Buridan. In: Corsi, G., Mangione, C., Mugnai, M. (eds.) Atti del Convegno Internzionale di Storia della Logica, le Teorie delle Modalità, pp. 93–111. CLUEB (1987)
15. Jacoby, P.: A triangle of opposites for types of propositions in Aristotelian logic. New Scholast. **24**, 32–56 (1950)
16. Keenan, E.: The semantics of determiners. In: Lappin, S. (ed.) The Handbook of Contemporary Semantic Theory, pp. 41–63. Blackwell, Oxford (1997)
17. Luke, D.: Stellations of the rhombic dodecahedron. Math. Gaz. **41**, 189–194 (1957)
18. Moretti, A.: The geometry of logical opposition. PhD thesis, University of Neuchâtel (2009)
19. Moretti, A.: The critics of paraconsistency and of many-valuedness and the geometry of oppositions. Log. Log. Philos. **19**, 63–94 (2010)
20. Pellissier, R.: "Setting" n-opposition. Log. Univers. **2**, 235–263 (2008)
21. Read, S.: John Buridan's theory of consequence and his octagons of opposition. In: Béziau, J.-Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition, pp. 93–110. Springer/Birkhäuser, Berlin (2012)
22. Sauriol, P.: Remarques sur la théorie de l'hexagone logique de Blanché. Dialogue **7**, 374–390 (1968)
23. Sesmat, A.: Logique II. Les Raisonnements. Hermann, Paris (1951)
24. Seuren, P.: The Logic of Language. Language from Within, vol. II. Oxford University Press, Oxford (2010)
25. Smessaert, H.: On the 3D visualisation of logical relations. Log. Univers. **3**, 303–332 (2009)
26. Smessaert, H.: Boolean differences between two hexagonal extensions of the logical square of oppositions. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrammatic Representation and Inference. Lecture Notes in Computer Science, vol. 7352, pp. 193–199. Springer, Berlin (2012)
27. Smessaert, H., Demey, L.: Logical and geometrical complementarities between Aristotelian diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrammatic Representation and Inference. Lecture Notes in Computer Science, vol. 8578, pp. 246–260. Springer, Berlin (2014)
28. Smessaert, H., Demey, L.: The unreasonable usefulness of bitstrings in logical geometry. Talk presented at the Fourth World Congress on the Square of Opposition (Vatican) (2014)
29. Smessaert, H., Demey, L.: La géometrie logique du dodécaèdre rhombique des oppositions. In: Ben Aziza, H., Chatti, S. (eds.) Le Carré et ses Extensions: Aspects Théoriques, Pratiques et Historiques. Presses Universitaires, Tunis (2014)
30. Smessaert, H., Demey, L.: The Logical Geometry of the Aristotelian Rhombic Dodecahedron. Book Manuscript (2015)

H. Smessaert (✉)
Research Group on Formal and Computational Linguistics, KU Leuven, Blijde Inkomststraat 21, 3000 Leuven, Belgium
e-mail: Hans.Smessaert@arts.kuleuven.be

L. Demey
Center for Logic and Analytic Philosophy, KU Leuven, Kardinaal Mercierplein 2, 3000 Leuven, Belgium
e-mail: lorenz.demey@hiw.kuleuven.be

On Metalogical Relativism

Vladimir L. Vasyukov

Abstract The conception of logical pluralism claims that there is not one true logic but there are many. The conception of metalogical relativism is based on the assumption that there is not one correct answer as to whether a given argument is deductively valid, but there are many—many non-classical answers depending of which non-classical metalogic we exploit: intuitionistic, relevant, quantum, many-valued, etc. Since this leads to the interplay between logics and metalogics the question arises: What is the nature of this interplay? The Universal Logics approach gives us hints at some answers to this question but at the expense of the exploitation of different combination of non-classical logical systems leading to the transition from metalogical pluralism to metalogical monism. The only problem in this case is the impossibility of the exploitation of an infinite combination of non-classical systems. There are also some semantic keys to the issue under consideration which are connected with the problem of the interplay of classical and non-classical universes: non-classical logics would be interpreted in the classical universe, and vice versa, the classical logic would be interpreted in non-classical universes.

Keywords Logical relativism · Logical pluralism · Logical monism · Metalogical pluralism · Non-classical metalogics · Universal metalogics

Mathematics Subject Classification 03A05 · 03B62 · 97E30

1 Introduction

The recent situation in logics features stable proliferation of non-classical logical systems, and this process, in all appearances, is irreversible by its nature. During more than 2000 years, the scholars construed Aristotelian and Stoic logic as only one possible logic; modern classical logic is a continuation of this tradition being different just by its tools. The emergence of non-classical logics seriously stroke the logical investigations while forcing to revaluate and cast doubt on many results which at that time were taken for granted. As a consequence, the point of view of the existence of not one but many true logics prevails in the modern philosophy of logics. This point of view is well-known under the name of logical pluralism.

The results of the project "Rule Following: Reasoning, Reason, Rationality", carried out within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) in 2014, are presented in this work.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_24

By the way, already Ch. Peirce thought that there are a lot of logics which can be determined, improved and exploited for examining one another (and even themselves)—everything depends on the goals of an investigation. Unlike him G. Frege endorsed only one unique calculus (*Begriffsschrift*) because for him all humans share one kind of thought. And this is the origin of Frege's formal language (*Formelsprache*)—in his opinion, it is not an additional language, but improved and clarified version of the ordinary one.

In the 20th century, the development of non-classical logics at the early stage featured the peaceful coexistence of logical pluralism and logical monism within the framework of one and the same philosophical community. A typical example is the Lvov–Warsaw logic and philosophical school whose two prominent representatives—St. Leśniewski and J. Łukasiewicz—adhered to the polar position in the quest for the unique foundation of logic. According to Leśniewski, the difference between mathematical systems and arbitrary deductive systems lies in that mathematical systems do not contradict 'logical intuitions' which, in turn, cannot describe the world at will. They should comply with common logic—the true proper logic of the world. The best, or rather, the only possible way this logic can be characterized as the classical one is as bivalent and extensional. For this reason Lesniewski, in particular, showed no interest in many-valued logics (and in other non-classical ones, in general).

As to Łukasiewicz, in 1936 he shared Leśniewski's opinion and wrote that the one and only logical system is proper logic of the actual world the same way as the one and only system of geometry is proper geometry, but in 1937 he claimed something contrary. Then his point was that all logical systems under the assumptions made are necessarily true. We just confirm the ontological commitments hidden somewhere in the foundation of logic while verifying by facts the consequences of these commitments. Finally, at the end of his life, in 1952, he came to the conclusion that there is no way to recognize which of n-valued logical system ($n \geq 2$) can be more relevant and implemented. The more applicable and rich the system is, the more truth values it has.

R. Carnap in his work 'The Logical Syntax of Language' [4] wrote that logic plays a big role in philosophical investigation, but it is not prescribed that in such studies we should only deal with one logic. He formulates the pluralistic 'Principle of Tolerance' saying that it is none of our business to forbid, but we need to obtain logical conclusions. According to him, there are no moral prohibitions, and everybody can yield and use that logic which is successful in his studies. Meanwhile, we need to take into account that we are interested not in the choice of logics but in the choice of the respective language.

Recently, despite the proliferation of non-classical logical systems' features in modern situations (which could be considered as an 'empirical' justification of the legality of logical pluralism), the debates on the nature of logical pluralism are still in progress. And the main problem in this connection is the question of the nature of the relationship between these logics: Are they rivals or do they form one friendly community? Scholars as Graham Priest regard all non-standard logics (i.e. intuitionistic, many-valued and quantum, relevant and paraconsistent, conditional and free) as correct, and their presence tells us that logic is not a collection of accepted truths but a discipline in which theories pretending to be valid are rivals. And at the same time, we need to distinguish theoretical logic, applied logic and informal one for their areas of competition are separate.

2 Logical Pluralism, Logical Monism, Logical Relativism

And what is the nature of logical pluralism from the point of view of modern logic? Carnap's 'Principle of Tolerance' which gives carte blanche to inviting and constructing logical systems in practice turns out to be not entirely pertinent and leading to some ambiguity. G. Restall [14] shows that Carnap's 'tolerance' reducing pluralism to an opportunity of choice of formal languages is incorrect since there are different, equally 'good' logical consequence relations in each language. So we should consider not the relationship between languages but the relationship between different kinds of logical consequence relation.

The latter is very important because, according to J.C. Beall and G. Restall [1, p. 475], not long ago logic was dominated by the Frege–Russell picture which treats logical truth as the primary character and consequence as secondary. Maybe this explains the failure of Carnap's principle as an argument for logical pluralism. The contemporary picture reverses the cast: consequence is the leading character. But what is then logical consequence? What does it mean that a conclusion A follows from premises Σ? Beall and Restall suppose that, within the framework of the tradition to which almost everyone subscribes, the nature of logical consequence is captured in the following principle:

(**V**) A conclusion A follows from premises Σ if and only if any case in which each premise in Σ is true is also a case in which A is true. Or equivalently, there is no case in which each premise in Σ is true, but in which A fails to be true.

This principle allows introducing validity as it was done by R. Jeffrey: 'A valid argument is one whose conclusion is true in every case in which all its premises are true. Then the mark of validity is absence of counterexamples, cases in which all premises are true but the conclusion is false' [7, p. 1].

But it is notable that Jeffrey ends his definition with the remark that difficulties in applying this definition arise from difficulties in canvassing the cases mentioned in it. These difficulties should be apprehended more than seriously for they are the source of logical pluralism having his effect on the whole contemporary logics and being the core of most recent studies.

This specific version of logical pluralism comes with three tenets:

1. The pretheoretic (or intuitive) notion of consequence is given in (**V**).
2. A logic is given by a specification of the cases to appear in (**V**). Such a specification of cases can be seen as a way of spelling out truth conditions of the claims expressible in the language in question.
3. There are at least two different specifications of cases which may appear in (**V**).

In Point 2, the most transparent treatment of cases can be obtained by identifying them with possible worlds so popular in modern logics. Then our clauses for truth in a case, or truth in a world, will look like this:

- $A \wedge B$ is true in a possible world w iff A is true in w and B is true in w;
- $A \vee B$ is true in a possible world w iff A is true in w or B is true in w;
- $\neg A$ is true in a possible world w iff A is not true in w;
- $\forall x A(x)$ is true in a possible world w iff for each object b in w, $A(b)$ is true in w;
- $\exists x A(x)$ is true in a possible world w iff for some object b in w, $A(b)$ is true in w.

However, in a more standard way and in the language of the first-order logic, the cases would be treated within the framework of the so-called Tarskian models. They are the structures that comprise the following constructions:

1. A nonempty set D, the domain; and
2. A function I, the interpretation, satisfying the following conditions:
   (a) $I(E)$ is an element of D, if E is a name (in the given language);
   (b) $I(E)$ is a set of ordered n-tuples of D-elements, if E is an n-place predicate.

Then we exploit a model M to interpret the language.

- If α is an assignment of D-elements to variables, then $I_\alpha(x) = \alpha(x)$. If a is a name, $I_\alpha(a) = I(a)$;
- $F(t_1, \ldots, t_n)$ is true in M, α iff $\langle I_\alpha(t_1), \ldots, I_\alpha(t_n)\rangle \in I_\alpha(F)$;
- $A \wedge B$ is true in M, α iff A is true in M, α and B is true in M, α;
- $A \vee B$ is true in M, α iff A is true in M, α or B is true in M, α;
- $\neg A$ is true in M, α iff A is not true in M, α;
- $\forall x A$ is true in M, α iff A is true in M, α' for each x-variant α' of α;
- $\exists x A$ is true in M, α iff A is true in M, α' for some x-variant α' of α.

These conditions give us validity of arguments in formal language by way of (**V**). An argument is valid if and only if in every model in which the premises are true, so is the conclusion. We will call this account the Tarskian account of validity.

If we will consider, for example, situations instead of models and worlds, then we obtain the definition of consequence and validity in relevant logics. If cases are treated as computational constructions (proofs), then we will get intuitionistic logics, and so on. Clearly, some conditions and definitions should be changed. For example, situational condition for negation will be as follows:

- $\neg A$ is true in a situation s iff A is not true in s' for any s' where sCs'.

Here sCs' means that s is compatible with s' (i.e. C is a binary relation of compatibility). In other words, the negation $\neg A$ is true in s' only when all situations in which A is true are incompatible with s. In this case, the situational reading of (**V**) will be

- The argument from Σ to A is relevantly valid if in any model, in any situation in which all premises in Σ are true, so is A.

For intuitionistic logic there is a Kripke-type semantics in which truth is relativized to points (with model constructions) which are partially ordered by strength (written '$\sqsupseteq$'). In particular, truth conditions for implication and negation are as follows:

- $A \supset B$ is true in c iff for any $d \sqsupseteq c$, if A is true in d then so is B;
- $\neg A$ is true in c iff A is not true in d for any $d \sqsupseteq c$.

So, a construction proves $A \supset B$ if and only if when combined with any construction for A you have a construction for B. The condition for negation becomes clear if we take into account that $\neg A$ should be defined as $A \supset\bot$ and for $\bot$ the construction does not exist. The respective version of (**V**) for the case of constructions will be:

- The argument from Σ to A is *intuitionistically* valid if in any model, in any construction in which all premises in Σ are true, so is A.

Emerging intuitive objection is due to the uncertainty of the limits of such interpretation of cases. It turns out that there are a lot of consequence relations generating different logics and that none of these, in any objective, universal sense, is better than the others. They are equal in rights, so does it not follow that 'anything goes'?

Logical pluralists usually answer that the goal of their studies is exactly establishing of the fact of logical pluralism and clarification of respective situation. Many non-classical systems can be challenged by logic, philosophical and technical reasons as well, but it adds nothing to the striking fact of the possibility of existence of many such systems. As a matter of fact, the disagreement arises just when the task of your logical system is formulated, when we have established what our system should do.

A more serious argument against pluralism was proposed by G. Priest [11, p. 203] while considering the hypothetical case of two equally good definitions of deductive validity K_1 and K_2 where β follows from α in case of K_1 but not of K_2, and we know that α is true. Were β true, would the truth of β be determined by the information adduced? In case of K_1, this is evident according to (**V**) but in case of K_2 we have a problem here. If one considers the truth as an absolute and not a relative notion and takes into account that validity does not mean to be true in general (i.e. in case of K_2 one does not assert that β is false) then it turns out that we should accept β to be true which leads to that K_1 is informatively preferable over K_2. Thus, the logic corresponding to K_1 exceeds the logic corresponding to K_2.

The further analysis of the situation arising in this connection leads S. Read [13] to the conclusion that the shortcoming of such a pluralistic approach is the acceptance of the classical metatheory since in this case, due to formalization of (**V**) as

$$(\mathbf{V}_{\supset})\quad \varSigma \vdash_{\supset} A \text{ iff } (\forall w)((\forall B \in \varSigma) w \models B \supset w \models A),$$

we have that from $\alpha \vdash_{\supset} \beta$ and from α being true does not follow that β is true. As an example one can take the case of disjunctive syllogism ($A \vee B, \neg A \vdash B$) which is well-known (so-called Lewis argument, cf. [12, § 2.6]) leading to *Ex Falsum Quodlibet* not agreeing with (**V**). Actually, let us consider the following inference (cf. [2]):

$$\dfrac{\dfrac{\dfrac{A \wedge \neg A}{A}\,(\wedge\mathrm{E})}{A \vee B}\,(\vee\mathrm{I}) \qquad \dfrac{A \wedge \neg A}{\neg A}\,(\wedge\mathrm{E})}{B}\,(\mathrm{DS})$$

It turns out that the disjunctive syllogism (by means of an elimination of conjunction and an introduction of disjunction) leads to the strange conclusion that everything (B in our case) follows from contradiction ($A \wedge \neg A$). It seems to be wrong, especially for advocates of the relevant logic.

But if we replace the classical metaimplication with the relevant one then the respective modification of (**V**)

$$(\mathbf{V}_{\rightarrow})\quad \varSigma \vdash_{\rightarrow} A \text{ iff } (\forall w)((\forall B \in \varSigma) w \models B \in w \models A)$$

warrants us the correct preservation of truth and correct account of validity together with the classical case. From here Read concludes that there is only one true logics, and this would be just the relevant logic with the relevant metalanguage (i.e. relevant metalogics) since it is exactly the constituent of the request for truth preservation from premises to conclusion: if the conclusion really logically follows from the premises then they should be relevant to the conclusion.

At first sight, Read's point of view seems very attractive especially since as J.-Y. Beziau denotes "Logical pluralism ... is not a general theory of logics. The logical pluralist does not make the distinction between reasoning and the theory of reasoning; both are put in the same bag" [3, p. 9]. But the situation turns out to be more complicated.

Roy Cook writes: "One is a relativist about a particular phenomenon X if and only if one thinks that the correct account of X is a function of some distinct set of facts Y. Thus, relativism about X amounts to acceptance of the following schema: The correct account of X is relative to Y" [5, pp. 492–493]. The presence of classical, intuitionistic and relevant metalogics in our considerations could be qualified exactly as relativism. Actually, the correct account of the argument from Σ to A in different versions of (**V**) is relative to one of this metalogic. Again, if we take into account that relativism about some phenomenon and pluralism about this phenomenon are often run together in discussions then pluralism of our situation should be specified according to Cook even as a *dependent pluralism* (an account of X is a dependent pluralism if it involves there being multiple 'correct' accounts of X, but this multiplicity is the result of an underlying relativism—otherwise it is a simple pluralism) (see [5, p. 493]).

At the same time, the shortcoming of relativism as such is that "relativism, by contrast with pluralism, excludes ways of choosing between or ranking positions differentially. It takes such form as that all theories are equally good, or more weakly, that there is no way of choosing between them" [17, p. 81]. Nevertheless, speaking of logics instead of theories, relativism does, like pluralism, envisage a bundle of logical systems while logical monism, by contrast with both, does not, but would eliminate other logical systems since according to simple logical monism there is just the one correct logic.

It is interesting to note that also one more position—logical eclecticism—is possible, about which, paraphrasing R. Sylvan, one can say that it involves exclusive selection, so it too is monistic; but it includes, as logical pluralism may, selection from the doctrines of other logics (see [17, p. 82]). Of course, unlike monism, eclecticism presupposes a pluralistic setting as selection basis but patently an inadequate setting, as its selection basis, on the one hand, presupposes not a limited set of logical systems and, on the other hand, there are, in principle, no rules or prescriptions for involving right logical systems.

3 An Issue of Non-classical Metalogics

The arising question concerns the following circumstance. The general scheme of (**V**) would be as follows:

$(\mathbf{V}_{\Rightarrow})\quad \Sigma \vdash_{\Rightarrow} A \text{ iff } (\forall w)((\forall B \in \Sigma) w \vDash B \Rightarrow w \vDash A)$

Read remarks that, according to Beall and Restall, we obtain different variants of '$\vdash$' by varying the range of 'w'—cases may be worlds, constructions or situations, for example. But the range of 'w' should be universal, and therefore different theories of consequence rather result from varying the interpretation of '$\Rightarrow$'. In the classical metalogic, there is really only one possibility for '$\Rightarrow$', namely, material implication '$\supset$'. In the relevant metalogic, there are two possibilities since relevant logic clearly distinguishes a classical implication from a relevant one. Strictly speaking, Read's monism is based on the second choice—we do not have more possibilities in our account.

However, should metalogic necessarily be either classical or relevant? This point of view is no longer generally accepted. The same G. Priest considering Tarski's theory of truth and his T-construction writes that "sometimes it is said that Tarskian theory must be based on classical logics: this logic is required for the construction to be performed. Such a claim is just plain false. It can be carried out in intuitionistic logics, paraconsistent logics, and, in fact, most logics" [11, p. 45].

Thus, we need to investigate the problem of non-classical metalogics emerging in view of the metalanguage formulation within the definition of a logical consequence relation. For the sake of simplicity, ($\mathbf{V}_{\Rightarrow}$) could be verbally paraphrased as follows: from A follows B if and only if 'A is true' implies 'B is true'. The word 'implies' points to metalogical implication and, following Read in this case, we can reformulate this definition as 'from A relevantly follows B iff 'A is true' relevantly implies 'B is true' '. Hence, we obtain relevance logics on the metalevel, too. But then we are tempted to consider the status of the following formulations: 'from A intuitionistically follows B iff 'A is true' intuitionistically implies 'B is true' ', 'from A quantum logically follows B iff 'A is true' quantum logically implies 'B is true' ', etc.

Actually, what does it mean 'from A intuitionistically follows B iff 'A is true' intuitionistically implies 'B is true' '? Since for intuitionistic logic there is a semantic of computational constructions (proofs) then validity here means truth in all metaconstructions, i.e. cases here mean metaconstructions linked with the specific ordering relation. Thus (**V**) (now the metaprinciple) is attached to truth in 'intuitionistic' metaconstructions whose peculiarity is determined by the semantics of intuitionistic metalogic.

The same way ' 'A is true' quantum logically implies 'B is true' ' is determined by the nature of 'quantum' cases which are subsets of the set of states of a quantum system. Quantum logics is featuring the lack of implication which, e.g. in case of R. Goldblatt's system of quantum orthologic [6], is replaced with 'deducibility' whose interpretation coincides with the ordering relation (the set-theoretical inclusion) on the orthomodular lattice of state subsets. In this case, in ($\mathbf{V}_{\Rightarrow}$) metaconnective '$\Rightarrow$' is precisely the similar connective determining truth of respective formulas.

But what is interesting is that in a paradoxical way Read's monistic argument for relevant logics can be strengthened while accounting for the possibility of 'varying' the metaconnective '$\Rightarrow$'. If we advert to the semantic of relevant logic **R** with the ternary accessibility relation then, as it is well-known, we obtain classical logic within the framework of **R** while extending the system of semantic postulates with the postulate $0 < a$ [15]. Thus if one considers the set K of situations having this additional property then ($\mathbf{V}_{\supset}$) is obtained from ($\mathbf{V}_{\rightarrow}$) by restricting quantification on w, i.e. it is given by the principle

$$(\mathbf{V}^{\rightarrow}_{\supset})\quad \Sigma \vdash_{\supset} A \text{ iff } (\forall w \in K)((\forall B \in \Sigma) w \vDash B \rightarrow w \vDash A)$$

Moreover, if the same postulate is added to the system of postulates for positive relevant logics (without negation) then as a result we obtain positive intuitionistic logic (cf. [15]). By restricting the respective quantifier to some set P of situations having required property, we obtain relevant-intuitionistic version of the principle (**V**) as

$$(\mathbf{V}^{\rightarrow}_{\supset_{\mathrm{Int}}})\quad \Sigma \vdash_{\supset_{\mathrm{Int}}} A \text{ iff } (\forall w \in P)((\forall B \in \Sigma) w \vDash B \rightarrow w \vDash A)$$

On the other hand, one can obtain a version of ($\mathbf{V}_{\rightarrow}$) on the ground, for example, of quantum logics. A. Kron, Z. Marič and S. Vujosevič [9] have shown that an algebraic structure of relevant logics (in particular, the system **R**) can be obtained by way of introducing on orthomodular lattice an additional operation $\rightarrow$ and at the same time if $\vdash_R A \rightarrow B$ then $v(A) \leq v(B)$ where v is an evaluation on orthomodular lattices and $\vdash_R$ is the sign of deducibility of the system R. Here A, B should strictly be the formulas without the sign of implication. But the axiom of distributiveness should be eliminated by replacing it with $A \wedge (\neg A \vee (A \wedge B)) \rightarrow B$. Thus, if one restricts quantification on w with the situations–states forming an orthomodular lattice **OML** then a quantum-relevant version of (**V**) would be obtained:

$(\mathbf{V}^{\Rightarrow_Q}_{\rightarrow_R})$ $\Sigma \vdash_{\rightarrow_R} A$ iff $(\forall w \in P)((\forall B \in \Sigma) w \vDash B \Rightarrow_Q w \vDash A)$

At the first glance, it seems that $(\mathbf{V}^{\rightarrow}_{\supset_{\mathrm{Int}}})$ and $(\mathbf{V}^{\Rightarrow_Q}_{\rightarrow_R})$ are simply the consequence of the violation of universality of the area of w (in both cases we limited the area of quantification on w). But nothing forces us in the case of $(\mathbf{V}^{\Rightarrow_Q}_{\rightarrow_R})$ to refuse treating w in general as cases considering states of a quantum system instead of situations.

Let us consider one more example of obtaining a ‘mixed’ principle (**V**). For the system of Łukasiewicz’s infinite-valued logic $Ł_{\aleph_0}$ in [18], the semantics of possible worlds with the ternary accessibility relation was yielded. In this semantics, to every formula truth values were assigned from the usual matrix [0, 1] of Łukasiewicz’s logic, and the postulates for ternary relation coincided with those for the relevant logic **R**, but possible worlds were linearly ordered in respect to the binary ordering $<$ obtained by means of the definition $a < b =_{\mathrm{def}} ROab$. If we restrict possible worlds to those in which formulas have values 0 or 1 then for such a set of possible worlds LW the $Ł_{\aleph_0}$-relevant version of (**V**) will be reduced to the following principle:

$(\mathbf{V}^{\rightarrow_L}_{\rightarrow_R})$ $\Sigma \vdash_{\rightarrow_R} A$ iff $(\forall w \in LW)((\forall B \in \Sigma) w \vDash B \rightarrow_L w \vDash A)$

Moreover, enriching the list of semantic postulates with the postulate $0 < a$ we, as in case of relevant logic, will obtain an $Ł_{\aleph_0}$-classical version of (**V**)

$(\mathbf{V}^{\rightarrow_L}_{\supset})$ $\Sigma \vdash_{\supset} A$ iff $(\forall w \in RLW)((\forall B \in \Sigma) w \vDash B \rightarrow_L w \vDash A)$

The suggested conclusion is that the extended ‘metalogical’ version of the principle (**V**) would be turned into the following formulation:

- The argument from Σ to A is valid in some logic if in the respective metalogic ‘premises in Σ are true’ implies ‘A is true’.

The question is just what should be or might be the respective metalogic determining consequence relation in an initial logic. The matter of fact is that we should consider two metalogical versions of (**V**):

- The argument from Σ to A is valid in some logic if in some metalogic ‘premises in Σ are true’ implies ‘A is true’;
- The argument from Σ to A is valid in some logic if in all metalogic ‘premises in Σ are true’ implies ‘A is true’.

The second version seems to be non-realistic—it is hard to consider all metalogics, in general. The question also arises whether the choice of a metalogic is determined by the existence of ‘translation’ from some logic into given logic since the existence itself of ‘mixed’ principles of type (**V**) is the result of the intervention or substituting the semantic of one logic instead of another one, an immersion of one semantic conception into another limited them (which causes the restriction of quantification on cases). We need also to think whether this leads to a restriction of some arguments for the benefit of choice between monistic (when logic coincides with metalogic) and pluralistic (when logic and metalogic are conceptually different) position.

And finally, we need to pay attention to one more case concerning the notion of metalogics. The point is that the consideration of one or another non-classical metaimplication can generate a meta-metalogical definition of a consequence. If we consider the metalogical part of (**V**) separately then since the case in point is the metalogical implication then it requires an implementation of this principle. Thus, the necessity appears on the meta-metalogical version of (**V**) of the following kind:

($\mathbf{V}'$) A conclusion A follows from premises Σ if and only if in any case in which each premise in Σ is true is also a case in which A is true iff in any case in which 'each premise in Σ is true' is true is also a case in which 'A is true' is true.

On the one hand, it leads to the bad infinity. On the other hand, this reminds of the situation with Kripke's truth theory (cf. [8]) leading to the following formulation:

($\mathbf{V}^{\text{Meta}}_{\Rightarrow}$) $\Sigma \vdash_{\Rightarrow} A$ iff $T(\Sigma) \Rightarrow T(A)$ iff $T(T(\Sigma)) \Rightarrow T(T(A))$

with the respective modifications of $\vdash_{\Rightarrow}$ for 'mixed' principles. In this case, besides Kripke's consideration of cases of truth and false on the respective metalevels, we need to envisage also truth and false on the pluralistic variants of metalevels.

4 Universal Metalogics

The problem of the status of 'mixed' versions and metaversions of the principle (**V**) is obviously decided as the result of a particular investigation but we can notice, however, that, in a sense, the choice of a metalogic is analogous to the choice of a non-classical universe: by choosing one or another non-classical logic as metalogic in the principle ($\mathbf{V}^{\text{Meta}}_{\Rightarrow}$), we indirectly determine the results of deducibility of one or another formula by metalanguage considerations which are not relevant for the given logical system. One would attempt to assess possible relationship between metalogics and logics in view of exploiting the notion of global and local logics ('global logic' here means some metalogic while 'local logic' means a logical system formulated in object language whose consequence conditions are unambiguously determined by the consequence conditions in the given metalogic, i.e. in global logic). But the bad infinity in ($\mathbf{V}^{\text{Meta}}_{\Rightarrow}$) makes this perspective problematic, though we can accept some conventions: in case of a 'mixed' version of ($\mathbf{V}^{\text{Meta}}_{\Rightarrow}$), we would stop on the 'last' metalevel after which '$\Rightarrow$' does not change anymore.

From the point of view of logical relativism, if one tries to compare some formulations of the type 'B follows from A in logic X iff 'B is true' follows from 'A is true' in logic Y_1' and 'B follows from A in logic X iff 'B is true' follows from 'A is true' in logic Y_2' then some theory is required which gives us the criteria of possible identification (coincidence) of Y_1 and Y_2. As such a theory we can suggest Universal Logics (cf. [20]) within the framework of which the mutual translatability and combinations of all logical systems are discussed. Then identification or connection of given formulations would be accomplished on a base of translation from logic (metalogic) Y_1 into logic (metalogic) Y_2. The result could be: 'B follows from A in logic X iff 'B is true' follows from 'A is true' in logic Y_1' is true iff 'B follows from A in logic X iff the translation of the conditions of 'B is true' follows from the translation of the conditions of 'A is true' in logic Y_2'. If such identification is appropriate for the series of translations then in this case we would say of the phenomenon of a local metalogical monism.

Instead of identification we can employ the join of two formulations with the help of methods of universal logics. In this case, in the join $Y_1 \oplus Y_2$ of the two logics Y_1 and Y_2, common consequence relation $\vdash_{1\oplus 2}$ would be defined using the following condition:

- $T_i(\Sigma) \vdash T_i(A)$ implies $T_{1\oplus 2}(\Sigma) \vdash_{1\oplus 2} T_{1\oplus 2}(A)$ for all $T_i(\Sigma) \cup T_i(A) \cup Y_i$ $(i = 1, 2)$

where $T_i(A)$ means A is true in logic Y_i. This case would be characterized as forcing an additive metalogical monism if we take into account that we can get the joined consequence by means of the join of all consequences for the formulations of the type above, i.e. joining all metalogics of the same logic.

Along with this one can consider the multiplicative metalogical monism resorting to the product of two formulations. For the product $Y_1 \otimes Y_2$ of two logical systems Y_1 and Y_2, the consequence relation $\vdash_{1\otimes 2}$ would be defined with the help of the following condition:

- $\langle T_1(\Sigma_1), T_2(\Sigma_2)\rangle \vdash_{1\otimes 2} \langle T_1(A_1), T_2(A_2)\rangle$ implies $T_i(\Sigma_i) \vdash T_i(A_i)$ for all $T_i(\Sigma_i) \cup T_i(A_i) \cup Y_i$ $(i = 1, 2)$.

This gives us the product of all consequences for the formulations of the type above and thus the product of all metalogics of the same logic.

Finally, we can obtain the coexponential metalogical monism while exploiting coexponential logical consequence relation $\vdash_{1\Leftarrow 2}$:

- $T_1(\Sigma) \vdash_{1\Leftarrow 2} T_1(A)$ iff $g[T_1(\Sigma)] \vdash_2 g(T_1(A))$ for all translations $g : Y_1 \to Y_2$

and the exponential metalogical monism using exponential logical consequence relation $\vdash_{2\Rightarrow 1}$:

- $T_1(\Sigma) \vdash_{2\Rightarrow 1} T_1(A)$ iff there are translations $g[T_1(\Sigma)] \vdash_2 g(T_1(A))$ for all translations $h : Y_2 \to Y_1$ and $g : Y_1 \to Y_2$ such that $h(g[T_1(\Sigma)]) \vdash_1 h(g(T_1(A)))$.

One can try to obtain both coexponential and exponential monism by means of the formulation 'B follows from the translation of A in logic X_2 iff the translation of conditions of 'B is true' follows from the translation of conditions of 'A is true' in logic Y_2' by using the (co)exponential logical consequence relation and the (co)exponential metalogical consequence relation simultaneously but along our way the problem of omniscience arises—alas, we cannot explicitly know all possible logics and hence we cannot accomplish all combinations described above.

5 Semantic Aspects of Logical Relativism

It is worth noting that an issue of confrontation of logical pluralism and logical monism mainly came to the discussions between logicians concerned with the underpinning of one or another viewpoint on the question of entailment, the nature of logical connectives, ways of the elimination of paradoxes, etc. The transition to the more broad horizon specified by the consideration of pure ontological and epistemological consequences of the decisions taken came up only in the context of the philosophy of science. However, the general considerations of those aspects can lead to the consideration of the hybrid 'logic-ontological' pluralism induced by the issue of the nature of ontological commitments of formal languages of different non-classical logics and to the hybrid 'logic-epistemological' pluralism caused by optional acceptance of those commitments (since we can deduce them from previous ones—and this raises a question of their recognition).

As to the specific problems called into existence by the metalogical relativism, it is worth reminding that the metalanguage aspects of the first-order classical logic are usually determined by the interpretation using models (Tarski models). The collection of all sets (the set universe) provides us with all kinds of models required for the interpretation of our logic. Hence, in a sense, the first-order logic is determined by the universe of sets (models).

But there is some subtlety here concerning the set-theoretical identity of this universe. Gila Sher writes: "In standard mathematical logic the metalanguage consists of a fragment of natural language augmented by first-order set theory or higher-order logic. In particular, models are set-theoretic constructs, and the definition of 'satisfaction in a model' is accordingly set-theoretical. This feature of contemporary metalogic is, however, not inherent in the nature of the logical enterprise, and one could contemplate a background language different from the one currently used" [16, p. 53].

Our previous consideration dealt not with a fragment of natural language but with much more narrow and poor formal language of one or another logical theory. And strictly following this course, we should not simply augment our language by first-order set theory but instead employ the respective non-classical set theories based on these logical systems.

In this case, we can also consider metalanguage as a language of the first-order classical set theory ZF while modifying the required definitions of operations on sets (interpreting the non-classical logic connectives) in such a way that they will represent the interpretation of non-classical logical connectives. This gives us an opportunity of the interpretation of the respective non-classical logic in the given metalanguage. The situation emerges where in the classical universe (model of classical logic) an interpretation of non-classical logic is realized. Such a situation is typical for the semantics of the non-classical logic. We can assimilate in our classical universe as many non-classical logics as we need.

But we can also take a non-classical universe, i.e. the model of non-classical logic, and then introduce in it some classical set-theoretical operations. In this case, we obtain an interpretation of the classical logic in a non-classical universe. Moreover, we can proceed over and over again employing other non-classical connectives, obtaining new interpretations of non-classical systems. As a matter of fact, we deal with the situation where within the framework of a non-classical universe we have an interpretation of the classical logic along with other logical systems.

It is significant that we will not have any arguments in favor of the classical universe as compared to a non-classical one. We can at most claim that there is one underlying (global) logic determined and determinate by our universe while there is a set of (local) logics inhabiting universe which is not determinate by them. Globality and locality in such a context are merely metaphoric markers fixing the state of affairs.

The category theory provides us with similar results, especially the topos theory. Here the well-known construction of the topos of functors from a small category into category of sets *Set* serves as a categorical semantic for intuitionistic logic where Heyting algebra (respectively transformed) plays the role of such a small category. Following this course and expanding the scope of consideration, we can obtain a category of functors from the so-called *CN*-category (being a categorical counterpart of da Costa algebra) into *Set*.

This category turned out to be a topos, and completeness of the paraconsistent system of da Costa logic could be proved exactly in respect of such a kind of topos (cf. [19]). An analogous approach could also be taken in the case of a relevance logic **R** where *RN*-categories (a categorical counterpart of the relevant system **R**) are used in order to obtain a topos of functors from the *RN*-category into *Set* (cf. [22]).

Specifically, this can be described in the following manner. First, it is worth taking into account that an intuitionistic nature of toposes displays itself in the construction of internal logic of topos. Such an internal logic for any topos E is typed with each E-object as one type. A formula is a term of type Ω or else an arrow $1 \to \Omega$ with domain type 1 while the logic of such formulas will be a version of intuitionistic logic Int [10, p. 128].

On the other hand, a topos interpretation of intuitionistic logic is yielded by taking a Heyting algebra P as the preorder category (in fact, a bicomplete preorder category with an exponentiation) and considering a topos of functors Set^P. A Set^P- valuation $V : \Phi_0 \to Set^P(1, \Omega)$ assigns to each sentence α a truth-value $V(\alpha) : 1 \dashrightarrow \Omega$ in Set^P. But, according to the definition of formulas of internal logic of topos, this truth-value coincides with the formula of internal logic, and hence one can conclude that $\alpha \models_{Set^P} \beta$ whenever $V(\alpha) \vdash_{TL} V(\beta)$ (where *TL* is a intuitionistic topos logic) and since $\vdash_{\text{Int}} \alpha$ iff $\models_{Set^P} \alpha$ then $\alpha \vdash_{IL} \beta$ if $V(\alpha) \vdash_{TL} V(\beta)$ or, in previous terms, $\alpha \vdash_{\supset_{\text{Int}}} \beta$ iff $\models_{Set^P} \alpha \supset_{\text{Int}} \models_{Set^P} \alpha$.

In the case of da Costa paraconsistent logic, a similar construction could be described following [19] where we replace Heyting algebra P with the da Costa algebra A. To obtain an interpretation of C_1 in a topos Set^A in [19] as the categorical counterpart of da Costa algebra, the so-called *CN*-categories have been implemented. But since $x \leq y \Rightarrow y' \leq x'$ is not a valid property concerning the paraconsistent negation in C_1, we need to improve the definition of *CN*-categories in the following way:

Definition 5.1 A *CN*-category is a preorder category C such that

(i) C has finite products $\langle -, - \rangle$, coproducts $[-, -]$ and C is distributive relative to those, i.e. $\langle [a, b], [a, c] \rangle \cong [a, \langle b, c \rangle]$ for any objects a, b, c in C;
(ii) C allows exponentiation;
(iii) $a \to b$ is an arrow in C iff $a \Rightarrow b \cong 1$, for any two objects a, b in C where $a \Rightarrow b$ is an exponential;
(iv) C has a terminal object 1 and an initial object 0;
(v) For any object a of C there is an object Na such that we have arrows $NNa \to a$ and $a^o \to (Na)^o$ in C where $a^o = N\langle a, Na \rangle$ and for any arrow $d \to a$ there is an arrow $d \to Na$ in C;
(vi) For any two objects a, b in C there is an arrow $a^o \to (b \Rightarrow a) \Rightarrow ((b \Rightarrow \mathbf{N}a) \Rightarrow \mathbf{N}b)$;
(vii) $1 \cong [a, \mathbf{N}a]$ and $0 \cong \langle a^o, \mathbf{N}a^o \rangle$.

It is easy to check that any *CN*-category has the following properties:

(a) An exponential $a \Rightarrow b$ in C will be a residual,
(b) C is Cartesian closed,
(c) $y \to x$ is an arrow in C iff $\langle x, y \rangle \cong y$ and $[x, y] \cong x$,
(d) $\langle \langle \mathbf{N}a, a^o \rangle, a \rangle \cong 0$, $[\langle \mathbf{N}a, a^o \rangle, a] \cong 1$,
(e) Every *CN*-category has at least three objects.

Starting from a categorical representation of the da Costa algebra in the form of *CN*-categories, one may proceed directly to a topos Set^A (where A is a *CN*-category) as a model for paraconsistency logic C_n. As in the case of intuitionistic logic, for any functor $\mathbf{F} : A \to Set$ we denote by $\mathbf{F}_p$ the value $\mathbf{F}(p)$ of functor $\mathbf{F}$ for an object p from A. For any q and p such that $p \leq q$, a functor $\mathbf{F}$ defines the function from $\mathbf{F}_p$ to $\mathbf{F}_q$ which we denote $\mathbf{F}_{pq}$. A functor $\mathbf{F}$ will be treated as the collection $\{\mathbf{F}_p : p \in A\}$ of sets indexed by elements of the set A from an algebra A and endowed with the transition mapping $\mathbf{F}_{pq} : \mathbf{F}_p \to \mathbf{F}_q$ under $p \leq q$ (in particular, $\mathbf{F}_{pp}$ will be the identity function on $\mathbf{F}_p$).

A classifying object in topos Set^A can be defined as $\Omega_p = [p)^+$ (where A^+ is a set of all principal filters, i.e. sets $[p) = \{q : p \leq q\}$ where $p, q \in A$) and for p and q such that $p \leq q$ the function $\Omega_{pq} : \Omega_p \to \Omega_q$ maps every $S \cap [p)^+$ into $S \cap [q) \in [q)^+$, i.e. $\Omega_{pq}(S) = S_q$.

Exploiting a Set^A-valuation $V : \Phi_0 \to Set^A(1, \Omega)$ which assigns to every formula α some truth value $V(\alpha) : 1 \rightarrow \Omega$, an interpretation of C_n in topos Set^A is obtained.

And now as in the case of intuitionistic logic one can conclude that $\alpha \models_{Set^A} \beta$ whenever $V(\alpha) \vdash_{TL} V(\beta)$ (where TL is an intuitionistic topos logic) and since $\vdash_{C_n} \alpha$ iff $\models_{Set^A} \alpha$ (cf. [19]) then $\alpha \vdash_{C_n} \beta$ if $V(\alpha) \vdash_{TL} V(\beta)$ or, in previous terms, $\alpha \vdash_{\supset C_n} \beta$ iff $\models_{Set^A} \alpha \supset_{\mathrm{Int}} \models_{Set^A} \alpha$. But we need to take into account that if we use one and the same set as the carrier both of Heyting algebra P and da Costa algebra A then in the second case we exploit just those elements of P for which $\Omega_p = [p)^+$ will be a da Costa algebra. Denoting the set of these elements as CP, we finally obtain $\alpha \vdash_{\supset C_n} \beta$ iff $\forall p \in CP (p \models_{Set^A} \alpha \supset_{\mathrm{Int}} p \models_{Set^A} \alpha)$.

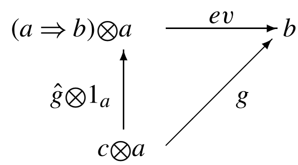
A relevant logic provides us with one more case of an interplay of logic and metalogic. Here we again need to consider a transformation of an algebra into a category and a suitable starting construction turns out to be the so-called *RN*-categories (cf. [22]).

Definition 5.2 An R-category $\mathcal{C}$ is a preorder category endowed with a covariant bifunctor $\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ such that

(i) $\mathcal{C}$ has finite products $\langle -, - \rangle$, coproducts $[-, -]$ and $\mathcal{C}$ is distributive relatively to those, i.e. $\langle [a, b], [a, c] \rangle \cong [a, \langle b, c \rangle]$ for any objects a, b, c in $\mathcal{C}$;
(ii) For any objects a, b, c in $\mathcal{C}$ there are the following natural isomorphisms:

$$a \otimes [b, c] \cong [a \otimes b, a \otimes c], \quad [b, c] \otimes a \cong [b \otimes a, c \otimes a],$$

i.e. bifunctor preserves coproducts;
(iii) $\mathcal{C}$ allows exponentiation relative to $\otimes$, i.e. the following diagram commutes



where $\Rightarrow$ is an exponential;
(iv) The following functorial equations are satisfied:

(a) $(g_1 f_1) \otimes (g_2 f_2) = (g_1 \otimes g_2)(f_1 \otimes f_2)$;
(b) $1_A \otimes 1_B = 1_{A \otimes B}$.

It is easily can be seen that our exponential $\Rightarrow$ plays the role of a residual relative to $\otimes$. To wit, there is the arrow $c \to a \Rightarrow b$ whenever we have an arrow $c \otimes a \to b$ and the other way round.

An R-category $\mathcal{C}$ is *monoidal* if:

(v) $\mathcal{C}$ has an object 1 such that $1 \otimes a \cong a$ and there is an arrow $a \to a \cong 1$ in $\mathcal{C}$ for all a in $\mathcal{C}$;
(vi) For any objects a, b, c in $\mathcal{C}$ $a \otimes (b \otimes c) \cong (a \otimes b) \otimes c$.

A monoidal R-category is *symmetric monoidal* if:

(vii) For any objects a, b in $\mathcal{C}$ there is an arrow $a \otimes b \to b \otimes a$.

A symmetric monoidal R-category is *relevant* if:

(viii) For any object a in $\mathcal{C}$ there is an arrow $a \to a \otimes a$.

An *RN*-category $\mathcal{C}$ is an R-category together with a contravariant functor $\mathcal{N} : \mathcal{C} \to \mathcal{C}$ such that

(ix) $\mathcal{N}^2 a \cong a$ for any a in $\mathcal{C}$;
(x) For any arrow $a \otimes b \to c$ there is an arrow $a \otimes \mathcal{N}c \to \mathcal{N}b$ in $\mathcal{C}$.

After accepting *RN*-categories as the initial point for a Set^C-construction in case of relevance logics, we need the following definitions and facts.

It is known that for a bounded distributive lattice L the dual space of L, $S(L)$, is the ordered topological space in which the set of points S is the family of all prime filters of L, ordered by containment. A relevant space is a structure $\mathcal{R} = \langle S, R, *, t\rangle$ where S is our family of all prime filters, R is a ternary relation on S, $*$ is a unary function on S, $t \in S$. For $A, B \subseteq S$ let $A \circ B$ be $\{z : \exists xy(Rxyz \;\&\; x \in A \;\&\; y \in B)\}$, and let $A \to B$ be $\{x : \forall yz((Rxyz \;\&\; y \in A)$ implies $z \in B)\}$. The following conditions are satisfied:

1. If $A, B \in L(S)$ then $A \circ B$ and $A \otimes B$ are clopen;
2. $(Rxyz \;\&\; x' \leq x \;\&\; y' \leq y \;\&\; z' \leq z)$ implies $Rx'y'z'$;
3. $\forall xyz(\neg Rxyz$ implies $\exists A, B \in L(S)(x \in A \;\&\; y \in B \;\&\; z \notin A \circ B))$;
4. The function $x \mapsto x*$ is a continuous decreasing map on S;
5. t is in $L(S)$ and satisfies the condition:

$$\forall yz\big(y \leq z \text{ is equivalent to } \exists x(x \in t \;\&\; Rxyz)\big),$$

where $L(S)$ is the dual lattice of S.

The dual algebra of R, $A(R)$ is defined on the lattice $L(S)$ by adding the operations $A \circ B, A \to B, \neg A = \{z : z* \notin A\}$ and defining the constant 1 as t. For a relevant algebra L the dual space of L, $\mathcal{R}(L)$, is defined by adding to the lattice of all prime filters of a bounded distributive lattice the ternary relation $x \cdot y \subseteq z$, defining $x*$ for $x \in \mathcal{R}(L)$ to be $\{a : \neg a \notin x\}$, and letting $t = \{x \in S(L) : 1 \subseteq x\}$. The following statements are true:

1. The dual algebra of a relevant space is a relevant algebra;
2. The dual space of a relevant algebra is a relevant space.

And moreover:

1. If A is a relevant algebra, then $\mathcal{A}$ is isomorphic to its second dual, $\mathcal{A}(\mathcal{R}(L))$, under the mapping $\eta(a) = \{x \in \mathcal{R}(L) : a \in x\}$;
2. If $\mathcal{R}$ is a relevant space, then $\mathcal{R}$ is r-homeomorphic (i.e. order-homeomorphic and isomorphic with the respect to the ternary relations, $*$-operation and unit) to its second dual $\mathcal{R}(\mathcal{A}(\mathcal{R}))$, under the mapping $\theta(x) = \{B \in \mathcal{A}(\mathcal{R}) : x \in \mathcal{R}\}$.

For our purposes it is convenient to use the following notions: If x, y are elements of a relevant space $\mathcal{R}$, then define $x \circledcirc y$ to be the set $\{z \in \mathcal{R} : Rxyz\}$, $x \twoheadrightarrow y$ to be the set $\{z \in \mathcal{R} : Rzxy\}$ and $\overline{\neg} x = \{z : z \notin x*\}$; for X, Y subsets of the space $\mathcal{R}$, define $X \circledcirc Y$ to be the set $\cup\{x \circledcirc y : x \in X, y \in Y\}$, $X \twoheadrightarrow Y$ to be the set $\cap\{x \twoheadrightarrow y : x \in X, y \in Y\}$ and $\overline{\neg} X = \{\overline{\neg} x : x \in X\}$. Elements of the relevant space $\mathcal{R}$ would be, in fact, identified with the principal filters, i.e. the element $x \in \mathcal{R}$ would be identified with $[x) = \{y : x \leq y\}$. In this case, the set $x \circledcirc y$ is identical with the set $[x) \circledcirc [y)$, $x \twoheadrightarrow y$ is identical with the set $[x) \twoheadrightarrow [y)$ and $\overline{\neg} x = \overline{\neg}[x)$.

Now we proceed directly to a topos $Set^{\mathcal{C}}$. First, consider the functor $\Omega : \mathcal{C} \to Set$ (where $\mathcal{C}$ is a RN-category) which will represent a classifying object in topos $Set^{\mathcal{C}}$. For any functor $F : \mathcal{C} \to Set$ we denote by F_p the value $F(p)$ of functor F for an object p from $\mathcal{C}$. For any q and p such that $p \leq q$, a functor F defines the function from F_p to F_q which we denote F_{pq}. A functor F will be treated as the collection $\{F_p : p \in \mathcal{C}\}$ of sets indexed by elements of the set of objects of $\mathcal{C}$ and endowed with the transition mapping $F_{pq} : F_p \to F_q$ under $p \leq q$ (in particular, F_{pp} will be the identity function on $\mathcal{F}_p$).

We continue in this fashion putting $\Omega_p = [p)^+$ (= a relevant algebra of all principal filters in $[p)$) and for p and q such that $p \leq q$ the function $\Omega_{pq} : \Omega_p \to \Omega_q$ maps every $S \in [p)^+$ into $S \cap [q) \in [q)^+$, i.e. $\Omega_{pq}(S) = S_q$.

Again, by introducing a Set^C-valuation $V : \Phi_0 \to Set^C(1, \Omega)$ which assigns to an every formula α some truth value $V(\alpha) : 1 \dot{\to} \Omega$, an interpretation of C_n in topos Set^C might be obtained. And then as in the case of intuitionistic and paraconsistent logic one can conclude that $\alpha \models_{Set^C} \beta$ whenever $V(\alpha) \vdash_{TL} V(\beta)$ (where TL is an intuitionistic topos logic) and since $\vdash_R \alpha$ iff $\models_{Set^C} \alpha$ (cf. [22]) then $\alpha \vdash_R \beta$ if $V(\alpha) \vdash_{TL} V(\beta)$ or, in previous terms, $\alpha \vdash_{\to_R} \beta$ iff $\models_{Set^C} \alpha \supset_{\mathrm{Int}} \models_{Set^C} \alpha$. But again we need to take into account that if we use one and the same set as the carrier both of Heyting algebra P and relevant algebra C then in the second case we exploit just those elements of P for which $\Omega_p = [p)^+$ will be a relevant algebra. Denoting the set of these elements as RP, we finally obtain $\alpha \vdash_{\to_R} \beta$ iff $\forall p \in RP (p \models_{Set^C} \alpha \supset_{\mathrm{Int}} p \models_{Set^C} \alpha)$.

All these results can be evaluated as parallel to a set-theoretic construction of non-classical universes where models of non-classical logic are yielded 'inside' the model of the classical (herein intuitionistic) logic. To this end, special categories were introduced like toposes rendering the structure of various non-classical logics (in particular, paraconsistent da Costa logic and relevant logic) and a project of the general method of obtaining such 'algebra-valued' categories is on agenda.

But if in the first case the matter concerns a semantic embedding of non-classical logics into intuitionistic universe—topos, then in the second case we deal with the construction of other non-classical category-theoretical universes playing the role of 'non-classical' topos. In particular, da Costa topos (or potos) appears to be a paraconsistent universe in which paraconsistent mathematics would be developed the way it was done in toposes for intuitionistic mathematics (cf. [23]). But if in the case of topos of functors from da

Costa algebra into category of sets, all instances of paraconsistency are the specific constructions within intuitionistic universe, some local artefacts, then in da Costa topos this paraconsistency would be absolutely organic by its nature, underlying all constructions and would be global and fundamental. Here already classical mathematics emerges as an artefact within paraconsistent universe, as some local deviation from paraconsistent patterns of relations. Thus, e.g. to interpret da Costa C_1–C_n systems it is required just to employ non-truth-functional valuation while truth-functional valuation becomes specific only for the Boolean toposes which are now just the particular case of da Costa toposes.

One more example would be the quantum topos or quantos [21]. It is a category which is structurally oriented to quantum logic but being similar to topos in that it has a subobject classifier and products (its subobject algebra and subobject classifier have an orthomodular lattice structure). Every quantos encloses Boolean topos and in the meantime it would be a universe for the interpretation of quantum logic.

References

1. Beall, J.C., Restall, G.: Logical pluralism. Australas. J. Philos **78**, 475–493 (2000)
2. Beall, J.C., Restall, G.: Defending logical pluralism. Logical consequence: rival approaches. In: Woods, J., Brown, B. (eds.) Proceedings of the 1999 Conference of the Society of Exact Philosophy, pp. 1–22. Hermes, Stanmore (2001)
3. Beziau, J.-Y.: The relativity and universality of logic. Synthese (2014). doi:10.1007/s11229-014-0419-0. Published Online: 20 March 2014.
4. Carnap, R.: The Logical Syntax of Language. Little eld, Adams and Co. (1959). Translated by Amethe Smeaton
5. Cook, R.: Let a thousand flowers bloom: a tour of logical pluralism. Philos. Compass **5/6**, 492–504 (2010)
6. Goldblatt, R.I.: Semantic analysis of orthologic. J. Philos. Log. **3**(1–2), 19–35 (1974)
7. Jeffrey, R.C.: Formal Logic: Its Scope and Its Limits, 3rd edn. McGraw Hill, New York (1991)
8. Kripke, S.: Outline of truth theory. J. Philos. **72**(9), 695–717 (1975)
9. Kron, A., Marič, Z., Vujosevič, S.: Entailment and quantum logic. In: Beltrametti, S., Van Fraassen, B. (eds.) Current Issues in Quantum Logic, pp. 193–207. Plenum, New York (1981)
10. McLarty, C.: Elementary Categories, Elementary Toposes. Clarendon Press, Oxford (1992)
11. Priest, G.: Doubt Truth to Be a Liar. Clarendon Press, Oxford (2008)
12. Read, S.: Relevant Logic. Blackwell, Oxford (1988)
13. Read, S.: Monism: the one true logic. In: DeVidi, D., Kenyon, T. (eds.) A Logical Approach to Philosophy: Essays in Honour of Graham Solomon, pp. 193–209. Springer, Berlin (2006)
14. Restall, G.: Carnap's tolerance, meaning and logical pluralism. J. Philos. **99**, 426–443 (2002)
15. Routley, R., Meyer, R.: Semantics of entailment. In: Leblanc, H. (ed.) Truth Syntax and Modality, pp. 199–243. Amsterdam–London (1973)
16. Sher, G.Y.: The Bounds of Logic. A Generalized Viewpoint. MIT Press, Cambridge (1991)
17. Sylvan, R.: Philosophy, politics and pluralism: 1. Relevant modellings and arguments. J. Non-Class. Log. **4**(1), 57–107 (1987)
18. Vasyukov, V.L.: The completeness of the factor semantics for Łukasiewicz's infinite-valued logics. Stud. Log. **52**(1), 143–167 (1993)
19. Vasyukov, V.L.: Paraconsistency in categories. In: Batens, D., Mortensen, C., Priest, G., van Bendegem, J.-P. (eds.) Frontiers of Paraconsistent Logic, pp. 263–278. Research Studies Press, Baldock (2000)
20. Vasyukov, V.L.: Structuring the universe of universal logic. Log. Univers. **1/2**, 277–294 (2007)
21. Vasyukov, V.L.: An ontology of quantum mathematics. Bull. Peoples Friendsh. Univ. Russ. Philos. **3**, 57–70 (2009) (in Russian)

22. Vasyukov, V.L.: Paraconsistency in categories: case of relevant logic. Stud. Log. **98**(3), 429–443 (2011)
23. Vasyukov, V.L.: Potoses: paraconsistent categories for paraconsistent logics. Manuscript (2014)

V.L. Vasyukov
National Research University Higher School of Economics, Moscow, Russia

V.L. Vasyukov (✉)
Institute of Philosophy, Russian Academy of Sciences, Volkhonka 14, 119991 Moscow, Russia
e-mail: vasyukov4@gmail.com

Constructivism and Metamathematics

Jan Woleński

Abstract This paper discusses the problem of availability of constructive metamathematics for constructive theories. Since intuitionism is usually considered as the most important kind of constructivism, we have the following question: Is it possible to give intuitionistically acceptable proofs of metamathematical theorems? This issue is illustrated by the completeness theorem for intuitionistic predicate logic and the consistency of arithmetic. The conclusion is that hitherto collected evidence justifies the claim that the universal intuitionistic metamathematics is very problematic.

Keywords Intuitionism · Logic · Completeness · Provability

Mathematics Subject Classification 03A5

Constructivism in the foundations of mathematics comprises a broad variety of views on mathematical theories and their properties. Constructivism is not a uniform proposal in the foundations of mathematics (see [1, 2, 8, 13]). It comprises, even disregarding some older views as, for instance, Kronecker's account of mathematical reasoning or French semi-intuitionists, many special versions, in particular, intuitionism (as represented by Brouwer and his followers), Russian constructivism, Bishop's constructivism, computable analysis, constructive analysis, predicativism, Lorenzen's operationalism, ultrafinitism (ultra-intuitionism), and constructive type-theory. Each brand of constructivism offers a criterion **C** of constructivity. In my further remarks, I will concentrate on intuitionism and Russian constructivism. Let us assume that **C** refers to a given fixed criterion of constructivity. Such a criterion always generates a kind of logic considered as constructive. For example, **C** can postulate logic in which the law of excluded middle does not hold, that is, the standard intuitionistic logic, or recommends the Markov principle (the symbol $\neg$ expresses negation, but its meaning depends on the accepted system; in the case of the Markov principle, $\neg$ formalizes classical negation)

(MP) $\forall x(A(x) \vee \neg A(x)) \wedge \neg\forall x \neg A(x) \Rightarrow \exists x A(x)$,

as admissible (intuitionism rejects (MP) as not sufficiently constructive). Consequently, a theory **T** is constructive if and only if it satisfies C. Typically, constructive systems as

This paper is essentially based on my talk delivered at the 14th International Congress of Logic, Methodology and Philosophy of Science held in Nancy in 2011. An extended abstract of this talk is available at http://www.clmps2011/en/editorial.html.

A. Koslow, A. Buchsbaum (eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-10193-4_25

based on a kind of constructive logic as its formal base are contrasted with theories having classical logic as their logical foundation.

How to compare, evaluate or justify foundational approaches to mathematics? One way consists in taking some philosophical–methodological insights as guiding. For instance, although a Platonist says that a mathematical object exists if and only if it is consistent, an intuitionist claims that consistency forms the necessary condition for existence which must be supplemented by the possibility of construction. This justification is both philosophical and methodological because it concerns existence, the main ontological concept, but, at least in the case of constructivism, invokes a method of proving existence; this means that a methodological (or perhaps even epistemological) parameter deeply enters into the stage. The opposition between Platonism and intuitionism appears as crucial. In fact, both philosophical (or foundational) positions exclude each other. The difference between intuitionism and formalism is not so transparent because the latter (at least as represented by Hilbert) is, on the one hand, a quite radical constructivism as admitting finitary metamathematical reasoning only, but, on the other hand, the principle of excluded middle holds in logic accepted by the formalists. In other words, whereas formalism favors classical logic as generally correct (it shares this attitude with Platonism), it restricts logical tools used in metamathematics to effective (just finitary), intuitionism assumes its version of constructive logic as suitable in all cases of mathematical reasoning.

Any exposition of philosophical foundations of mathematics cannot abstract from some mathematical ideas. This fact finds its illustration, for instance, in the use of the phrase "finitary mathematical reasoning" as employed in the above characterization of formalism and its relation to intuitionism or in the character of logic accepted by this or that particular view in the foundations of mathematics (logic is a mathematical theory). Yet, formal mathematical ingredients in justification of the foundational views can be more or less explicit, although we should always remember that the border between mathematics and philosophy (in the discussed domain) is somehow vague. In what follows, I would like to point out some problems related to the role of metamathematical properties in justification of constructivism. I restrict my considerations to intuitionism and Russian constructivism. Independently of the fact that ordinary or hard mathematicians are usually more or less skeptical about metamathematical investigations, all genuine concepts used in metamathematics have an explicit mathematical content and function in theories which satisfy typical mathematical requirements. In particular, these concepts are definable (or not) by resources and devices used in mathematics, occur in theorems stated under fully specified constraints and are provable by definite methods. The domain of such studies is frequently called mathematical foundations of mathematics. Thus, facts derived from metamathematics as a legitimate part of mathematics can (and should) be employed in (possibly) the justification of constructivism (and any other declared position in the philosophy of mathematics as well).

If we characterize mathematical theories, we say that they are (or not) consistent, complete, decidable, independent, etc. Not all metamathematical (or metalogical) properties are equally important. Without entering into a discussion about the meaning of importance in this context, we can consider consistency as a mandatory requirement for any mathematical theory and completeness as good everywhere where it is available but required for pure logic (I will incidentally mention other properties as well).[1] Let **T** be a

[1]Since Jean-Yves is one of the champions of paraconsistency, I am sorry that I ignore paraconsistent logic.

theory, that is, a set of formulas closed by the consequence operation *Cn* (of course, this operation must be relativized to assumed logic, for instance, classical or constructive). Thus, $Cn\,\mathbf{T} \subseteq \mathbf{T}$ (since the reverse inclusion is trivial, we have $Cn\,\mathbf{T} = \mathbf{T}$). The first task consists in investigating which metamathematical properties hold (or not) for **T**, e.g. if it is consistent, complete, decidable, etc. However, except possessing or not some properties, we have the second issue, namely the way of proving whether **T** is consistent or complete. More specifically, the issue concerns the possibility of constructive proofs of such as '**T** is consistent' or '**T** is complete'. In particular, if **T** is constructive, we have a more concrete problem, namely:

(∗) Does **T** possess a constructive metatheory (let us denote it by **MT**)?

For instance, one can ask "If **T** is intuitionistic, is **MT** intuitionistic as well?"

The problem posed by (∗) is interesting. Clearly, each fact provable constructively (on any understanding of constructivism) is also provable classically. A simple illustration is provided by the relation between direct and indirect proofs. A constructivist does not accept all forms of indirect proofs (details are not relevant here), but the classical position considers such reasoning as legitimate (if formally correct). Consequently, every direct proof can be converted into an indirect proof, although such a manoeuvre appears rather strange because direct proofs have a better reputation than indirect ones. The converse problem is much more important. Many theorems have indirect proofs and one can ask whether their transformation into direct reasoning is possible at all. A classical logician considers the issue as open, but he adds that if a direct proof is not available, we should accept an indirect one. The constructivist rejects *ad absurdum* indirect proofs and adds that theorems not proved constructively are not acceptable. The most radical constructivist regards such results as meaningless, but a more moderate position recommends waiting for constructive proofs.[2] This latter methodological proposal takes constructivism as a program.

I will argue that we have arguments for regarding the programmatic constructivism as limited. As I have noticed earlier, intuitionism and Russian constructivism function as the points of reference. Intuitionism has a distinguished position in the entire constructivist camp. It began the story of the mature constructivism, and its logical foundations are particularly well-established. Many logicians consider intuitionist logic as the most important rival of classical logic. Russian (Markov's) constructivism represents a moderate position. (MP) exhibits a clear formal difference holding between the both mentioned kinds of the constructive philosophy and metatheory of mathematics. Another interesting distinction between intuitionism and Russian constructivism concerns the Church thesis (CT). Whereas the latter school accepts (CT) in its fully general form (roughly speaking, every computable function is recursive), the former restricts (CT) to the principle that only arithmetical functions are computable (see [13], 28–30, 192–193 for further details). Since CT has its main relevance in developing constructive analysis than in metamathematics, I will not appeal to it in further consideration. I repeat that my main focus consists in examining consistency and completeness in the framework of constructivism (more

[2] This is the picture from the side of ideology. In fact, constructivists sometimes offer non-constructive proofs. Brouwer's fixed-point theorem in topology is a very good example in this respect. In general, philosophical ideology is not always coherent with the needs of mathematics.

specifically, within intuitionism). I will do that via deriving some consequences from basic metalogical results, in particular by using the completeness theorem and the first and second incompleteness theorem.

The completeness property can be defined either syntactically or semantically (see [12] for a more comprehensive treatment of metamathematical concepts and results used in this paper). Syntactically speaking, the completeness property is defined by the statement:

(1) A theory **T** is syntactically complete if and only if for any formula A of the language of **T**, $A \in \mathbf{T}$ or $\neg A \in \mathbf{T}$.

If **T** has the property defined by (1), we say that the syntactic completeness theorem holds for it. The syntactic character of this definition (and the related theorem) becomes evident, if we observe that only syntactic notions occur in its formulation because only notions of syntax are involved in it. Another concept of completeness equates the set of theorems (formulas provable by admitted rules of inference) in **T** and the collection of all sentences of **T** true in each of its models. It is a semantic notion because it involves the concept of truth. More formally,

(2) A theory **T** is semantically complete if and only if ($A \in \mathbf{T}$ if and only if A is true in all models of **T**).

In fact, the completeness property is defined by the if-part of the right part of (2); the only-part of the definition expresses the property of soundness. If **T** is semantically complete, then it enjoys the completeness theorem, that is, the statement

(3) If A is a true in **T**, then A is provable in **T**; in symbols, $\models A \Rightarrow \vdash A$.

This proposition is equivalent to

(4) A theory **T** is consistent if and only if it has a model,

that is, the Gödel–Malcev completeness theorem.

(1)–(4) can be applied to an arbitrary formal system. However, since higher-order logic plays a minimal role in intuitionism, it is reasonable to limit further analysis to first-order (that is, elementary) systems. Moreover, the situation of pure logic (denoted by **LOG**) understood as $Cn\,\varnothing$ and consisting of propositional calculus (**PC**) plus quantification calculus (**QC**) is peculiar to some extent. Neither **PC** nor **QC** are syntactically complete; this holds for classical logic as well for intuitionistic logic. On the other hand, we have yet another concept of completeness, namely the Post-completeness, which is more suitable for logic. Let **LOG** be a logic. We say that **LOG** is Post-complete if and only if for any formula A of the language of **LOG**, A is a theorem of **LOG** or the set $\mathbf{LOG} \cup \{A\}$ is inconsistent. Now, **CPC** (classical propositional calculus) and **IPC** (intuitionistic propositional calculus) are Post-complete, but **CQC** (classical quantification calculus) and **IQC** (intuitionistic quantification calculus) are not. (1)–(4) can be trivially generalized to **LOG**, although model theory for intuitionistic theories, including logic, is not as uniform as in the classical case; yet I take for granted that intuitionistic semantics is satisfactory. Thus, **LOG** is syntactically incomplete, partially Post-complete, semantically complete, and the equivalence of (3) and (4) holds for it; these conclusions concern $\mathbf{LOG}^{\mathbf{C}}$ (classical logic) and $\mathbf{LOG}^{\mathbf{I}}$ (intutionistic logic). (3) and (4) and the equivalence of both are valid for extralogical theories. There are extralogical syntactically complete first-order theories, for instance, Presburger arithmetic, but their expressive power is rather small. The arithmetic of natural numbers, perhaps the most important extralogical mathematical theory, is syntactically incomplete (the first incompleteness theorem). The above observations are true with respect to intuitionistic as well as classical systems.[3]

Where is the problem?—one can ask. The metamathematical parity of intutionistic and classical systems essentially breaks at the point of the way of proving some results. Clearly, all results concerning classical systems have classical proofs, that is, they are executed by deductive devices available in $\mathbf{LOG^C}$. Some of them also have intutionistic proofs, but this fact is not relevant at the present moment.[4] The real problem occurs when we ask whether all metamathematical theorems valid for the systems based on intuitionistic logic have proofs acceptable from the intuitionistic point of view. No problem is provided by propositional calculus because we can prove Post-completeness, the completeness theorem and their equivalence by means of $\mathbf{LOG^I}$. The situation becomes radically different in the case of predicate calculus. Since the completeness theorem implies (MP), the intuitionistic proof of the former seems problematic (see [5, pp. 154–155], [7] for a discussion) as intuitionistically acceptable; recall that an intuitionist does not accept (MP).[5] Of course, the same holds for the completeness property of extralogical theories. Anyway, if a constructive proof of (3) and (4) is to be available even in the case of $\mathbf{LOG^I}$, constructivism as modeled by intuitionism must be weakened. Yet, the equivalence of (3) and (4) can be intuitionistically established. Since both Gödel incompleteness theorems have intuitionistic proofs, the completeness theorem, in its both versions, is actually a critical problem for any constructivism. The role of (MP) shows that Russian constructivism constitutes a very interesting proposal. In particular, the case of the completeness theorem for **IQC** shows that metamathematical arguments are fairly important for understanding what the content of constructivism as a foundational project is. However, the issue of metamathematics in the context of constructivism is not limited to comparing particular versions of constructivism. Incidentally, similar comparisons are also possible between constructivism and other foundational views.

I pass to the problem of consistency by relating it to Peano arithmetic (**PA**; it uses classical logic as its formal skeleton) and Heyting arithmetic (**HA**; it has the same arithmetical axioms as **PA**, but intuitionistic logic forms its logical base). Since **PA** is expressible in **HA**, the latter is subjected to incompleteness theorems. In particular, the consistency of **HA** cannot be proved either in **HA** or in **PA**. On the other hand, the consistency of **HA** entails the consistency of **PA** (this fact was established by Gentzen in [6], see also [4, p. 30], [5, p. 24] for justification of this fact). This assertion fact appears as interesting in itself because it suggests that consistency of arithmetic is independent of choosing classical or intuitionistic logic as the indented logical base. One could interpret this situation as providing a good point for intuitionism because it suggests that weakening of classical

[3]The difference concerns decidability. **CPC** and **IPC** are decidable. Classical monadic quantification (predicate) logic is decidable, but intuitionistic is not. The full **CQC** and **IQC** are undecidable.

[4]I will return to the problem of constructive proofs within the classical position at the end of this paper.

[5]Some authors (see [3, 14])) claim that they offer purely intuitionistic completeness proofs. However, a deeper discussion (see [13, pp. 729–730], [4, pp. 186–201]) demonstrated that the related arguments were based on non-standard semantics and/or concerned the so-called $\bot$-free fragment of intuitionistic predicated calculus. As Michael Dummett says [5, p. 154], the proof via (MP) is "a best possible result" from the constructive point of view. In fact, some proofs of the completeness of **IPC** use classical logic at the level of **MT** (see [5, p. 87]).

logic by the rejection of the principle of the excluded middle (or other critical logical rule contested by the intuitionist) does not increase the danger of inconsistency of arithmetic.

However, the following reasoning shows that the issue has further aspects which make the success of intuitionism very problematic as far the issue concerns consistency of **HA** and its significance for the question of existence in intuitionism. The implication (Cons(**PA**)—'**PA** is consistent'; Cons(**HA**)—'**HA** is consistent', $\vdash^{\mathbf{I}}$—intuitionistic provability)

(5) $\mathrm{Cons}(\mathbf{HA}) \Rightarrow \mathrm{Cons}(\mathbf{PA})$

can be established intuitionistically (all further steps are intuitionistically sound as well). Thus, (5) can be rewritten as

(6) $\vdash^{\mathbf{I}} (\mathrm{Cons}(\mathbf{HA}) \Rightarrow \mathrm{Cons}(\mathbf{PA}))$.

Consequently, we a have (via the Bernays provability conditions)

(7) $\vdash^{\mathbf{I}} \mathrm{Cons}(\mathbf{HA}) \Rightarrow \vdash^{\mathbf{I}} \mathrm{Cons}(\mathbf{PA})$.

Transposing (7) gives

(8) $\neg\vdash^{\mathbf{I}} \mathrm{Cons}(\mathbf{PA}) \Rightarrow \neg\vdash^{\mathbf{I}} \mathrm{Cons}(\mathbf{HA})$.

Since the proof of the consistency of **PA** cannot be carried out in **PA** itself, the consistency of **HA** is unprovable by the devices available in **HA**. The fact (8) makes troubles for intuitionism because it shows that the intuitionist is not able to establish intuitionistically the most fundamental property of mathematical theories in which arithmetic is expressible.[6] In general, since consistency forms a necessary condition of existence in intuitionism, the intuitionist cannot prove this condition by methods approved by him, at least in arithmetic and stronger theories. Thus, faith and empirical arguments remain, similarly as in the case of the classical position. Although they might be considered as more or less convincing, it is difficult to consider them as constructively coherent with intuitionism. *Pace* intuitionism, (5), suggests a relative proof of Cons(**PA**) modulo Cons(**PA**). It can be viewed as an advantage of intuitionism because if all arguments for consistency of a weaker theory become arguments for consistency of a stronger theory, our faith into the latter could be considered as better grounded. Yet, it is still an act of faith based on empirical evidence, and not a rigorous mathematical reasoning.

Three general conclusions can be derived from the above analysis. Firstly, there are various degrees of constructivism in metamathematics. Due to (MP), Russian constructivism is less radical, but more flexible than intuitionism; similar comparisons are possible with respect to the whole constructivist camp. Secondly, given a fixed **C**, a universal constructive **MT** relatively to **C** seems improbable because known informal treatments, like forwarded in this paper, use classical logic and this situation seems to be fairly stable. Thirdly, constructive arguments are to be preferred even in the domain of classical metamathematics, similarly as in the case of ordinary mathematics, quite independently of a

[6] I communicated this conclusion and arguments supporting it to Ann Troelstra. He replied that I was not right because since the intuitionists do not ascribe a major importance to formalization, they are not particularly sensitive to the argumentative force of metamathematical questions. However, I consider this answer as begging the question. In fact, metamathematics plays a considerable role in [13], co-authored by Troelstra.

particular view in the foundations of mathematics. It can be illustrated by the fact that the equivalence of (3) and (4) can be constructively provable, although both completeness theorems have classical or less constructive proofs, by the availability of the proof of (3) via (MP), and by (5). Some radical constructivists, for example, Brouwer, argue that non-constructive mathematics is meaningless. Consequently, they cannot say, for instance, that intuitionistic logic is a part of classical one. Of course, a classicist has no problem with that assertion because it is known that $\mathbf{LOG^C}$ is the only maximal extension of $\mathbf{LOG^I}$. Thus, it is possible to develop a detailed account of intuitionism in the classical framework (see [4, 11]). An intuitionist can eventually employ the technique of negative translation and replace every classical assertion A by the formula $\neg\neg A$, where the symbol $\neg$ refers to intutionistic negation. However, this move does not help the intuitionist very much because it leads to negative statements of the type "it is absurd that A is absurd" (or more accurately "if (if A implies absurdity), then absurdity" and symbolically "$(A \to \bot) \to \bot$; the symbol $\bot$ denotes absurdity) and we further have a problem of the exact meaning of A under intuitionistic meaning-rules. Since, according to intuitionism, the content of $\neg\neg A$ should be generally considered as different from the content of A, provided that the assertion A is established classically, and not constructively, there are no universal tools for comparing the content of arbitrary statements of ordinary mathematics from the intuitionistic point of view. On the other hand, a classicist can distinguish various degrees of constructivism (this way out was suggested by Andrzej Mostowski in Poland in the 1950s; see [9], [10, p. 106]) and look for constructive proofs of various theorems without violating his own theoretical position. This means that the program of constructive foundations of mathematics does not require the radical constructivism and its reach should be regarded as an empirical question which cannot be decided *a priori*.

References

1. Beeson, M.J.: Foundations of Constructive Mathematics. Springer, Berlin (2004)
2. Bridges, D., Richman, F.: Varieties of Constructive Mathematics. Cambridge University Press, Cambridge (1987)
3. de Swaart, H.C.: Intuitionistic Logic in Intuitionistic Mathematics. Faculteit der Wiskunde en Natuurwtenschappen, Nijmegen (1976)
4. Dragalin, A.G.: Mathematical Intuitionism. Introduction to Proof Theory. American Mathematical Society, Providence (1987)
5. Dummett, M.: Elements of Intuitionism, 2nd edn. Clarendon Press, Oxford (2000)
6. Gentzen, G.: Über das Verhältnis zwischen intuitionistischer und klassischer Arithmetik. Arch. Math. Log. Grundl.forsch. **16**, 119–132 (1974) (this paper was originally submitted to Mathematische Annalen in 1933 but it was withdrawn; Eng. tr. (on the base of the galley proof) in Szabo, M.: The Collected Papers of Gergard Gentzen. North-Holland, Amsterdam, 53–67 (1969))
7. McCarty, D.: On theorems of Gödel and Kreisel: completeness and Markov's principle. Notre Dame J. Form. Log. **35**, 99–106 (1994)
8. McCarty, D.: Constructivism in mathematics. In: Gabbay, D.M., Thagard, P., Woods, J., Irvine, A.D. (eds.) Philosophy of Mathematics, pp. 311–343. Elsevier, Amsterdam (2009)
9. Mostowski, A.: On various degrees of constructivism. In: Heyting, A. (ed.) Constructivity in Mathematics, pp. 178–194. North-Holland, Amsterdam (1957)
10. Mostowski, A.: Thirty Years of Foundational Studies. Lectures on the Development of Mathematical Logic and the Foundations of Mathematics in 1930–1964. Societas Philosophica Fennica, Helsinki (1965)
11. Novikov, P.S.: Mathematical Constructive Logic from the Point of View of Classical Logic. Nauka, Moscow (1977) (in Russian)

12. Pogorzelski, W.A.: Notions and Theorems of Elementary Formal Logic. Warsaw University—Białystok Branch, Białystok (1994)
13. Troelstra, A.S., van Dalen, D.: Constructivism in Mathematics. An Introduction. North-Holland, Amsterdam (1988)
14. Veldman, W.: An intuitionistic completeness theorem for intuitionistic predicate logic. J. Symb. Log. **41**, 159–166 (1976)

J. Woleński (✉)
Mickiewicza 26, 34-200 Sucha Beskidzka, Poland
e-mail: j.wolenski@iphils.uj.edu.pl