Studies in Universal Logic

Arnold Koslow
Arthur Buchsbaum
Editors

The Road to Universal Logic

Festschrift for the 50th Birthday of Jean-Yves Béziau

Volume II

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 the 50th Birthday of Jean-Yves Beziau
Volume II

Birkhäuser

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

Arthur Buchsbaum
Federal University of Santa Catarina
Florianópolis, SC, Brazil

ISSN 2297-0282
ISSN 2297-0290 (electronic)
Studies in Universal Logic
ISBN 978-3-319-15367-4
ISBN 978-3-319-15368-1 (eBook)
DOI 10.1007/978-3-319-15368-1
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014953260

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

Birkhäuser

Printed on acid-free paper

Springer International Publishing AG Switzerland 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 50th birthday.

The chapters in each of the two volumes (of which this is the second) 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 chapters, we want to express 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 ensure that tradition will continue and flourish. Many thanks, Jean-Yves. Our heartfelt thanks on your 50th 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

[1] Beziau [2].

defined on undetermined abstract objects that can be propositions of any science, but also data, acts, events." (p. vi)

(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 us 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 catalog 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 valid 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. The second observation suggests that not only are comparisons offered, but there may be also 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: 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

[2] Beziau [1].

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.

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 observations in (iv) **Beyond Syntax and Semantics** permits the extension of the program beyond the more traditional range of contemporary logical systems. As he stated, not only can we have the notion of consequence for scientific propositions, and nonpropositional, nonsentential 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 Peter Gärdenfors,[6] who developed a logic of propositions on 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

[3] Gabbay and Hunter [4].
[4] Private communication from D. Gabbay, 2005.
[5] Westerhoff [6]. The implication relation proposed for pictures is similar to one that Corcoran [3] proposed for propositions, as noted by Westerhoff.
[6] Gärdenfors [5].

functions is a consequence relation provided that functional composition is commutative and idempotent.

More specifically, (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 $f_1, f_2, \ldots, f_n$ and g in P, let (G) be the condition that

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

(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

Mai Ben Adar Bessos Bar-Ilan University, Ramat Gan, Israel

Marcos Antonio Alves Campus Jacarezinho – UENP, State University of Northern Parana, Jacarezinho, Brazil

Hajnal Andréka Alfred Renyi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary

Irving H. Anellis Peirce Edition Project, Indiana University, Indianapolis, USA
Purdue University at Indianapolis, Indianapolis, USA

Ignacio Angelelli Philosophy Department, The University of Texas at Austin, Austin, USA

Régis Angot-Pélissier Le Mans, France

Pedro Baltazar SQIG – Instituto de Telecomunicações Dep. Matemática – Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

Diderik Batens Centre for Logic and Philosophy of Science, Ghent University, Ghent, Belgium

Jonas R. Becker Arenhart Department of Philosophy, Federal University of Santa Catarina, Florianópolis, Brazil

Hilan Bensusan Campus Universitário Darcy Ribeiro, Instituto de Ciências Humanas, Departamento de Filosofia, ICC Ala Norte, Universidade de Brasília, Brasilia, Brazil

Jean-Yves Beziau University of Brazil, Rio de Janeiro, Brazil

Otávio Bueno Department of Philosophy, University of Miami, Coral Gables, USA

Carlos Caleiro SQIG – Instituto de Telecomunicações Dep. Matemática – Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

Mihai Codescu Institute of Knowledge and Language Engineering, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany

Alexandre Costa-Leite Campus Universitário Darcy Ribeiro, Instituto de Ciências Humanas, Departamento de Filosofia, ICC Ala Norte, Universidade de Brasília, Brasilia, Brazil

Didier Dubois Institut de Recherche en Informatique de Toulouse (IRIT), Université Paul Sabatier, Toulouse Cedex, France

Luis Estrada-González Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México, Mexico City, Mexico

Dov M. Gabbay Department of Computer Science, King's College, London, UK
Bar-Ilan University, Ramat Gan, Israel

Katarzyna Gan-Krzywoszyńska Adam Mickiewicz University, Poznań, Poland

Yvon Gauthier Department of Philosophy, University of Montreal, Montreal, Canada

Edelcio Gonçalves de Souza Departamento de Filosofia, Universidade de São Paulo, São Paulo, Brazil

Ivor Grattan-Guinness Middlesex University Business School, Hendon, London, UK

Dany Jaspers KU Leuven campus Brussels, Brussels, Belgium

Décio Krause Department of Philosophy, Federal University of Santa Catarina, Florianópolis, Brazil

Oliver Kutz Institute of Knowledge and Language Engineering, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany

Azriel Laufer Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel

Itala M. Loffredo D'Ottaviano Unicamp, Philosophy Department, Centre for Logic, Epistemology and the History of Science – CLE, University of Campinas, Campinas, Brazil

Sérgio Marcelino SQIG – Instituto de Telecomunicações Dep. Matemática – Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

Amirouche Moktefi Chair of Philosophy, Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Tailin, Estonia

Alessio Moretti Nice, France

Till Mossakowski Institute of Knowledge and Language Engineering, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany

Istvan Németi Alfred Renyi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary

Fabian Neuhaus Institute of Knowledge and Language Engineering, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany

Sergei Odintsov Sobolev Institute of Mathematics, Novosibirsk, Russian Federation

Hitoshi Omori The Graduate Center, City University of New York, New York, USA

Olga Pombo Faculdade de Ciências, Centre for Philosophy of Science of the University of Lisbon, Universidade de Lisboa, Lisboa, Portugal

Henri Prade Institut de Recherche en Informatique de Toulouse (IRIT), Université Paul Sabatier, Toulouse Cedex, France

Giuseppe Primiero Department of Computer Science, Middlesex University, London, UK

Fabien Schang Higher School of Economics, Department of Philosophy, National Research University, Moscow, Russian Federation

Toshiharu Waragai Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Tokyo, Japan

Christian de Ronde Philosophy Institute “Dr. A. Korn”, University of Buenos Aires, CONICET, Buenos Aires, Argentina
Center Leo Apostel and Foundations of the Exact Sciences, Brussels Free University, Brussels, Belgium

Personal Recollections About JYB by Newton da Costa and Others

Katarzyna Gan-Krzywoszyńska

Abstract The aim of this chapter is to present the personal recollections of some of Jean-Yves Béziau's friends and collaborators from all over the world. The chapter is divided into sections written by his supervisors, collaborators, and disciples, describing their first encounters and collaboration, and the inspiring work of Jean-Yves Béziau. They reflect on his rich personality, many different interests and talents besides logic, i.e., philosophy and art (music, film, paintings, photography) from the perspective of his endless travels and global organizing activity.

Keywords Universal logic · Paraconsistent logic · Philosophy of logic

Mathematics Subject Classification (2010) Primary 01A70 · Secondary 03B22, 03B53

1 Licence to Think – My Friend Jean-Yves Béziau

I met Jean-Yves Béziau (aka JYB) for the first time in January 1991 during one of my systematic visits to Paris. He was then a young student interested in logic, particularly in paraconsistent logic. We discussed several topics related to these subjects and I became very much impressed with his personality and some of his ideas. Therefore, during my stay in Paris, with the help of Michel Paty, I tried to get a French scholarship for him in order that he would be able to spend some time in Brazil at the Department of Philosophy at the University of São Paulo, where I had a group of young colleagues and graduate students in logic and philosophy of science. Béziau got the scholarship and spent one academic year in Brazil.

He was an active participant in my seminar at the University of São Paulo, he decided to enter in our Ph.D. in Philosophy program, we collaborated in some joint papers, and he passed his Ph.D. thesis on philosophy of logic under my guidance in 1996. However, earlier he had returned to France for some time and obtained his Ph.D. in Mathematics at the University of Paris.

One of his first results that I cannot forget owing to personal reasons concerns the cut-elimination theorem and Gentzen's formalization of paraconsistent calculi, really a nice result. But the fact is that he devoted himself to various logical and philosophical themes.

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_1

His field of research may be classified in three basic areas: paraconsistency, universal logic, and the philosophy of logic and science. In these three domains, he made significant contributions.

However, Béziau is not only a logician and a philosopher, but he is also an excellent administrator, organizing several meetings and congresses of logic and philosophy in Brazil and abroad. This fact is of fundamental relevance for a country such as Brazil, contributing to its cultural development in domains of knowledge in which it is so lacking.

Béziau was one of my best graduate students, and today is one of my best friends. I believe that his independence of judgment and tendency to see the central problems of a given field of knowledge are evident in the following portion of a letter he sent to me from Wrocław, Poland, in November 1993 (the original was written in French):

> "Considering the concept of structure as fundamental, we always start with a structure already given, which presupposes a notion of trivial identity, each element is identical to itself, two different elements are different. Then we can consider different notions of identity, among which the one mentioned above seems the most interesting.
>
> What about quasi-objects? I think the quasi-objects can be precisely be those objects that are logically identical, but not in reality, hence the paradoxes of quantum physics. I think we can reject Leibnizian identity simply by noticing that it is not an absolute identity but a relative one. If we make additional determinations, if we complexify the structure, then objects that seemed to us identical will turn to be different. If we then consider that there is no fundamental structure, but a hierarchy of more or less complex structures each corresponding to a certain representation of reality, there is therefore no absolute concept of identity, since to change the structure is to change the notion of identity.
>
> A different approach about the theory of proposition is what we might call the *axiomatic theory* of proposition, recently developed in particular by Suppes. We consider the proposition as a primitive term and we are looking for axioms characterizing it. The circularity of this approach is even more striking: the construction of an axiomatic system presupposes the notion of proposition, in particular an axiom defining what a proposition is.
>
> Getting back to Curry, he takes as a starting point the notion of formal system and he considers that mathematics and logic in particular are part of a general theory of formal systems. His definition of formal system is based on a "formalist" approach, more or less close to the conception I called the materialist conception. The ontological nature of what he calls an "ob" is not very clear.
>
> When certain objects are given, we can give them such or such names, following the interpretation that we have in mind and to which we are aiming at, in fixing some determinations, in one way or another, to these objects. The philosophical choice is performed at the first level, we are therefore not really committed when speaking of "proposition," "line," etc.
>
> In my thesis, I developed to the full the tautological conception of logical truth, to see up to which point we can follow this direction. However, I think the tautological conception is not the only one, and can be included in a much more general context, considering structures of type $\langle L; \vdash \rangle$ where $\vdash$ is consequence relation defined on a set of objects L without a proper structure.
>
> The term *abstract logic* seems to me to fit well for these kinds of studies, because it means in some sense we make *abstraction* of the nature of the elements of L by opposition to the usual method, *formal logic*, where the nature of the objects is fixed by the distinction form/content. The expression *universal logic* seems also excellent to me, especially in relation with the idea of *universal algebra* and the Bourbachic conception of mathematics. We can be *universalist* without thinking that there is only one logic.
>
> I believe that if we want to go further and better understand the Bourbachic conception notion (or to reinterpret it), we must try to clarify the concept of *type* of structure type as opposed to *species* of structure (Warning! I am not using these notions of species and type according to the definitions of Bourbaki; cf. pages 130 and seq.). We should in particular precisely define the notion of type of structure which intuitively corresponds to a thought frame.

Regarding the new solution, and its relation to the usual method, I would like to make some additional remarks. The ordinary method today may seem natural, but it is rather a matter of habit. In fact this method did not appear suddenly, it is the result of a long development and the confluence of two very different streams: on the one hand, the linguistic stream (Frege, Peano, Carnap) and on the other hand, the stream of "structuralist" abstract mathematics. Frege is now considered as the father of the theory of quantification, but at the time he was writing the *Begriffsschrift*, the notion of structure was still almost inexistent. It is only in the 1950s that these two streams merged with the development of model theory. However, this wedding is not quite harmonious, and contains certain defects, the result of the superposition of two disparate ways of thinking.

Regarding quantification, everyone agrees that the intuitive interpretation of quantification corresponds to infinite conjunctions and infinite disjunctions. However, according to the usual method, there is no simple formalization of this intuitive interpretation, because it would be necessary to consider the proposition $\exists x Rxa$ not only as an infinite disjunction, but also as an infinite conjunction of infinite disjunctions, because the undetermined constants R and a must range over the whole class of structures.

My method permits one to avoid this problem by starting to work within a particular structure. As the fundamental idea of the tautological concept of logical truth is that logical truth should not depend on the nature of the structure, the notion of logical truth that will be defined will be the same as if we had chosen another structure, so we do not lose any generality."

Newton da Costa
Federal University of Santa Catarina
Florianópolis, Brazil

2 A Letter to Professor Jean-Yves Béziau

Department of Logic and the Methodology of Sciences
University of Wrocław
3 Koszarowa Street
PL 51-149 Wrocław
Poland

December 5, 2014

Dear Professor Béziau, mon cher collègue et ami,

I am writing to you on the occasion of your upcoming 50th birthday on January 15, 2015. On this day, I would join the many logicians and philosophers from every continent who will be sending you their best wishes and congratulations...

I would also take this opportunity to sketch a few reminiscences of how we became acquainted, and our work together. The first time I encountered your name, though not you in person, was when the Section for International Collaboration at my university, the University of Wrocław, approached me in a letter dated April 8, 1992, asking me to give you a 10-month academic fellowship with the Department of Logic and the Methodology of Sciences (on a French government stipend). After reviewing your research proposal (projet d'étude), which began with the words,

> «Les logiciens polonais furent au début de ce [XX^{e}] siècle les principaux fondateurs de la logique moderne. ... Une des particularités de l'école polonaise de logique est qu'elle a toujours considéré comme primordiales les questions philosophiques alors même que la logique se faisait de plus en plus mathématique. Il en résulte une conception très générale de la logique étroitement liée au problème des fondements de cette science.»,

... and noting your extensive training in the foundations of logic and philosophy (from Lycée Henri IV, Université Paris 1 & Paris 7, and the University of São Paulo), it did not take me long to agree: on April 23, 1992, I informed my university's authorities that you would be a research fellow in my charge from October 1, 1992 to July 31, 1993. Things went well, and the following year the term of your fellowship was extended by another five months, to December 31, 1993.

We met in person in early October, 1992, at Wrocław's main train station, Wrocław Główny; after exchanging greetings I drove you to your hotel in my Fiat 126p, known in Poland as a "little" (*mały*) Fiat. I do not remember exactly, but it was probably the hotel attached to the Polish Academy of Sciences, at 75 Podwale St., near the German consulate. From mid-October on, you lived in Apt.18–21 Plac Grunwaldzki, a small flat the University has for visiting academics, at the κέντρον of Wrocław's old intelligentsia neighborhood, where the polytechnic and university communities meet. My *mały* Fiat again got you and your bags there. But I will leave those details for another day, interesting though they be, and try to tell you how I view the fruits of your time in Wrocław from the vantage point of today.

You learned about Polish logic, especially its approach to the theory and methodology of sentential calculi. At your insistence I suggested lots of contemporary authors you should read, such as (in alphabetical order) G. Bryll, J. Czelakowski, W. Dzik, W. Dziobiak, J. Kabziński, J. Łoś, G. Malinowski, M. Omyła, J. Perzanowski, W.A. Pogorzelski, T. Prucnal, S.J. Surma, R. Suszko, P. Wojtylak, and A. Wroński. Your work frequently cites their investigations and results; hence you are their heir and successor.

You established personal contacts with many logicians and university centers of research into logic in Poland. I remember that the first such contacts I helped you establish were with people at universities in Łódź (Malinowski), Kraków (Perzanowski & Wroński), Kielce (Prucnal), and Katowice (Dzik & Wojtylak). Later, over the next two decades, you expanded your contacts and collaboration to other logicians and centers of logic research in Poland, e.g., the University of Toruń (Jaśkowski's Memorial Symposium in 1998), Adam Mickiewicz University in Poznań, and Jan Długosz University in Częstochowa (the conference "Application of Logic in Algebra and Computer Science," Zakopane 2006).

In 1992–1993 you were an active participant in my seminar in the Logic Department in Wrocław. I can confirm this by recalling the titles and dates of your lectures (in a historic building and room: 36 Szewska St., 4th floor)...

(a) "On abstract universal logic" (10 December 1992);
(b) "Rules and derived rules" (21 January 1993);
(c) "On Russell's paradox and the liar antinomy" (1 April 1993);
(d) "On a problem posed by T. Prucnal" (6 May 1993);

(e) "What is negation" (14 October 1993);
(f) "On logical truth" (18 November 1993).

Notice that all those dates were Thursdays, which was the traditional day for the seminar to meet. Some of your lectures were delivered to joint meetings of the Logic Department and the Wrocław branch of the Polish Philosophical Society. It may be worth adding that 1993 abounded with foreign guests; the Department enjoyed talks by Peter I. Bystrov (Moscow), Thomas Uebel (Boston), Antonio de Freitas (Covilhã, Portugal), V.V. Rybakov (Krasnoyarsk), and Jan Tarski (Berkeley).

By the way, do you remember the trip we made to the sacred Mount Ślęża in mid-May, 1993, partly on account of Rybakov's being just then in Wrocław? It has had so many different names over the ages: *monte Silentii*, *monte Silencii*, *Monte Slenz*, *Zobtenberg*, *Góra Sobótka*. Remember the prehistory of it: how, in the Neolithic period, at least as far back as the 7th century BC, the heathen tribes of the Lusatian culture were living there? And then it was settled by Celts, and later by Germanic Lugians? Our group, consisting of you, Rybakov, Max Urchs, Jacek Hawranek, the late Dr. Krzysztof Zielnica and myself, went to explore the remains of those cultures in the sanctuary, and then we "conquered" the summit. A nice photograph still survives, documenting that experience, recording the bright green spring colors. Hawranek recalls that during the trip, perhaps under the influence of the history of the place, you gave us a mini-lecture on *La Chanson de Roland* – the oldest surviving major work in French literature.

While in Wrocław, in 1993, you wrote two papers: "Recherches sur la logique abstraite: logiques normales," and "More about the connection between the Axiom of Choice and Lindenbaum's extension lemma." The first was published in the "*Logika*" Series of our University's journal *Acta Universitatis Wratislaviensis – Logika*, vol. 18 (see Zbl 1023.03523). The second came out later under a somewhat different title: "La véritable portée du théorème de Lindenbaum-Asser" 'The real import of the Lindenbaum-Asser theorem' (see Zbl 1013.03033).

Over the time you were here, 1992–1993, you also completed a doctoral dissertation, titled "*Sur la vérité logique*," which you submitted for a Ph.D. in philosophy at Université Paris 1 – Panthéon–Sorbonne in October, 1993. [Your advisor was Philippe de Rouilhan, who maintained close relations with logicians from the University of Łódź. A copy of your dissertation, which I keep to this day, was graciously given me by its author on June 24, 1993, along with another typescript, "*La logique paraconsistante*", by Newton C. A. da Costa and Jean-Yves Béziau.] I did not have any direct influence on the content of this dissertation "on logical truth". I recall discussions with you on the topics raised in it – discussions which were conducted at 36 Szewska Street over quantities of coffee and tea. My role consisted mainly of suggesting relevant readings from the literature on the subject, which can be found in the list of references at the end of the thesis. Despite the fact that my contribution was slim, you did me a great honor; the last sentence of the dissertation reads:

> "L'auteur tiens à remercier MM. les professeurs Newton C.A. da Costa et Jan Zygmunt qui l'ont accueilli respectivement au Brésil et en Pologne."

The fruits of your Wrocław sojourn are further evident in your second dissertation, which you developed and finished in the first half of 1994 in Paris, where you went after

Wrocław. This second dissertation, titled "*Recherches sur la logique universelle (excessivité, négation, séquents)*" (see MR 1645129 (99f:03008)), was submitted to the Université Denis Diderot – Paris 7 in partial fulfillment of the requirements for a *Doctorat de Logique et Fondements de l'Informatique* – a type of doctorate degree, existing in the French system, distinct from a Ph.D. Professor Daniel Andler was your advisor, and I was honored to be a member of your examining committee for this work. The discharge of my duties in this role, which would have been a pleasure for me under any circumstances, involved on this occasion the additional delights of visiting Paris at the beginning of July, 1995, and being able to meet Marcelo Tsuji and Richard Zuber while I was there.

Dear Professor Béziau – *Drogi Jubilacie*,
We could reminisce on and on – But this letter should come to a point, so I shall save further memories for your next jubilee. Nevertheless, I must add one more thing: It is said that no man lives by bread alone... and in our case "bread" can be taken to mean "logic." I and my Wrocław colleagues remember your many and varied interests: in art, in literature, and film. In your spare time you painted in oils, and wrote a novel. You were passionate about film, and the history of film. You established close working relationships with the Wrocław division of Alliance Française, and its director, Madame Teresa Pękalska. Through them, you arranged and led enthusiastic classes on French cinema. You held Polish filmmakers in high esteem, particularly Krzysztof Kieślowski, and you were adamant in your assessment that his *Trois couleurs: Rouge* was by far his best picture. I also remember your writing to me, that once at a film festival somewhere in far-off Brazil you were profoundly touched by a particular Polish entry – a psychological drama about loneliness, directed by Dorota Kędzierzawska, titled *Wrony* (Crows).

Nearly 23 years ago, when you were writing your *projet d'étude* with the intention of moving to Poland and studying Polish logic, you stood, in fact, before a great unknown. But even then it was clear that your plans and intentions were underpinned by an unshakeable belief in their rightness. Later you would go on to harness this same strength of conviction in the service of your supreme idea, the idea of *universal logic*, and to bring the idea to life in the form of an international journal, *Logica Universalis*, a book series, "Studies in Universal Logic," and the creation of World Congresses and Schools on Universal Logic. Since then many Polish logicians have been invited by you to collaborate in the field of universal logic and are deeply grateful to you for the opportunity. In particular, I, and three of my colleagues, Janusz Czelakowski, Piotr Wojtylak, and Robert Purdy, sincerely thank you for providing us with this chance to be part of your enterprise, and we send you our heartfelt *félicitations*.

From all of us I would say: Stay forever free-spirited and young in your soul. Keep forever true to your innermost compass. There are precious few Rimbauds in this Hobbesian world. Let us not lose one of its last lights.

Yours sincerely,
Jan Zygmunt

Translated by Robert Purdy

3 Dr Yes-and-No

Between 1990 and 1992, I worked at the Department of Computer Science of the University of Buenos Aires (UBA), in the Group for Artificial Intelligence (GIDIA), led by Adolfo Kvitca, in close collaboration with Carlos Alchourrón's group of logic at the Law Faculty of UBA. During that time, we received many visiting researchers from a variety of countries, including Newton da Costa (who inspired me to move to Brazil), David Makinson, Peter Gärdenfors and, in April 1992, a young, brilliant, and enthusiastic French logician, named JYB. After his very nice talk on a general theory of valuations (which pleased all the audience, including Carlos Alchourrón), we talked about logic in my office at UBA for about an hour. I was surprised at his broad knowledge of the subject, despite his being young.

We met again in July 1994 in São Paulo. At that time, both of us were studying for our respective Ph.D. degrees at the University of São Paulo (USP). Jean-Yves was concluding his first Ph.D. in Paris under the supervision of Daniel Andler while working on his second Ph.D. thesis at the Philosophy Department of USP under the supervision of Newton da Costa, and I was working on mine at the Department of Mathematics under the supervision of Francisco Miraglia. Our friendship began then, spending long nights at the exotic bars of downtown São Paulo, or watching classic movies and listening to music together with other friends in my small apartment at Estação da Luz, over a period of several years. He moved to Rio de Janeiro and Petrópolis by the end of 1995, but our cultural meetings have continued through to the present day, due to periodic visits in both directions.

From an academic perspective, I learned a lot of Logic from Jean-Yves' papers and personal communications. He has plenty of original ideas, and his insights are remarkable and inspiring. He always surprises me with original perspectives and observations. His tireless work as a promoter and disseminator of Logic as a broad discipline, interacting with several other areas of knowledge, is also outstanding.

I would like to send my warmest congratulations to Jean-Yves for his remarkable academic work and career (observing that he is only 50!), and I would like to thank him for his sincere friendship during these years.

Marcelo Coniglio
UNICAMP
Campinas, Brazil

4 From St. Petersburg with Snow

I first met Jean-Yves through my colleague Vladimir Vasyukov; I cannot now remember the exact dates and circumstances but most likely this happened in 2003 in St. Petersburg. In 2005, I've got an invitation from Jean-Yves to give a tutorial on Categorical Logic at the *First World Congress and School on Universal Logic (Unilog)* in Montreux (Switzerland). In the intellectual and political climate of the 2000s, which especially emphasized the need to tolerate multiple systems of thought as well as multiple systems of logic, the title

"Universal Logic" sounded very provocative. This name could be used for a certain system of logical norms offered as the only "true" logic – against the current trend to logical pluralism. However, I quickly learned that Jean-Yves' idea was different and actually more interesting. It was about bringing together many different people working in various areas of logic (including people coming from outside the current academic mainstream), making them to talk to each other and finally asking them to describe a hypothetical notion of "universal logic," which, if it makes any sense at all, remains invariant through a wide variety of things found under the name of logic on the contemporary intellectual market. I am very grateful to Jean-Yves for this kind invitation, which I used for systematizing my own knowledge and ideas about the subject and for learning a lot of contemporary philosophical logic.

This and some other meetings organized by Jean-Yves and our personal discussions gave me quite a lot of inspiration for my further work. And it still inspires an important part of my research today. Among other things, this concerns the concept of logical translation, i.e., translation between conceptual frameworks or "different logics." This sort of translation not only allows different systems of logic to talk to each other but also determines the core structural features of these very particular logical systems. See my abstract "Logical Forms versus Translational Categories" in the Unilog2007 Handbook and also, for a keen approach, "Homotopical Categories of Logics" by Peter Arndt in the first volume of the present Festschrift, pp. 13–58. I believe that this translational view on logic is very promising and deserves to be further developed. I also believe that Jean-Yves has made a great contribution to this and other research projects, which fall under the scope of universal logic in Jean-Yves' very liberal and always friendly sense of the word.

Happy Birthday, dear Jean-Yves, and Many Many Happy Returns!

Andrei Rodin
Saint Petersburg University
Saint Petersburg, Russia

5 "Torpedo" Béziau

I have been kindly asked to accomplish the following task, namely writing some words about the hero of the day, viz. Professor JYB ("Beez-you," with the English-worldwide accent), at the occasion of his half-centenary. It is my pleasure to reply positively to such a proposal, all the more that a lot has to be said about this public character. Let me just say a couple of words about him; however, English is not my *Muttersprache*, and this is only a sketchy portrait, but here it is, as far as I can honestly picture it.

A well-known football club in Moscow is called "Torpedo Moscow,"[1] where the first word has a blatantly military meaning that everyone can guess. But another reading also refers to the sea-fish philosophers are very familiar with: it is the fish which served as a nickname for Socrates, due to his legendary capacity to address people in the street with

[1] Note to the purists: write "торпеда" in Russian, but " Торпедо Москва" for the whole name of the team with a genitive declension of the word.

simple questions before leaving them with serious headaches. Jean-Yves is exactly like one of these provocative dialecticians, juggling with concepts just as some humorists may play with boundaries.

But Socrates also struggled against the Sophists, and Jean-Yves seems closer to these when dealing with paraconsistent flows. Like a Sophist, but in a neutral sense of the word; not the pejorative one which so harshly affected the official memory of genuine thinkers from ancient Greece such as, e.g., Heraclitus. And yet, the Sophists flirted with conventions; whether for want of money in life, or for any other reason. Anyway, it is much fairer to say that Jean-Yves' sarcastic mind nicely matches with that of someone like Diogene of Sinope. Had Kripke asked him anything, Jean-Yves would have ordered Saul to stand out of his light; certainly, because he could not read his non-truth-functional matrices on the wall, or the like.

Indeed, one of the best moments I experienced during Jean-Yves' talks was the very first one in Montreal on September 2003. We were attending a conference organized by the Francophone Society for Analytic Philosophy (SOPHA). Jean-Yves talked about possible-world semantics, a famous "hot topic" among those one is delighted to put on the table for very philosophical purposes. What of Jean-Yves? He simply threw the modern idol on the floor, mocking those who venerate the Holy Accessibility Relationship and claiming that this machinery was nothing but an extensive by-product of algebraic bi-functions. I have to point out that it was my first time as a conference speaker. Not sure that most of the usual speakers had such a subversive mission in mind, speaking in a distant way to astonished people while eating grape breads. Then the discussion came to be somehow "dynamic," once the idol had been hurt and underestimated by an incredulous torpedo.

Likewise, all of Jean-Yves's talks play with pedagogical pictures and so general ideas that most of the listeners do not want to take seriously. Wasn't it the same with the foolish Diogenes in ancient Greece, whether for right or wrong reasons? Identity, Proposition, Truth-Value, Opposition, Logical constants … our Professor is used at launching insightful discussions with simple ideas, and it is not the least courageous philosophical act to do this. Borrowing from Timothy Williamson's words: "To be precise is to make it as easy as possible for others to prove one wrong. That is what requires courage."

From excessively simple words to really strong points, let me quote some main contributions of Jean-Yves: the S5-translation of paraconsistent negation, following Gödel's S4-translation of intuitionist negation, non-truth-functional matrices non-classical constants, and a reply to philosophical objections concerning the nature of paraconsistent negation (cf. Slater's challenge), a revival of the ancient theory of opposition, following Robert Blanché's developments, and the launching of a bunch of new results (especially with one of his former students, Alessio Moretti).

Finally, he is an everlasting globetrotter going throughout the world and spreading in exotic countries some exciting events like *Unilog*, *The Square of Opposition*, together with a top-level scientific journal (*Logica Universalis*) and a collection of book series for logicians, mathematicians, and philosophers, and a clear editorial success *à la Gabbay*.

I have only talked about what I am most familiar with, being aware that a huge number of technicalities Jean-Yves faced largely go beyond my mental capacities. However, I still take him to be the one who stimulated my first pure reflections about formal semantics, negation, or the logical concept of opposition, until the day when he agreed to be a jury member for my Ph.D. defense on March 2007. Something about epistemic modalities … here is just a way to refresh his mind with Kripkean possible worlds, or Hintikka's epistemic logic.

Jean-Yves, in a nutshell is a logical comet between Molière and Diogenes, and not far from Montherlant in attitude. Not regarding his style of writing, I must confess, due to Jean-Yves's inner struggle with the French syntax and his clear taste for rational abstractions rather than lexical subtleties. Someone who never forgets to point at those Mr Jourdains of philosophical logic, i.e., those playing seriously with truth valuations just as kids wage war with tin soldiers.

Now I should return to my own conceptual toys. I just sketched our beloved scratching powder in a very subjective *Gonzo* style that some readers might find appropriate or not. *Tertium non datur*, notwithstanding the blatant affinity between many valuedness and our dear Professor, alias Jean-Baptiste Torpedo Béziau of Sinope.

To the Baron of Chambourcy, pending another future heading for our common master of situationist logic.

До скорого, дорогой Жан-Ив![2]

Fabien Schang
National Research University Higher School of Economics
Moscow, Russia

6 The Man with the Golden Thought

I do remember the Jean-Yves of the old days – pretty much the same as he is now but with time enough on his hands for long rides toward artificial cities on Brazilian coaches. He was an enthusiast of new breeding grounds: Schopenhauer, universal algebra, Badiou and non-classical logics were his ingredients at the time. I recall universal logic being roughly ready in his head as a road if not as a castle. I suppose he likes castles less. Then he hit the road. I met him again years later for a night out in Palo Alto. He was sort of a habitué. He spent the evening showing me the core of the place. And how unfulfilling it was! I suppose it is a logician's way to go places – see the (hard) core of it and you see its imbrications. His approach to seeing places, I suspect, was like the ideas of universal logic he had been cherishing: make a mosaic of these cores. No place, like no logic, is fully and universally satisfactory. But you can hope for something out of a cubism of (hard) cores.

Hilan Bensusan
University of Brasília
Brasília, Brazil

[2] "See you soon, dear Jean-Yves!"

7 Jean-Yves Béziau's Imprint on Universal Logic

I first met Jean-Yves in 2005 in Mumbai when we both were on our way to a conference in Pune. He had rented a car and we traveled there together, and afterward we went on for a spontaneous roadtrip. Altogether we spent around two and a half weeks together, connecting our thoughts, discussing logic and life, and encountering incredible India. This was a marvelous start and our paths kept crossing, and will hopefully continue doing so for a long time.

When I first encountered the idea of universal logic (this was before meeting Jean-Yves), it struck me as natural and convincing. However, I could not imagine how such a grand and general undertaking could be pursued in a focused way, how it could be more than just disconnected research, driven by individual curiosities, and maybe somewhat randomly forming a bigger picture over decades. Seeing Jean-Yves in action taught me otherwise.

With the efforts he put into articles, lectures all around the world, creating conference series and book and journal series, Jean-Yves managed to transmit his ideas widely and to convey the impression that there is indeed a big picture to be understood, a goal to reach for, and that conscious effort can bring us closer to it. He got a lot of people engaged in the project.

One reason that he succeeded is, of course, that the idea of universal logic itself is good.

Another reason is his remarkable openness to ideas, his huge willingness to learn from others who have thought about a subject, and his ability to connect the different things that he learns. He must have opened the eyes of quite a few people to the fact that what they are doing is related to different activities in ways that they never suspected (the square of oppositions is a prime example). This got them interested in universal logic, and made them see where they had something to contribute.

And the third reason is his own style in research (which is something that is not easily found in science). Careful analysis of situations and solid arguments should be part of any scientific action.

But Jean-Yves has two further trademarks: he asks fruitful and intriguing questions and puts a major emphasis on them. They are ambitious but just within the reach of making progress. And they are not rhetoric; he wants people to try and give answers (as exemplified by the Unilog contests). He is also great at bringing up examples, not just for clarification but also for pointing out our confusion about some ideas, and showing places where we can strive for a better understanding (think of the collapsing problem in the fib ring of logics).

All of this made universal logic the project of many, while bearing Jean-Yves's personal imprint. So, Jean-Yves, keep up your inviting and challenging style; it has served us well! Happy birthday!

Peter Arndt
University of Regensburg
Regensburg, Germany

8 The World Is not Enough

I first met Jean-Yves in 2004 in Portugal. That was also my first encounter with his universal logic project, which was advertised on that occasion. I immediately had a strong feeling that I had met with the long awaited home in logic for my activity. Gradually, we have also become good friends. Jean-Yves is a highly unusual character in contemporary academia. In a scientific world dominated by narrow specializations, autism, and ruthless competitiveness, he brings in a wide perspective and an impartial approach. He persistently provides dedicated support for ideas outside mainstream scientific activity. Many of them, albeit very interesting, can be marginalized or even ignored just as an effect of the toxic competitiveness that dominates our academic world. For this, I regard Jean-Yves as a sort of intellectual hero.

This kind of intellectual generosity may also bear some costs. The events around Unilog are too big compared to how much the highly substantialist Western scientific culture is capable of accommodating the nonsubstantialist thinking underlying universal logic. The latter seems to be related, at least in spirit, to other scientific traditions of humankind that modernity has put in the shadows. In my view, there is a danger that authentic universal logic thinking may become marginal within Unilog events. But with wisdom and care, I think that this can be avoided and universal logic will fulfill the prophecy of some paramount scientists, as a true renaissance of mathematical logic.

Răzvan Diaconescu
The Romanian Academy & Școala Normală Superioară București
Bucharest, Romania

9 On Logic's Secret Service

I had a lively image of JYB even before we first met, back in 2001, for a logic event in Las Vegas, USA that he was organizing. A number of common acquaintances had told me about his work and interests, but also about his colorful and outspoken French–Brazilian character, and I must say that I was not disappointed. I had recently become interested in his work on the theory of valuations, which I found extremely interesting and which ended up being very influential to my own research on generalized compositional meaning and the combination of logics. We got along very easily from the first moment, and have kept in regular contact ever since. JYB visited me in Lisbon, Portugal on several occasions, and I also visited him in Neuchâtel, Switzerland, and most recently in Rio de Janeiro, the Land of the Future. Over the years, we ended up spending many hours discussing this and that, some of the happiest over a glass of wine, ranging from logic, science, mathematics, philosophy and art, to politics, and life in general.

Along with his talent for thinking out of the box, JYB has an eye for simple but profound problems. As Lloyd Humberstone put it, and I could not say better, JYB "has a knack for noticing areas of logical theory where we find ourselves with intuitions pulling in different directions." I do not know where JYB's wit comes from, in part certainly from his multidisciplinary background, but I am sure that his inquisitive globe-trotting nature is an important part of the equation. He seems to have visited every logician around the world, knows about their research, and is able to pick all these scattered inputs (we all know how scattered they can be) and understand the common ground. His universal logic initiative, including the World Congress and School, the journal, and the book series, is the remarkable result of his vision, but also of his resolve. I have done it personally, but it is only fair to thank JYB again, on this occasion, for such a remarkable indelible effort.

JYB and I have collaborated on several initiatives, events, committees, and projects, and have discussed many ideas for joint work. As JYB would put it himself, all you need is "um prroblemo interressante." We have identified several such interesting problems together but, as I write, my JYB number is sadly still 2. I am sure that one of the ideas that we have in the freezer will eventually come to life, but now it has become urgent: we shall not wait for another 50 years, and no Icelandic volcano will prevent that!

Happy birthday, aquele abraço.

Carlos Caleiro
University of Lisbon
Lisbon, Portugal

10 Jean-Yves: Logician and Globetrotter

How many people know that JYB published a paper on the philosophical conception of suicide in Schopenhauer? Surely many know him nowadays as a serial organizer of conferences all over the world, and as a prolific editor of journals and books, with topics related to intersections of logic with philosophy, mathematics, computer science, linguistics, psychology, theology, and arts. But who else would think about doing a complete Ph.D. in Brazil, after having defended a successful doctoral thesis at Paris 7? And who else would dream of becoming a professor in Brazil after acting as a professor in Switzerland? Who in the world would strive to organize a scientific conference at Easter Island?

The man who counts maybe as the only incarnation of Bourbaki as a philosopher has been for years tirelessly circumnavigating the globe, madly sailing into all sorts of research centers. This is the man who stops and takes pictures at stop signs, who brought logical insights to penguins in the far south, as well as to hardly accessible tribes in the Amazon, the man who has been both praised and condemned for smuggling turbo-polar paraconsistent logic into Poldavia.

His influence in my work started as soon as I did my first proper courses on logic. I have learned a lot from him about the intricacies of bivalence, about the many varieties of paraconsistency and of modalities, about the complexity of identity, about the paradoxes hidden inside the notion of translations between logics and revealed by the combinations of logics… and logicians. Since then we've met (often by sheer chance) in the most unlikely places, from a secluded island in South Africa to an old mosque in Turkey, from the streets of São Paulo to the Costa Rican jungle.

It's been almost 20 years now, Jean-Yves and we have still not gotten bored of living this adventure. Many thanks for being a good friend all along, and for making our logical path so much more exciting!

João Marcos
Federal University of Rio Grande do Norte
Natal, Brazil

11 On JYB

Basically, I describe my impressions of an icon, an idol I have. I remember very well when I met Jean-Yves. Indeed, I had already read one of his articles and, for that reason I was pretty much interested in meeting him personally. He was sitting on a bed pool in a hotel in South Brazil talking with a woman. He had a weird hair, a red nose, and I tried to approach him asking some questions on modal logics. But he declined to answer arguing that he did not know anything about modal logics. However, some days later, to my surprise, I started receiving by post packages of his articles on many aspects of modal logics! What a paraconsistent character!

Enjoying since then we have had a prosperous collaboration, where I learned too much, and he invited me to join his research group on universal logic at the University of Neuchâtel, Switzerland. There I recognized that I got much more than a Ph.D. supervisor, but a real friend, the kind of friend we find only two or three times in life.

Zooming in, he has a universal, generalist, and rare mind which attracts all kinds of collaborators and persons, making him a natural ambassador of logic in the domain of culture, connecting it to even the most unexpected topic. Being a workaholic, he always organizes meetings, writes papers, edits books in nonstop mode.

In his great care and attention to people, we discover a holy being, who is a gentle and helpful person: each human being is super valued when in touch with him. He tries to make each one develop its own powers, stimulating research and creative thinking in a way that is very difficult to describe.

After all, what really attracts me in his personality is that his way of thinking is not a provincial one because he has a very good philosophical *ouverture* to discuss all tendencies. Without prejudices he is always open to debates, and does not try to impose his concepts and ideas.

Unique, original, creative, not dogmatic, and with a good sense of humor, JYB is my favorite philosopher and logician.

Alexandre Costa-Leite
University of Brasília
Brasília, Brazil

12 The Real Universal Thinker

I have known JYB for 11 years. *The First World Congress and School on Universal Logic – Unilog'05* in Montreux was the first major congress I had ever participated in, thanks to Jean-Yves, also as a member of the organizing committee. First, Unilog was a huge success, an absolutely fantastic event in a very beautiful place by Lake Geneva in the Alps. It was the time when I decided to participate whenever possible in all the events organized by Jean-Yves.

12.1 Some Unknown Facts About Jean-Yves

Fact 1 Because he was living some time in Corsica he was given the nickname "Napoleon," but I am sure it was also due to his strategic abilities and logical genius.

Fact 2 He once wanted to be a rock star. It is said that in Switzerland he had a band with some other now renowned researchers. I think rock and roll may have lost a great star, but surely logic has a much more fun spirit thanks to the failure of Jean-Yves' music career. Anyway, I think that in this way he has more devoted fans.

Fact 3 He loves Smurfs and has a collection of Smurf figures. Jean-Yves considers this to be a very philosophical cartoon, and after many philosophical debates, I still think that Peanuts beats Smurfs, but this remains an open metaphysical problem.

Fact 4 His father was the first person who crosses the Sahara desert on a BMW motorcycle.

Fact 5 His mother made her communion in a dress made from a parachute of Antoine de Saint-Exupéry, souvenir or an emergency landing of the aviator in the property of her family.

Fact 6 Deep down he is a Polish guy with a Slavic soul. How do I know this? First of all, he is a great logician, he has a great imagination, he is extravagant, he has a special absurd sense of humor, and he is quite repetitive, but he is still very funny, that is, when he is not being extremely repetitive (like telling the same joke 20 times in the same evening). He is very inspiring and a truly great friend. He enjoys good, strong alcohol, herrings, and poppy seed cake. And he never gives up. And here the proof ends.

12.2 The Art (of Life) by JYB

He is an art lover and an artist at the same time. Perhaps, it is not even a love, just an organic part of his existence. Whether it is photography, drawing, music, painting, film or just beautiful shoes, he is an expert, admirer, and practitioner. He is a professional and sincere, absolutely snob-free, authentic art-lover and artist. He always listens to music, though in my opinion, a little too melancholic. Too sad!

His photography exhibition "World in the wrong way" perhaps shows his interest: the limits and borders he wants to cross and the obstacles he loves to conquer.

He travels around the world together with Catherine promoting their ideas, art, and a true and rare sense of scientific and artistic cooperation and dialogue, facing very different people, cultures, and systems, and thanks to this, many unexpected international, intercultural, interreligious friendships, connections and cooperation have started. And as they should be, they are very welcomed everywhere.

Jean-Yves has great taste in film, and is a cinematic connoisseur and professional. If you haven't seen it yet, please find and watch every video he has made with Catherine, movie, interview or filmed performance. I think it was the first time in the history that the main organizer of a congress made a professional movie concerning the main topic of the event. This is the case with "The Square of Salomé," where Jean-Yves plays Jean-Baptiste.

12.3 Paradoxes

His life is full of paradoxes. Maybe that is why he is so keen on paraconsistency and is searching for way to go beyond dichotomies.

- He takes many risks, but he is not reckless.
- He has his own strong opinions, yet he listens carefully to everyone's opinion.
- He has a great imagination and fantasy, but works hard every day like a monk.
- He may sometimes be shy; however, he is fearless in expressing any controversial claim and in defending his own – even the most unusual – opinions.
- He has what all philosophers, scientists, artists must have – a great imagination. But he is also down to earth and understands everyday problems.
- He works at the same time on the most abstract and refined questions in logic, and also does the most fundamental, basic work on a universal, global level.

12.4 Logical Mogul that Never Goes 50/50

I do not really know if anyone is aware of how hard he works. He personally takes care of so many things concerning congresses, publications, searching for cooperation, helping young and old researchers from all around the world. When it comes to the organization of congresses (like Unilog, Square of Opposition, Paraconsistency, Logic in Questions, and so many more), he worries over absolutely every detail: from the venue where the event will take place and the invited speakers, to the food, climate, prices, distractions, excursions, etc. He let me work with him, which was one of the most important experiences of my life, and I always look forward to spending time with him and Catherine. They are so incredibly talented and creative, yet such very down-to-earth people. We spent weeks together working in many countries around the world, and afterward I was always totally exhausted but extremely happy. Jean-Yves usually has to invest all his time, sometimes his own money, and definitely he is paying with his own health for the success of all these events, publications, and enormous activity.

I also think that his idea of “université mondiale”[3] should be seriously studied and developed. It is an antidote for the bureaucratic pandemic in the academic world around the world. He has proved that a wonderful academic career can be at the same time a fascinating adventure with artistic flavor.

I truly admire him and cherish our friendship; however, we usually have very different tastes in music, philosophy, and fashion, and many varying opinions. Although I do not like to say this, sometime later I usually have to admit he was right. I have learned so much from him. He inspires me, as I am sure he does so many other people, as well. He is so patient and looks for the best in people. And hopefully he finds it.

At the end, on the one hand, poor Jean-Yves hits 50, but on the other hand, with all due respect to his unquestionable achievements, let me quote Jack Nicholson when he received the Life Achievement Award, because I am sure that with Jean-Yves it’s the same story: “You ain’t seen nothing yet!”

Happy Birthday Jean-Yves!

Katarzyna Gan-Krzywoszyńska
Adam Mickiewicz University
Poznań, Poland

13 A Logician Who Loves Me

There is the man, the researcher in logic, in philosophy, the artist, the photographer, the painter, I live with all the men in him, in a train, in a boat, in a plane, in a car, in Rio, in Paris, in Valparaiso, in Beirut, in Buenos Aires, in Lisbon, in New York, in Montreal, in Melbourne, in Montreux, in Bombay, in Rome, in Bucharest, in Budapest, in Moscow, in Copenhagen, in Athens, in Poznań, in Berlin, all over the world I live with him, in a double room, in a hotel, in a castle, in a guest house; it is the universal life, because this man loves open spaces; he constantly creates and thinks all the time; sometimes he would like to take holidays from thinking, but this is impossible; he is animated by research, his passion, to gather people from around the world from many different disciplines, to try to understand and organize congresses where everyone can meet, discuss, exchange new ideas, if possible, not so easy to have new ideas, but for him, no question of settling for a little life reduced to a miniature.

He thinks big, he sees large, he sees universal, he offers, he shares; his congresses and his journals show it; living with him is the great life, the wonderful life, but one must be strong, be obstinated, brave, generous, be creative, inventive; our thoughts are creations,

[3] See J.-Y. Béziau, “Les universités face à la globalisation: vers une université mondiale?,” In Journée de l’Unesco 2004, vol. 10, F. Naishtat (ed.), Unesco, Paris, 2006, pp. 207–211.

his thought is constantly active and creative, but above all he believes in potential because life offers only what we give; he gives everything, he believes in life, its always strange mystery he wants to understand and break through, he believes in man's power of reasoning, and he is a spiritual man hidden behind all his research.

Happy birthday Jean-Yves!
With all my love, Catherine Chantilly

K. Gan-Krzywoszyńska (✉)
Adam Mickiewicz University, Poznań, Poland
e-mail: katarzyna.gan.krzywoszynska@gmail.com

Logical Autobiography 50

Jean-Yves Béziau

Abstract After describing my family background and interest in mathematics and philosophy at school, I explain how I became interested in logic when studying at university in Paris. I describe how I discovered the work of Newton da Costa on paraconsistent logic, how I met him in Paris, and then go to do research in Brazil, Poland and USA before defending my PhD on universal logic at the Department of Mathematics of the University of Paris 7 and a PhD on logical truth at the Department of Philosophy at the University of São Paulo. I recount my postdoctoral years in Rio de Janeiro and Stanford and how I settled down for a couple of years in Neuchâtel, Switzerland, organizing the First World Congress on Universal Logic in Montreux and launching a journal and book series on this topic. I tell how I came back to the land of the future, starting more editorial projects, organizing logic events around the world (Paris, Beirut, Vatican, Lisbon, etc.) and continuing to develop research related to all aspects of logic (mathematical, philosophical, semiotical, historical). I finish with a short overview of my future projects, supported by a list of potential papers and books following a complete list of my writings to date.

Keywords Autobiography · Universal Logic · Paraconsistent Logic · Square of Opposition

Mathematics Subject Classification (2000) Primary 01A70 · Secondary 03B22, 03B53, 03A05, 03B45

> A designer knows he has achieved perfection
> not when there is nothing left to add,
> but when there is nothing left to take away.
> —Antoine de Saint-Exupéry

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_2

1 Why an Autobiography?

I[1] see two good reasons to write an autobiography: on the one hand, to reflect on and become more conscious of what we have done; and on the other hand, to inspire and provide guidance and advice to others. I have read a few autobiographies that were very useful to me. Let me give two important examples: *Confessions* [41] by Jean-Jacques Rousseau and *I Want to Be a Mathematician: An Automathography* [22] by Paul Halmos.

I read *Confessions – Part I* when I was a teenager and it continues to have a strong influence on me to this day. I do not know much about Rousseau's philosophy and this is the only book by him that I have read in its entirety. I feel close to Rousseau because I have been connected to several places where he lived: Geneva, Annecy and Neuchâtel. However, this is only one aspect. This geographical connection is not necessarily an emotional one – it is more as if he were there, part of the landscape. In fact, I do not feel a special affinity with Rousseau's personality; my own personality is perhaps closer to that of his rival, Voltaire. Yet I was much impressed by the way Rousseau conducted his life, always innovating, thinking about all kinds of topics while traveling on foot through Europe.

Paul Halmos's autobiography is quite different, since it is pretty much restricted to his intellectual life. As with Rousseau, I do not feel a real affinity with Halmos's personality. However, what I like is that he was not just a follower, and that he had many interesting ideas regarding both his main area of specialisation, mathematical logic, and the way he conducted his academic life. He not only gives a lot of useful tips, but also reflects on the activity of the mathematician, worthy of a thousand papers in the philosophy of mathematics. A typical controversial remark by Halmos is: "The best notation is no notation; whenever it is possible to avoid the use of a complicated alphabetic apparatus, avoid it. A good attitude to the preparation of written mathematical exposition is to pretend that it is spoken. Pretend that you are explaining the subject to a friend on a long walk in the woods, with no paper available; fall back on symbolism only when it is really necessary." ([23], §15. Resist symbols.) I like that, although I do not necessarily agree with it, since I believe in the power of (non-trivial) symbolism – cf. my recent book *La Pointure du Symbole* (2014).

Of similar interest to me is Gian-Carlo Rota's *Indiscrete Thoughts* [39] with its remarks about teaching, publishing, lecturing, accurate testimonies and provocative ideas about the relation between logic, mathematics and philosophy – including his essay "The pernicious influence of mathematics upon philosophy". I also enjoyed very much Suppes's "Intellectual autobiography, Part I, 1922–1978" [49] and *The Part and The Whole* by Werner Heisenberg [24]. Other autobiographical works, which I have read only in part, are also very interesting, like André Weil's *The Apprenticeship of a Mathematician* [50], Alexander Grothendieck's *Récoltes et Semailles* [19], Lévi-Strauss's *Tristes tropiques* [30], and Carl Jung's *Memories, Dreams, Reflections* [26].

Regarding consciousness, it is very important to write an autobiography especially when it is not a final autobiography at the end of life, because this is a reflection not only about the past, but also about the present and the future. In this sense I think it is worth

[1] Many thanks to Robert Purdy, Patricia McCaslin and Damien Bureau for proofreading this paper.

writing different autobiographies at different stages of our life. I wrote one in 2000, when I was 35 years hold. It is called "From paraconsistent logic to universal logic" (2000). It was a couple of years after my two PhDs, one in mathematical logic at the University of Paris 7 in 1995 and one in philosophical logic at the University of São Paulo in 1996. At this stage I was starting to develop a full research program but I did not have a permanent position – I was a visiting scholar at Stanford University. I had no precise idea about how I could or would be able to do that. However, shortly after that I got a position in Switzerland for a couple of years and was able to implement my research activities. The above autobiographical paper is a 30-page paper about the period 1990–2000, in which I explain in detail how I started to develop the idea of universal logic. It is complemented by a paper I wrote this year: "The relativity and universality of logic" (2014). I will not repeat here the details of the development of my research project on universal logic, nor those of my project on the square of opposition that I have described in "The new rising of the square of opposition" (2012).

In the present paper I will give a general vision of my life up to now, with more personal elements, although this is mainly an intellectual autobiography. I have tried to write something of interest, not only for people working in a particular area of research. This paper can be of interest for people working in areas related to the work I have conducted concerned with logic, mathematics, philosophy, computer science and semiotics. Moreover, I hope it can also be of interest for people working in any area of research.

2 Youth and Adolescence (1965–1983)

2.1 Birth and Family

I was born on January 15, 1965 in the city of Orléans, France, the last of a family of five children. My parents had previously given birth to four girls: Hélène (1954), Elizabeth (1956), Pascale (1958) and Françoise (1963). Orléans is a quite famous town, about 130 km south-west of Paris, associated in particular with Joan of Arc (nicknamed *The Maid of Orléans*). I was born in a hospital just near her famous statue. Orléans is connected to the French royal family d'Orléans, itself linked to the American city *New Orleans*, originally *La Nouvelle Orléans*. I only lived for 2 years in the region of Orléans. I am not of a definite origin. My father is from the west of France, my mother was born in Casablanca, Morocco, and we have never settled down anywhere; there is no place to which I belong.

My father, Jacques Béziau, was born in 1929 in La Guichère, a small village in the west of France, close to the Atlantic Ocean, near the boundary between Brittany and Vendée. The family name Béziau is from this region. There is a place called *La Bézilière* close to my father's birthplace. It is not a common name in France.[2]

[2] Presently there is officially no accent in my proper family name – my name is Jean-Yves Beziau – the accent having been lost through some Kafkaesque administrative process that I will not explain here. However, I am happy with that, I think accents are a useless complication. One of the advantages of the English language is that there are no accents or other parasites surrounding its letters.

This is one of the most Catholic regions of France. During the French revolution, the people there supported the King of France. The counter-revolutionary Catholic army of Vendée was directed by François Athanase Charette de la Contrie, who had at some point restructured his army in *La Béziliѐre*. Charette was shot and is considered a hero in this region. He was admired by Napoleon, who, when taking power over the revolution, stopped the anti-Catholic frenzy, and for this reason is cherished in the Vendée area. My father's aunt, Adèle, entered the Catholic Church as a nun. My father's brother became a priest: he was one of the worker-priests – a missionary initiative by the French Catholic Church. However my father himself, and my grandfather, were rather critical of this church.

My grandfather, Vincent, had a strong personality. Being a tough guy he was sent for 3 years of military service during the 1920s in the Middle East, in the region that is now divided into the countries of Lebanon, Israel and Syria. Back home he married a girl who lived in Paris but was from this region and returned there during the holidays. Her name was Alice. She was raised on Avenue Foch, near Champs-Elysées and her father was a policeman. He died at an early age and Alice had to earn money when she was still a teenager – working with fashion at Guerlain. She met Vincent, who was a member of one of the wealthiest families of her region of origin, married him and never lived in Paris again, although she liked this city very much and went there from time to time. Alice's brother lived in Paris with his wife Madeleine; he was a barber and had a hairdressing salon in rue Saint-Denis. Later in the 1980s Madeleine helped me to find a flat to rent nearby, in rue Tiquetonne.

At the time I lived here, at the end of the 1980s, this area was quite decadent, a mix of prostitution, clandestine workers in the Sentier (brands like Naf Naf and Kookaï were appearing), suburban gangs emerging from the *Forum des Halles*. Nearby was Beaubourg with a lot of tourists and street animators, the très chic Café Coste, the gay and Jewish quarters in the Marais, the Gothic church Saint-Eustache, the Bibliothèque Nationale (National Library), Passage Choiseul where Louis-Ferdinand Céline grew up, Port Royal, Paris Bourse, the Louvre, ... The heart of Paris (the 1st, 2nd and 3rd Arrondissments) is a part of my life.

My grandfather lost most of his money during the Second World War through bad transactions. My father wanted to be an airplane pilot, but due to the lack of money he was placed in a marine mechanics school in the nearby city of Nantes. Nantes has been a very important town in the history of France for marine activities (including the triangular trade) and this was a very good school. After the end of school my father was able to work on boats at an early age. First he sailed on oil boats going up to Kuwait, then on merchandise boats going to North-Africa. The boats stopped in various harbors, and my father had the opportunity to visit many different places, for example Beirut and so on. He sometimes met his brother, who as a priest went to the harbors to help mariners to not spend all their money on drinks and prostitutes in a few hours.

My father became fascinated by the desert and undertook the project to cross the Sahara by motorbike. A project that he realized. He was the first man to cross this huge desert by motorbike (a BMW). Form Morocco he went down to *Afrique noire* (Black Africa) and came back to Morocco by crossing the desert one more time. The German company BMW rewarded him with a brand new engine and invited him to visit their factory in Munich. My father and his brother liked motorbikes very much, a passion they inherited from their father. Unfortunately, his brother died in a motorbike accident when he was 23 years old. This was an emotional shock to my father because they were very close. His brother's name was Yves. This is why, later, my parents gave me the name Jean-Yves.

My father met my mother in Morocco, the place he chose as a basis for his Sahara expedition. My mother was born there and she and my father lived in this country for a couple of years. My elder sister was born in Morocco, before my parents went back to France. These were the last years before the independence of Morocco. As my father later described it to me, this was an amazing place, a mixture of Arabs, Frenchmen, Jews, Berbers, Italians, Spanish and also Americans. Americans went there during the Second World War, and my father at some point worked in an American military base. The famous movie *Casablanca* is about this epoch.

My mother was born Louise Desmeules in 1929 in Casablanca, Morocco. Her father, William Desmeules, was from a small town called Ropraz near Lausanne, Switzerland. William's mother had become pregnant at an early age, and William never met his father. Later his mother married another man with whom she had several children. Due to the English name she gave her first child, it is suspected that William's father was of English origin. When still young, William left Switzerland for a life of adventure. He settled in Marocco where he worked at the post office and later had a grocery shop. He married Marcelle Heitz, born in Oran, Algeria, whose parents were from Alsace. They were typical *Pieds-noirs* settlers, with a huge farm and a big family – my grandmother had about ten sisters and brothers. At some point, her family had moved from Algeria to Morocco, where she and William met. Antoine de Saint-Exupéry made an emergency landing on Marcelle's parents' farm, near Casablanca. William ran to meet the aviator, which is how he met Marcelle. Saint-Exupéry gave the family his parachute as a souvenir, and later this parachute was used to make a white dress for my mother for her communion.

One of the important connections between William and Marcelle was the Protestant religion and culture. They had three daughters: Louise, Marguerite and Germaine, my mother Louise being the eldest. As a child my mother was sent to Switzerland by boat and train through a Swiss Protestant organisation, crossing Europe during the Second World War. At the end of colonisation, my grandparents went back to Switzerland and lived in Lausanne, where I spent some time during the holidays when I was a child and of which I have strong memories. Their home was full of souvenirs from North-Africa and there were books of all kinds everywhere. At the time when my mother met my father she was working in a Montessori school, and the plan was for her to go to Paris to study pedagogy with Jean Piaget, but she did not go because they had their first child. Her sister Marguerite went instead and was a student of Piaget. Later Marguerite dedicated her life in France to helping people from Africa learn French so that they might integrate better into society. She married Jean Dybowski, a descendant of the famous Jean Dybowski, a French colonizer in Africa of Polish origin. On this side of my family we also have a Polish connection through cousins of my mother, who emigrated to Argentina and Australia and all married Polish descendants. The younger sister of my mother, Germaine, married a Swiss engineer, and they went to live, like her parents, in the region of Lausanne. Part of the family of her husband are Swiss immigrants in Brazil.

My mother was interested in education all her life; for some years she was very active in the parent–teacher association Cornec. Later in life she worked in Geneva in a private retirement home and was responsible for running its cultural activities and helping its residents to prepare to face death. She also helped a friend of hers who had the Atelier Arno Stern of therapeutic painting in Geneva. My mother always liked flowers and plants very much and had a strong interest in botany and she frequently went to the botanic garden in Geneva. She did not like travelling very much, and after leaving Morocco she never went back there and spent most of her life in the Genevois. At some point I did the trekking promenade around the four-forest cantons lake with her, which was designated to commemorate the 700 years of the Swiss confederation in 1991. My mother liked to say that she had attended all the three Swiss national expositions that happened during her lifetime: in Zurich in 1939, in Lausanne in 1964 (when she was pregnant with me), and in Neuchâtel in 2002, when I had just returned to live there. My mother never told me what was wrong or right to do, she had the idea that children have to try and discover things on their own.

My father would have liked me to become an airplane pilot but never forced me to go in that direction. He liked the sky and astronomy – observing the stars with telescopes – and also the sea very much; at one point he bought a sailboat and we used to go sailing on weekends. He did not like TV and we never had a TV at home, but my parents liked listening to the radio and music and going to the cinema to see all kinds of movies, and also attending sessions of *Connaissance du Monde* (Knowledge of the World) – a series of documentary movies about different parts of the world followed by a lecture and a discussion with the film director. They took me along with them and this opened me to the world.

2.2 Early Years – Forest and Mountain

I was born in the city of Orléans at the time when my parents lived in Orléans forest. My father had entered the ONF (Office National des Forêts) – the National Forests Office. In France most forests belong to the state. I do not remember almost anything about this forest. I have seen some photos (Fig. 1), my father liked taking photos, a passion which he inherited from his father. He was quite good at that, with a good knowledge of how to capture light. We lived in wild conditions at the middle of the largest forest in France with very old sequoias. We were in isolation in the middle of the forest in a house of the ONF, which in the beginning was without electricity and running water. Nowadays, the house is an abandoned ruin. My sisters did not go to school, my mother educated them.

When I was 2 years old my parents moved to Cruseilles, a small town in France 25 km South of Geneva, 20 km North of Annecy. It is from this place that I have my first memories. Because he worked in the administration of the forest, my father had the possibility to change location every 3 years.[3] He asked to come to this region because he likes the mountains and it was close to my mother's family in Lausanne, Switzerland. Cruseilles

[3] At the end of his career he worked in the French Caraibe, responsible for the natural park surrounding the volcano *La Grande Soufrière* on Guadeloupe island. I visited him there and had the opportunity to dive in Jacques Cousteau's underwater reserve–one of the most spectacular diving sites in the world.

Fig. 1 With my sisters – in the middle of Orleans forest

is located in the Genevois, close to the mountain Salève, which dominates Geneva. The forest of this mountain was administrated by my father and we used to go there. Cruseilles is also in the Savoie area (the name is related to pines, which are numerous in this region), formerly part of the Kingdom of Sardinia. This area became a part of France only in the second part of the 19th century.

In this region there is Mont Blanc, the highest mountain in Europe (4800 m), which lies between France, Switzerland and Italy. Mont Blanc was explored in particular by Horace-Bénédict de Saussure (1740–1799), the founder of alpinism, ancestor of the linguist Ferdinand de Saussure. When I organized the First World Congress and School on Universal Logic in Montreux in 2005, by lake Geneva, Kripke told me that he was happy to see Mont Blanc from his hotel room at the time he was writing a paper for a book commemorating 100 years of Bertrand Russell's paper "On denoting". References to this mountain can be found in the correspondence between Russell and Frege. Russell wrote to Frege on December 12, 1904: "I believe that in spite of all its snowfields Mont Blanc itself is a component part of what is actually asserted in 'Mont Blanc is more than 4000 metres high'." However, this was not Frege's idea, who replied to Russell on December 13, 1904: "Mont Blanc with its snowfields is not itself a component part of the thought that Mont Blanc is more than 4000 metres high" [16]. When I was in my twenties I did the tour of Mont Blanc with my sister Françoise. It is a spectacular trekking promenade of about 200 km circling the mountain going up (to 3000 m) and down to villages in three countries, attracting many people from all over the world. When we lived in Cruseilles, I started to do mountain trekking and skiing at a young age.

Cruseilles was in a booming region. Winter sports were flourishing and the nearby city of Annecy on the banks of a very nice lake is charming. Local people are good at business, and this is the place of origin of Carrefour, which is today one of the largest retail groups in the world. Since my youth I have known the Carrefour hypermarket in Annecy and also the original shop of Fournier, the founder of the group. In Cruseilles we had a big ONF chalet, the furniture was made from wood produced by my father. Cruseilles is the place I

Fig. 2 In Cruseilles

started school (Fig. 2). My first contact with school was to pass an exam to skip the first year. I remember a personal interview during which I had to answer some questions. When I arrived at school I already knew how to read and write; I had learned at home with my mother. My sisters liked to read tales to me, and I also soon started to read, we had plenty of story books at home. It made a strong impression on me and greatly developed my imagination. I also listened to music. My father and sisters went to attend a show of Los Calchakis in Geneva, who played music from the Andes, and came back with their disc, which I like very much. I also liked other, completely different panpipe music – Romanian music promoted by Gheorghe Zamfir, who became world famous through the Swiss musical explorer Marcel Cellier. In my twenties I attended a show by Zamfir at the Saint-Germain church in Paris, where panpipe was mixed with the organ, an idea of Cellier. At school we listened classical music and we had to recognize who the composer was. I was quite good at this and in particular I liked Russian composers: Rimsky-Korsakov, Moussorgski, Tchaikovsky and Prokoviev. Another memory I have is that we still learned to write with a fountain pen and an inkpot.

My family decided to leave Cruseilles, not because they did not like the region but because my father had some problems with his supervisor. He did not agree with spreading DDT over large areas, a very strong insecticide.[4] Since living in the forest my parents started to be concerned by the protection of nature. They were also in favour of organic food and were attracted by alternative medicine. This was at the end of the 1960s, long before it became fashionable. They formed a strong friendship with the pharmacist of Cruseilles, who was completely different from Homais, the famous character of Flaubert. He was Vietnamese and initiated them into Asiatic culture. My father liked it very much, and at home we incorporated a series of Asiatic types of behaviour. My parents were also

[4] On his wish list for a "mutation" (transfer) my father put Alsace as number one, Corsica at the end. He was transferred to Corsica as a kind of retaliation; it was a place that people were afraid to go to because of the harsh behaviour of Corsicans against outsiders, which sometimes led to murder.

against nuclear energy, and we used to take part in meetings. This was the beginning of the hippie period, there was a very good atmosphere, relaxed and friendly. Although my parents were not of the traditional and conservative type, they were never hippies with long hair, or strange clothes and they did not use drugs. My father at some point smoked cigarettes (Gauloises), which in fact at this time were freely provided by the ONF to their officers, but he stopped through acupuncture treatment. One of my sisters later became an acupuncturist.

France is a very rationalist country where all these things were dismissed at this time and the government promoted many wrong things based on some "scientific" dogmas. This was strongly criticized by Alexandre Grothendieck, in particular in his talk about scientism, "The new universal church" [20], presented in 1972 for Russell's centenary. I was raised during this period, in a culture quite close to these ideas, but my parents were never radical in the way Grothendieck turned out to be. I discovered Grothendieck's essay much later and I think it is very interesting, a good reflection about rationality and science, which points out the paradoxical irrationality of some scientists. It took more than 20 years, and quite a number of toxic deaths, before the French government took action about asbestos at the main scientific campus (Jussieu) within Paris. At the time this campus was constructed, there were already warnings about the danger of asbestos. For many years the authorities negated the possible negative effects on students and professors, but in the end the campus was completely evacuated for some years to remove asbestos. The campaign against asbestos was a long fight, and by some coincidence was led by mathematical logicians working on the campus, in particular my former teacher Michel Parigot – I had already left by this time. I spent only 2 years in this asbestos environment.

2.3 The Age of Reason on the Island of Beauty

I lived for 3 years on Corsica Island, between the ages of 7 to 10 – important years in the development of a child. This is where I entered the age of reason and started to have a real interest in mathematics. Paradoxically, this happened in a mythical location where I was surrounded by a natural not to say supernatural atmosphere. This had a very strong influence on my life, which lasted up to now. Corsica is incredible from a geographical point of view. In French it is known as *Ile de Beauté*, which means *Island of Beauty*. It is a mixture of mountains (the highest, Monte Cinto with permanent snow) and sea, a lot of contrasts in a very small area. Nowadays 3 million tourists visit Corsica, which is an island inhabited by 300 000 habitants, per year, mainly in the summer. Despite tourism, the island is preserved due to the fact that Corsicans resist invasion, in particular they do not let foreigners settle and/or buy houses or construct buildings.

During the first half of our stay in Corsica we lived in an isolated house at the middle of the island in the mountains, near the river Golo (Fig. 3); it was called Onia. It was a wild place and the nature was very impressive, in particular there were many different kinds of mushrooms. I walked a lot in the nature around our house, alone or with our tiger cat, Popolasca, who we had named after the mountains in this area. School was about 3 km away in Ponte Castirla, and I went there by bicycle with my sister Françoise. This was an

Fig. 3 In Corsica – with my sisters Françoise and Helene

elementary school with very few children. For this reason, we were all in the same class and room. What I remember from this time is that for recitation exercises we were able to choose the poems we wanted. I liked La Fontaine and to chose the longest fables to push my memory capacity to the limit.

Corsica was at the same time wild and protected from globalisation, but connected to the whole world, since there are Corsicans in many parts of the world and they always come back to their island. Our teacher at school was a Corsican lady who had spent several years in China. My sisters listened to the music of this time, the early 1970s. Two pieces of music impressed me. One was *Slag solution* by Joe Buffalo's band and the other one was *Ouverture to the sun* by Sunforest, incorporated in the soundtrack of Kubrick's movie *A Clockwork Orange*. My sister Hélène used to paint and I did my first painting under her supervision. I liked to paint Corsican villages in the mountains and boats on the sea.

The island attracted hippies and adventurous people. My father used to take hitchhikers in his car and sometimes invited them to our house. My parents had many friends and we went to their houses. I especially remember two Canadian families from Québec. One lived in the convent of Omessa – we were slept in the cells, and the other one lived in a small village in Cap Corse, which was completely abandoned, with houses with doors open and still full of furniture and things. I have a rare copy of the *Pensées* by Blaise Pascal, which we found in this village. My father was in charge of a big part of the Corsican forest and we travelled all around the island. The atmosphere of the ONF was quite good and there was the tradition of méchoui convivial parties (barbecued whole lamb) imported from North-Africa – at some point the Sultan of Morocco Mohammed V had been exiled to Corsica with his family.

At school we learned about the history and geography of Corsica. The lessons were held in French, but all the local children had Corsican as their mother tongue. Corsica is known because this is the place where Napoleon Bonaparte was born, the first French Emperor, and one of the most famous politicians in the history of humanity. His nephew

Napoleon III was the second and last Emperor of France. The writer Prosper Mérimée was part of Bonaparte's circle, a good friend in particular of the Empress Eugenia, the wife of Napoleon III. He made Corsica famous through his novel, *Colomba*. Mérimée is also the author of the novel *Carmen*, a story which became famous through Bizet's opera that is based on it. Napoleon, although supposedly a real Corsican, is not very popular in Corsica because the Corsicans had to fight for their independence. The island has been invaded and dominated through the years by foreigners. Pasquale Paoli (1725–1807) succeeded in creating a Corsican republic from 1755–1769. It was based on a constitution that is considered the first constitution of the modern world, probably the first constitution in history allowing women's suffrage. Jean-Jacques Rousseau supported Corsicans and was asked by them to work on a constitution – see [42]. The Corsican constitution has inspired that of the United States. In the USA five towns are called Paoli, remembering the Corsican politician. Paoli also created a university in Corte in 1765, the historical capital of Corsica, the town where we lived during the second part of our stay. In May 8–9, 1769 Corsicans lost the battle against French troops at *Ponte novu* and since then Corsica has been a part of France. Napoleon was born August 15, 1769 in Ajaccio, but he was conceived in Corte where his father was attending the university. Napoleon strongly contributed to the incorporation of Corsica within the French Empire that he himself developed.

The University of Corte was closed in 1768 and only reopened in 1981 under the name *Università di Corsica Pasquale Paoli*. I remember that when we lived in Corte, the reopening of this university was a main issue for Corsicans. Many years later in 2010 I had the possibility to organize the 2nd Word Congress on the Square of Opposition at this university with the help of my friend Pierre Simonnet, who was working at the Computer Science Department. We had the participation of Damian Niwiński from Warsaw University, Editor-in-Chief of *Fundamenta Informaticae*, and of the famous mathematician Pierre Cartier, one of the most active members of the second generation of Bourbaki and a close friend of Grothendieck. Cartier likes Corsica very much, in particular because he spent his honeymoon there. During the congress we filmed Cartier talking in detail about his meeting with Gödel in his house in Princeton in the 1950s. It was in Corte that I encountered modern mathematics, not at school but through some booklets by Georges Papy that my mother gave me. Pappy (1920–2011) was a Belgian mathematician, a great promoter of modern mathematics for young children but also for a wider audience. His booklets are full of pictures and it is stated that the age of the readers ranges between 7 to 77. These booklets present many important results about set-theory and infinity, like the one-to-one correspondence between natural numbers and a proper subset, Cantor–Schroeder–Bernstein's theorem, etc.

In Corte I was reading a lot of books – we lived next door to a library. It was then that the comic book *Astérix in Corsica* by Goscinny and Uderzo was released. I read many other comic books by Goscinny, my favorite being the series *Iznogood* (a joint work with Tabary producing the pictures) and I also liked *Le Petit Nicolas*, which he produced with Sempé. René Goscinny is of Polish-Jewish origin and spent his youth in Buenos Aires. He is a very talented and imaginative writer who has had a strong influence in France, rivaled only by the Belgian cartoonist Hergé with his famous *Tintin*, which is also excellent. By contrast to Goscinny, Hergé was pro-Nazi and it has been claimed that Tintin's dog, Milou, was inspired by Hitler's dog. This is certainly not the case of Astérix's dog Idéfix (Dogmatix in

English–Goscinny was very good at word play, most of them simply untranslatable), who is known to cry when a tree is cut. I also walked around the city by the rivers Restonica and Tavignano. The atmosphere at elementary school was very good, the students were smart and friendly, the girls beautiful. I enjoyed the exercises we had to do – inventing stories inspired by the music we listened to while writing. Here I was always number one at school without any special effort.

2.4 Thorens Glières and Annecy

We left Corsica in 1975 and went back to Haute-Savoie, to a small village called Thorens Glières. This is the place where Saint François de Sales (1567–1622) was born. He was the bishop of Geneva and is the author of *Introduction to the Devout Life*. We lived in an ONF house a few metres from the chapel and the castle related to his family. In this castle Napoleon III met the count of Cavour, minister of the Kingdom of Sardinia, to discuss the affiliation of the Savoie area within France, which was formalized in the Treaty of Turin in 1860, according to which the County of Nice also became part of France. The region was full of forests and mountains. I did a lot of hiking, skiing and bicycling. In the mountains there was a national park that my father took care of, which was accessible only by climbing and there were some mountain refuges where people from the ONF and their family and friends were allowed to stay for a couple of days. My father participated in the re-introduction of several extinct species of animals, and this park was also known for the western capercaillie – specialists went there to study its behaviour. I took part in observations of these birds.

Because I had come from Corsica, for several years my nickname was Napoleon, or Napo for short, or Bonaparte. I attended a middle school in the nearby village of Groisy for 4 years. I did not like the atmosphere much at that school, and I did not get very good marks, except in mathematics, and in "la composition française" – exercises in imagining things and then writing essays about them. The mathematics classes were strongly in the spirit of modern mathematics. We were taught abstract algebra, and the exams were much more directed toward proof than calculation. I was able to solve the most difficult questions. At this time mathematics saved me, because in France it was considered the number one discipline, the absolute proof of intelligence. I also got good results on IQ tests. At that time it was a popular practice to administer IQ tests in schools, and your IQ was taken into account when determining what track your further education should follow.

I started to read a lot of science fiction books, classics by Isaac Asimov, van Vogt (who I later learned was a follower of Korzybski, a promoter of non-Aristotelian logic), Philip José Farmer and Philip K. Dick. They gave me a lot of inspiration. The writings of Asimov about robots have a quite interesting logical aspect, and Asimov is thought to have anticipated the idea of Wikipedia.[5] My favorite author turned out to be Fredric

[5] Although Asimov may have anticipated Wikipedia, no science fiction writers anticipated personal computers and the web, two fundamental components of our present world. Reality is always more surprising than fiction.

Brown; I liked his style, sense of humour, ability for ingenious plot, incredible imagination and psychological subtlety. Before he started writing science fiction, Brown wrote crime stories – *The Far Cry* is my favorite. He was also much influenced by Lewis Carroll, whose works I also read at this time.

Our neighbor was a colleague of my father. He was about 30 years old and he liked to play chess. I started to play chess and after a couple of weeks I was able to win against him. I liked this game very much. For me it was connected to reasoning. Before I had a maths exam, I thought it was good to play a game of chess to fortify my mind. At this time I listened to all kinds of music. I liked singers like Eddie Cochran, Leonard Cohen and Peter Gabriel, and bands like the Beatles, Pink Floyd, Deep Purple, King Crimson and Genesis. Then disco music appeared with Cerrone and Boney M, and punk music with the Sex Pistols, Clash, Ramones, etc. I started to play music, composing different pieces. Also through my father, who has always been interested in cars, a passion that he inherited from his father, I began to take an interest in car racing, in particular Formula 1. My father had worked for Ferrari in Morocco and had met Ascari. My favorite driver was Niki Lauda.

After middle school I went to the nearest high school, which was in Annecy, about 20 km from Thorens Glières: the Lycée Gabriel Fauré, bearing the name of a famous French musician. I had to go by bus, and it was quite slow because the bus stopped in many villages along the way. During the winter there was lots of snow, so we went very slowly. I would wake up at 6 o'clock in the morning and come back home at 7 o'clock in the evening. All this had a negative effect on my schooling; I had to repeat my first year of high school, but this was also related to a general teenage feeling of dissatisfaction. I went to Paris for the first time and stayed with my sister Hélène who lived there. Paris made a very strong impression on me. I wanted to be independent and even thought of quitting school.

At some point, in Annecy there was a competitive exam to be a postman at the post office, and I decided to be a candidate. I had a positive idea of the post office, since this was related with my grandfather William and Saint-Exupéry, who had worked for the famous postal line from Europe to Chile – I had a collection of stamps. Also, I knew that Gaston Bachelard had been a postmaster, so it seemed to me that this kind of work was not incompatible with intellectual life. I had felt empathy with Bachelard, also influenced by Korzybski's non-Aristotelian logic. Later, when I was a student at the Sorbonne, Bachelard's daughter Suzanne was one of my professors. She held a class on the *Dioptric* of Descartes. She criticized Descartes, although she liked him, laughing at him in a funny way. I did not become a postman because I failed the post office exam – strangely enough because of mathematics. The general exam was not very difficult. We had to know a lot about geography, and I was good at that, and we had to write a composition française, which was also easy for me. Then there was a mathematics test, which was a rule-of-three problem. However, due to my modern mathematical training, I had had very little exposure to such questions, and I did not succeed in solving it. Modern maths again saved my life.

I, therefore, had to continue with high school and prepare for my "baccalauréat". The situation improved because I moved to Annecy and I found a nice girlfriend. At first I shared a flat with my sister Françoise and her boyfriend, but I soon lived on my own. This was in 1981; I was 16 years old, lived alone, and I never went back to living with my parents. For many years, up to 2002, I was without any kind of longer-term “home base”, changing addresses nearly every year and sometimes living in hotels. Later Pat Suppes would tell me I was like Kreisel, a good friend of his, but opposite in personality to him, since Kreisel moved around a lot while Suppes had been living at the same address on the Stanford campus since the 1950s.

Kreisel has been pivotal for the development of logic in France. In the 1960s under his influence a school of mathematical logic started up in France. Jean-Louis Krivine and Jean-Yves Girard were taught by him. Krivine and Girard were both teachers of mine, so in some way I have benefited from Kreisel spreading logic in the world. However, I must say that I have never been much attracted by Kreisel’s style, personality, or directions of work in logic – a “Tarskian dissident”, as depicted by Sol Feferman (see [14]). I have always felt close to Alfred Tarski. Yet Kreisel is definitely an interesting character. He was a good friend of Raymond Queneau, one of my favorite writers, and had many interests in life. Even if there is an apparent similarity between myself and Kreisel from the point of view of nomadic versus sedentary person, there is a strong difference in the sense that Kreisel was a rich man. In my case, nomadism was related most of the time to living in a precarious situation. This was not at all easy, but it was a conscious choice. I had a preference for a life of adventure; I wanted to avoid a boring everyday repetitive life. For me, reasoning and thinking are connected to this nomadism in two opposite ways: on the one hand, traveling is good for developing ideas, and on the other hand, reason makes a good defence against fear and despair, not so much in the sense that analyzing emotions by reasoning dissolves the problems, but rather that reasoning is good for the health.

Being short of cash I would take “summer jobs” during the holidays for pocket money during the school year. I had three very different kinds of jobs: one to do with apple trees, one with a security agency, and one in a cheese factory. These jobs gave me a clear idea of what I did NOT want to do, and motivated me to go on studying. In the cheese factory I worked on a production line 7 h/day, from 5:00 a.m. until noon, repetitively placing little pieces of cheese on a rail that carried them into a machine that automatically wrapped them in plastic.[6] Until doing that, I had had no real idea that there were people who do this their whole life. I had seen Chaplin’s movie *Modern Times*. Watching a movie is quite different from living through the experience. I watched lots of movies because my parents had registered me in Annecy’s cinema club, one of the oldest in France – and in Annecy there is the yearly international festival of animation films, which I also attended. I saw many great movies by Pasolini, Godard, Eisenstein, etc.

At Gabriel Fauré high school I had two friends, Alain and Jacques, with whom I talked about most everything and even started a music band. They were the two best students in the school. Alain was the very best. For him everything was easy, he always got top marks in all fields, i.e. 20/20 in the French system. In the last year of high school he had a philosophy teacher who was a Lacanian and who had a reputation for being very harsh; nobody understood what he was talking about and it was difficult to get a mark higher than 5/20. Alain regularly got about 15/20. For him this was the most interesting and challenging subject, exactly because it was difficult to understand. After high school

[6] I have seen the same little piece of cheese wrapped in ornamental plastic bags with different colourful brands, so that the consumer has the illusion of choosing among diversity.

he decided to go on in this direction and he did his PhD with Jacques-Alain Miller in Paris, the main Lacanian after Lacan's death. Lacan was strongly criticized by Sokal and Bricmont in their book *Impostures intellectuelles* – translated into English as *Fashionable Nonsense* – (see [48]).

Most of their book is rather *argumentum ad personam*, with no serious argumentation and/or philosophical basis. In 1998 I was invited with other colleagues to take part in a debate with Sokal and Bricmont at the University of São Paulo. In my talk I criticized their "argument" about Kristeva. At some point in their book they say that she speaks about the continuum and $\aleph_1$ in the context of poetry, and that this makes no sense because in literature there is only a finite number of characters or strings of characters. I said two things about that. First that it is possible to write a poem involving all the transfinite cardinals: $\aleph_0$ is a beautiful cardinal, etc. This is a poem close to the Oulipo spirit, mixing mathematics and poetry. On the other hand, I told them that in Hilbertian proof theory, a proof is considered as a finite string of characters but that nevertheless Gentzen proved the consistency of arithmetic using transfinite recursion up to ϵ_0. Bricmont answered me something like: "Maybe, but nevertheless Kristeva wrote a lot of nonsense". Sokal and Bricmont defend a return to the *Siècles des Lumières* – I am not sure that this is very interesting. Their book gives a distorted vision of French intellectual life in the decade after the Second World War, which is one of the most intellectually interesting periods. There was in particular a mixing of mathematics with other fields in a highly creative way. This was the time of structuralism. André Weil collaborated with Claude Lévi-Strauss. Le Lionnais, co-founder of Oulipo, edited a beautiful book, *Les grand courants de la pensée mathématique* [29], including in particular Bourbaki's masterpiece *The Architecture of Mathematics* but also a piece by le Corbusier and much more.

My other friend, Jacques, was principally interested in mathematics, and after he passed his "baccalauréat" he entered "Mathématiques supérieures" at the Lycée Berthollet, also in Annecy. I entered this school too, but in "Lettres supérieures". In France these are called "classes préparatoires", nicknamed "prépa", and they are taught in high schools even though you cannot enter them until after you have passed your "baccalauréat", which is the end exam of high school. In "prépa", students prepare for the entrance exam to the *Ecole Normale Supérieure*, which is considered to be the best post-secondary school in France – one dare not call it a mere "university" – and one of the best in the world.

At the time when I started high school it was standard practice for students to be assigned different "tracks" according to aptitude. Since I was good at mathematics I was placed on a scientific track. I did not like the atmosphere, particularly the fact that 80% of my class were boys and only 20% were girls, but also because we had to take physics and chemistry, which I found boring. Then I transferred to a section combining literature with mathematics, and that was fun. We had a very good literature teacher, and I started to write all kinds of things. In my "baccalauréat" there was an oral exam on literature where we had to comment on a text chosen by the examiner from a list that we had prepared. The list of texts was supposed to follow some rules, but it was allowed to include personal writings. Most students did not do this, but I put quite a few of my own compositions on the list, hoping to be asked to comment on something I had written. However, the examiner, a woman, asked me to comment on a poem by Arthur Rimbaud. I refused to do so, explaining that for me the whole business was absurd. The woman told me that I was wrong not to be interested in Rimbaud, that I would probably like his poetry, that I was a bit like him. In the end, I got quite a low mark for literature, and only passed my "baccalauréat" thanks to a high mark in mathematics. I always enjoyed mathematics at high school, real mathematics, based on reasoning. At some point when I was about 14 years old I had the feeling that reasoning was independent of the contents that were being reasoned about, and after that I did not have much interest in the specific contents. I was more interested in difficult types of reasoning. Rubik's cube was very popular at that time among students, but it never interested me; it was not the kind of difficulty that I was interested in.

In the last year of high school I had about 9 h of philosophy a week. The program included a wide range of topics. We had a very dedicated teacher, a young woman. I was very enthusiastic and got the highest marks. She recommended me to go to "Lettres supérieures" in Annecy and then to continue in Paris, where, she told me, I would have the best teachers. This is what I did. After 1 year of "Lettres supérieures" I applied to go to "Première supérieure" to the top ten best "Lycées" in Paris. For this second year we had to choose a specialty. For nine of them I applied with history as a speciality and only for the best, Henri IV, did I apply with philosophy as a speciality. I was selected to all of them and, therefore, went to Henri IV for philosophy. Before my class of "Lettres supérieures" I had had no particular interest for history, but during that year we studied Greek antiquity, Roman antiquity and 19th century French history. I liked it a lot and this was the discipline where I had the best marks. I was also good in philosophy, but traditionally in France, it is more difficult to get a good mark in philosophy than in any other area. In history I had average marks of 17/20, in philosophy 14/20. I preferred philosophy to history, but since I wanted to be sure to be selected in a good "Lycée" in Paris, that is the reason why I applied to all the "Lycées" with history, excepted one, the best. That was a kind of wager. A strange wager, quite different from that of Pascal's.

During that year in philosophy we studied Blaise Pascal and I liked it. Our professor was an old Christian lady, and we studied in detail the *Book of Job*. She also liked Plato and, in particular, we studied the *Symposium*. During this year I learned a lot about Greek culture, writing in particular a 40-page essay analyzing in detail religion in the Iliad and Odyssey. During the Easter holidays a trip was organized for all the students to the region of Tuscany in Italy.We visited all the museums in Florence and Siena and many historical things spread over small villages. It was a fantastic trip in a region that is a wonderful mix of beautiful nature and culture. This was 2 weeks of relaxation, much welcome because in "prépa" we had to work hard all the time, writing lots of dissertations in the section that I was in.

3 Student in Paris (1984–1991)

3.1 Lycée Henri IV – Focusing on Philosophy at the Top of Paris

In 1984 I moved to Paris and attended "Première Supérieure" at Lycée Henri IV. This was originally called Lycée Napoleon and is the highest rated "lycée" in France together with the nearby Lycée Louis Le Grand. There is a legendary rivalry between the two. Louis le Grand is better for science and Henri IV is better for humanities. Many famous people are alumni of these two "lycées". One reason for the objective superiority of Henri IV is both geographical and physical. Henri IV is at the top of the Montagne Saint-Geneviève and has a tower, the Clovis tower, after the name of Clovis the first king of the Franks. This tower was the highest building in Paris for many centuries. I was "interne" (boarder) in the "lycée" and we used use to climb to the top of the tower (which was not open to the public) at night for the nice view of Paris. We also explored underground passages and the many buildings of the "lycée" (several movies have been shot there, such as *La Boume*, or in English: *The Party or Ready for Love*). We were allowed to go outside at night and come back any time, so for me Henri IV was not at all a prison unlike the nearby Lycée Saint-Louis (a third famous "lycée") was for Jean Van Heinjenoort. As reported by Anita Feferman: "We were walking on Boulevard Saint-Michel, towards the Seine, when he pointed to a large building and said directly to me, "That's the prison where I was locked up for two years." "Prison?" I said, "What do you mean, prison?" "Well, it is a school, I was an interne, which means I wasn't allowed to go out except on Sundays." [13][7] Contrary to a prison, Henri IV was a nice and quiet refuge including a beautiful flowered cloister. I was happy to go back there after wandering around discovering Paris. The "lycée" attracted students of many different backgrounds and origins; in particular Henri IV has a tradition of welcoming the children of French expatriates living in the ex-colonies of France. I had good friends and the atmosphere was quite extraordinary.

As far as the lessons were concerned I was not very motivated because it was a lot of cramming for entrance to the *Ecole Normale Supérieure*, especially since I was preparing for the entrance exam for ENS Saint-Cloud, where there is a specific program changing every 2 years. In history the program was colonisation and in geography the Maghreb (Tunisia, Algeria and Morocco) in which, despite of, or maybe because of, the history of my family, I had little interest. During this year of "Première Supérieure" I was only interested in philosophy. Since philosophy was my area of specialisation, I had two classes of philosophy: one which was common to everybody, and one which was only for those who were specializing in philosophy. The one for everybody was given by Pierre Raymond and the topic of the program was "La morale"; the other was given by Pierre Jacerme and there were two topics, "L'existence" and "Le symbole". I especially liked the topic on the symbol – years later I am still working on that topic, organizing a congress and publishing a book on it – cf. *La Pointure du Symbole*. Independently of the topics these philosophy classes were interesting because of the strong personalities of the two teachers. They were two opposite characters and figures. Raymond was a former student of Althusser and Jacerme a former student of Beaufret, the two main leading figures of philosophy in France in the 1960s and 1970s.

Raymond was a nice guy but his classes were rather boring and not easy to understand because it was not clear what he was aiming at. One thing I remember is that he told

[7] As it is known, Heijenoort left the Lycée Saint-Louis to join Trotsky in Turkey, spent 10 years as his secretary and bodyguard in Mexico, and then went to the USA where he became a good friend of Quine and the first historian of modern logic, promoting the myth of Frege as the founding father – the detailed story is related in Anita Feferman's book [12].

us about Wittgenstein, which was the first time I heard about this philosopher. My first contact with Wittgenstein was not with his *Tractatus* or logic but with ethics. I read his writings about ethics and esthetics and quite liked them, in particular his criticism of the way we use language, how our mind can be directed by language without understanding what we are thinking about. Later I found similar criticisms in Schopenhauer, one of the major influences on Wittgenstein.

Through Jacerme I learned about Heidegger. The first book I read was *What is a Thing?* and I liked it very much, especially the first part where he talks about mathematics explaining the original meaning of the word. To date I think this is one of the best introductions to philosophy. Later I also read *The Principle of Reason*, one of his best books, and various of his essays of the 1940s and 1950s, in particular those collected in the book *Off the Beaten Track*. On the other hand, I never had any interest in *Being and Time*. Later, I read his essay "Plato's doctrine of truth", which led me to write a Master's thesis on Plato's cave.

I presented myself for the ENS exam but without much motivation. On the one hand, I had no interest in disciplines other than philosophy, and on the other hand, even for philosophy, I was not much interested in entering ENS, because at that time it meant following a not very exciting track: preparing "agrégation" (another cramming thing), teaching in high school, etc. A track connected with quite a secluded life. I wanted to have more contact with real life. With research in philosophy in mind, I decided, therefore, to go to university.

3.2 Studies at the Universities Paris 1 and 7 – The Logical Way

After doing 2 years of "prépa" it is possible to get an equivalence certificate to enter the third year of university. The Sorbonne was just nearby. When I speak here of Sorbonne, I mean the building. Historically, the Sorbonne was a university located in one building, "rue de la Sorbonne", but nowadays the Sorbonne is spread over many buildings and it is not only one university, but three universities have the name Sorbonne: *Université Panthéon-Sorbonne* (Paris 1), the *Université Paris-Sorbonne* (Paris 4), and *Université Sorbonne Nouvelle* (Paris 3). After May 1968 the University of Paris was split into many universities. Today there are 14 universities, each with a number and a name: from Paris 1 to Paris 14. I have studied in Paris 1, Paris 3 and Paris 7. From my entry to the university in 1985 up to definitively leaving Paris in 1991, I received seven diplomas:[8]

- Licence de philosophie, 1986
- Licence de logique, 1987
- Maîtrise de philosophie, 1987
- Maîtrise de logique, 1988
- Diplôme d'études approfondies de philosophie, 1988

[8] I have indicated the names of the diplomas in French because there are no straightforward translations. These university diplomas do not exist anymore in the new Bologna system that has been implemented in Europe. A "Licence" was a 3-year degree, a "Maîtrise" a 4-year degree, and a "Diplôme d'études approfondies" (="DEA") a 5-year degree, equivalent to a Master's degree.

- Diplôme d'études approfondies de logique et fondements de l'informatique, 1990
- Licence d'études cinématographiques et audiovisuelles, 1991

I came back to Paris in 1995 for a few days just to defend my eighth French diploma, a PhD in mathematical logic.

I will try here to explain how logic became my main topic of study. Logic most of the time does not appear as the name of a field of study or a discipline. It is very difficult to find a university with a department of logic and/or a diploma of logic. In Paris there were no (and there still aren't any) departments of logic, but some diplomas of logic: "Licence" and "Maîtrise" in the Department of Philosophy at Paris 1, and a Master's degree and PhD in the Department of Mathematics at Paris 7. I obtained all four of these degrees. When I applied for an equivalence certificate to enter the "Licence" of philosophy at Paris 1, I was required to pass some exams in logic because this field was a strong part of the philosophy curriculum in that department and did not exist in "prépa". I had no precise idea of what was beyond the substantive word "logic". In French there is an adjective and a noun, which are the same word: "logique". As with the English adjective "logical", the French adjective is part of natural language. The French noun, however, is not widely used – the layman does not exactly know what it means, and among intellectuals there is a lot of ambiguity surrounding it. I did not know exactly what logic was about but I was attracted by the very word. After asking colleagues about the program, I studied by myself and passed these exams. They were about basic propositional and first-order logic, with some exercises of translation from natural language sentences into formal language.

I then entered the "Licence" of philosophy. There was a bit of everything: I remember that I studied Sartre with Hélène Védrine and Plato with Sarah Kofman. It was also possible to do a "Licence" of logic based on an advanced class of logic centred on the completeness theorem for first-order logic, together with some epistemology and philosophy of science. I completed this second "Licence" while I was doing my "Maîtrise" of philosophy. Such a "Maîtrise" was a compound of two classes of philosophy and a dissertation of 100–150 pages. I wrote a dissertation about the philosophy of modern physics under the direction of Bernard d'Espagnat, a very famous physicist who also taught the philosophy of physics at Paris 1.[9] My dissertation is focused on David Bohm's holomovement (cf. [5]) but with a strong part about the Copenhagen interpretation and on the background Heidegger's analysis of the relation between thing and thought. I was quite impressed by the writings of Werner Heisenberg. In 1987 Bohm was supposed to come to Paris but he was not allowed to because of a visa obligation, and I went in London to have a discussion with him. Besides the dissertation and a class with d'Espagnat, I followed an interesting class held by Sarah Kofman on the comparative study of the Presocratics viewed by Hegel, Nietzsche and Heidegger.

After my "Maîtrise" I first thought of doing a Master's of Philosophy on logic and foundations of mathematics with Jacques Bouveresse. But then I read "Plato's doctrine of truth" by Heidegger and decided to work on Plato's cave. Since I liked the classes of Sarah Kofman I asked her if she would agree to direct my work on this topic, and she gave me a positive reply. Kofman did her PhD on Nietzsche with Deleuze and then met Derrida with whom she became a good friend and was much influenced by. I never really

[9] Later d'Espagnat won the Templeton prize.

read texts by Derrida but I read several of Kofman's books. My favorite one is *Comment s'en Sortir?* (How to find a way out?) [28] which is about the notion of "aporia" in Plato. Plato's cave is one of the most famous texts of philosophy, and I wanted to understand its profound meaning (if any). My initial idea was to examine and classify all the various interpretations and also, since I liked images, the different pictorial representations of the cave. A book by the Canadian Yvon Lafrance was very helpful for this work–in this book he indexed all texts about Books 6 and 7 of Plato's *Republic* with short abstracts of their contents. On the one hand, I was much surprised to discover that there were very few different interpretations of the cave – and also very few pictures – but on the other hand, I discovered some interesting things, in particular about a cave in Crete that inspired Plato.

I decided to write a dissertation in three parts, each being a caricature of three different interpretations. The first is a class given by a high school teacher commenting on Plato's cave, explaining the theory of ideas and epistemology with a touch of morality. The second is a dialogue of Jean Beaufret with "one who is questioning", a parody of the books of Beaufret *Dialogues avec Heidegger*. The third is a mix of new age, neo-Platonism and Hellenic history that emphasises the rituals happening in the cave discovered in Crete, in which Pythagoras and Plato supposedly took part. I defended my Master's dissertation at Sarah Kofman's own house – she had some difficulties of locomotion. She liked it and gave me a good mark. Later on, when I had already left Paris, she committed suicide and I dedicated a paper I wrote about Schopenhauer's analysis of suicide to her.

I was then thinking of doing a PhD of philosophy about Schopenhauer. I started to have a strong interest in this philosopher after reading various books by Clément Rosset [10] about him. However, Rosset was in Nice and I did not succeed in getting in touch with him. Finally I decided to matriculate for a PhD with Philippe de Rouilhan to do something about the philosophy of logic. I knew de Rouilhan from the "Maîtrise" of logic I was doing in parallel with my "Maîtrise" of philosophy. He was a kind of provocative dandy from an aristocratic family – one of his ancestors had been secretary to the King of France. His style contrasted that of traditional philosophers, whether continental philosophers or analytic philosophers like Bouveresse, his former PhD advisor. Initially a mathematical logician, he shifted to the philosophy of logic, or rather the history of logic, and became a specialist of Frege and Russell. We studied Frege's "Über Sinn und Bedeutung" and various paradoxes of this time with him. De Rouilhan was a strong admirer of Heijenoort, whom he had the opportunity to meet. In this "Maîtrise" of logic Bouveresse also gave a class about philosophy of logic, but the core of this degree was modern logic. We had four classes for each of the main topics: model theory by Joël Sakarovich, set theory by Michel Eytan, recursion theory by Jean Mosconi and Susana Berestovoy, and proof theory by Joël Combase. There was also a class of computation by Susana Berestovoy – we learned LISP, the famous computer programming language invented by John McCarthy in 1958.

This was a fine and serious background for logic studies given by a group of eclectic people (Sakarovich was also interested in architecture – see his book [43]). Eytan liked category theory and he used the book *Axiomatic Theory of Sets and Classes* by Murray Eisenberg [11]. I wrote my first philosophical essay about logic for his class on the topic he gave us about the truth of assertions in set theory. Susana Berestovoy was from Argentina, she was very dedicated and it was possible to discuss with her for hours after

[10] For me Rosset is the best French philosopher of the second part of the 20th century.

class. Combase had studied with Kreisel and then Feferman at Stanford, he was intelligent and friendly but had some psychic troubles.

After the “Maîtrise” of logic at the Department of Philosophy of Paris it was possible, but not easy, to be admitted to the Department of Mathematics of Paris 7, to go on to study logic, doing a Master’s and a PhD in logic. The University of Paris 7 is the most interdisciplinary university in Paris (the others are generally focused on some specific topics), which is why it was named “Denis Diderot”, after the name of the famous encyclopedist. However, it is especially known for its Department of Mathematics, one of the best in France. At that time it was directed by a former student of Grothendieck, Jean-Louis Verdier, but unfortunately he died in a crash accident shortly after I arrived, in 1989. The department was strong in algebra, category theory and logic. At this time the logic group of this department gathered the main French logicians, Jean-Louis Krivine, Jean-Yves Girard, Daniel Lascar, etc. A few years later it was split in various groups, and Girard developed his own group in Marseille.

In this department there were the degrees Master and PhD in Mathematical Logic, which had just been renamed Master and PhD in Logic and “Fondements de l’Informatique”, corresponding to some fashion and the shift of interest of some members of the group from model and set theories to lambda calculus and foundations of programming. These people focused on the relation between proofs and programs, the Curry–Howard correspondence being a kind of motto that they were putting forward at each talk they gave (a litany which is still going on). These logicians had strong personalities; they were different from each other and did not necessarily have good relations with each other. We had a very dedicated teacher, René Cori. He and Lascar later published a two-volume book corresponding to what they were teaching to us in the first semester (see [6]). As the reader can check, it is a lot of material and a very intensive class about the basic elements of modern logic. The academic year was divided into two semesters, and in the second semester we could choose some optional classes. I did one with Girard on proof theory and also one with Daniel Andler on non-classical logics.

Then during the summer we had to write a monograph corresponding to research work showing (or not) our capacity to do research and go on with a PhD. Very few people intended to go on to do a PhD and we were not especially encouraged to do so. I wanted to do a PhD but it was not clear on which subject I would work. This was the case of other students who generally were incorporated to one of the groups, in particular those working on linear logic with Girard. I did not follow this linear road. An interesting subject naturally appeared to me.

In Paris I had a nice girlfriend whose father was a psychoanalyst who had been analysed by Lacan. At some point we were relaxing in his house in the countryside on the banks of the Loire river. In these circumstances I read an interview of Newton da Costa in the Lacanian magazine *L’âne* (The donkey). I had never heard about this Brazilian logician and his work before. The title of the magazine article was something like, “Paraconsistent logic: a logic for the unconscious?”. I liked the personality of da Costa, the way he replied to the interviewer’s questions and the way he maintained that he did not consider himself only as a technician of logic. At the end of the interview he quoted Pierre Curie saying: “Faire de la vie un rêve, et d’un rêve une réalité” (To turn life into a dream, and a dream into reality). Furthermore, I was very much interested in the topic: paraconsistent logic, a logic in which the principle of non-contradiction does not hold. I wanted to understand how this could work, mathematically speaking. Moreover, it seemed to me very interesting from a philosophical point of view, since the principle of non-contradiction is often presented as the most basic law of thought.

Back in Paris I went to the library to look for da Costa's papers. I had no difficulty finding them because his initial papers have been published in French in the *Comptes Rendus de l'Académie des Sciences de Paris*. How this happened is explained by Marcel Guillaume in [21], a Bourbachic mathematician friend of da Costa. Though these papers had been published in France, nobody knew about them in Paris. I started to work on that and asked Daniel Andler if he would be interested supervising my Master's on this subject. He had given us a class on non-classical logics including modal logic and non-monotonic logic but did not know paraconsistent logic, although he knew Guillaume. I explained to him what paraconsistent logic is, and he gave a positive reply. Andler had done his PhD in Berkeley in Tarski's group at the beginning of the 1970s – see [1]. Later, his interest turned to artificial intelligence, cognitive science and the philosophy of science. At the time I met him he was researcher at CREA (Research Centre for Applied Epistemology), which was attached to the polytechnical school, and was working on the launching of the cognition department at ENS-Ulm, which he would succeed to create. He is now Professor of Philosophy of Science at Paris 4, *Université Paris-Sorbonne*.

I enjoyed working on that Master's thesis very much. I read in detail the papers by da Costa, trying to understand everything. I reformulated the main concepts, developed some new techniques, in particular using sequent calculus, which I had just learned, and presented some new proofs. I very much liked the work of Gentzen and the cut-elimination theorem that I had studied in detail when reading the Gentzen's original paper and following the class of Girard, who told us that this was one of the most fundamental theorems in modern logic. I succeeded in developing a proof of the cut-elimination theorem for a sequent calculus I constructed for the paraconsistent logic C1 of Newton da Costa and variations of it. A former student of Paul Bernays, Raggio, had presented an incomplete work in that direction 20 years before. By doing this work on paraconsistent logic I had a better understanding of how classical logic works. I tried to see what was similar to both of these logics and what was different–both from the semantical viewpoint and the proof-theoretical viewpoint. This was the first step in the direction on my work on universal logic that I later developed in my PhD, as I explained in details in my 2001 paper "From paraconsistent logic to universal logic".

While doing my Master's of Logic at Paris 7, I was going to the seminar of de Rouilhan at IHPST,[11] where interesting logicians gave talks. I remember in particular a very attractive talk by George Boolos (1940–1996). Sol Feferman also presented a talk, about Gödel. The atmosphere was relaxed; at this time, after the talks, there were cocktail parties with plenty of champagne. I think nowadays they serve Coca-Cola.

At Paris 7 I had a friend who was studying logic with me and who liked category theory very much and the philosophy of Alain Badiou. At this time Badiou's book *L'être et l'événement* had just been released – a book in which Badiou deems the invention of

[11] IHPST stands for Institut d'Histoire et Philosophie des Sciences et Techniques and is an institute that is affiliated with Paris 1 and CNRS, located in the Odeon area of Paris, close to Saint-Germain des Prés. It was originally created in 1932 and directed by Gaston Bachelard from 1940 to 1955.

the notion of forcing, by Paul Cohen, to have been the main event for philosophy in the 20th century. My friend started to organize some meetings with a small group of people to discuss the book gathering Badiou and some mathematicians, mainly people working in category theory, in particular René Guitart. This was useful to reinforce the link between mathematics and philosophy. After that Guitart was invited to give a seminar at the *Collège International de Philosophie*, which I attended and which was very interesting. Guitart had good relations with Jacques Riguet (1921–2013), a mathematician and friend of Lacan, who taught Lacan graph theory and did a nice PhD on relational structures.

While I was studying logic I also studied cinema. I had always been interested in cinema from an early age and thought of becoming a film director. France is a very important country for cinema; one can argue that this is where it started with the projection of movies of the Lumières brothers at *Salon indien du Grand Café*, place de l'Opéra in 1895. This is also the first country in the world where cinema was introduced at the university. I matriculated for "Licence" of cinema, which was a joint degree of Universities Paris 1 and Paris 3. I attended very interesting classes on the history of cinema, critical analysis of movies, script writing, soundtrack, etc. I saw many, many movies, at one stage an average of three movies a day. The university film library in Paris is the biggest in the world with lots for rare and original movies, which were projected at those universities, and there was also the general film library of Paris and many movie theaters in the Latin Quarter. I also attended a retrospective of Swiss movies at the Swiss Cultural Centre in the Marais, which was close to my house. I thought of entering the newly created school of cinema, FEMIS, directed by Jean-Claude Carrière, who had written many scripts of Luis Buñuel's movies, in particular *That Obscure Object of Desire*, his last and best movie. However, for some reason I followed the logical road with the idea that I could work in cinema later. This is what I have done to such an extent, producing movies related to logic and philosophy. On the other hand, at some stage, I completely stopped watching movies because of a lack of time and also because I think it is more interesting to live in reality than in fiction – these two reasons being connected.

During these years in Paris I progressively focused on logic. Logic did not appear suddenly to me as a kind of new creature of which I became fascinated or obsessed, running after it. Before really studying logic I liked abstract mathematics and philosophy. The substantive word "logique" appeared to me, I liked it and discovered that it was not just an illusory reality, that there is a real substance in it. I was lucky to be in a place and time where modern logic was taught and where there were even degrees of logic. I went on doing research in logic and I am now a professor of logic. I do not see logic as a field among others. The meaning of the Greek word "logos" is related to science, reasoning, language, relation. Logic is a very rich notion. As Rougier put it: "Le *logos*, voilà la création du génie grec, dans les sciences, les arts, en morale et en politique; et le logos veut dire tout à la fois *discours*, *raison* et *raisonnement*, *rapport* et *proportio*" (The *logos*, here is the creation of the Greek genius, in sciences, arts, moral and politics; and *logos* at the same time means *discourse*, *reason*, *reasoning*, *relation* and *proportion*) [40]. I gave a talk about that in 2013 during a congress organized by my friend Olga Pombo at the occasion of the unification between the two main Universities of Lisbon, and the corresponding paper has been published under the title "Philosophy, mathematics, logic: three sisters".

3.3 *Meeting Newton da Costa and Leaving Paris*

I had had the idea of studying abroad for a couple of years. I thought mainly of the United States, because I was attracted by the country itself, in particular California, and because I know that there are very good universities there. However, no concrete opportunity appeared for me to go there at this time and instead of going to North America, I went to South America – different continents but nonetheless America. For someone from Europe they certainly have something more in common than just the name. Before knowing the work of da Costa, I had not thought of going to Brazil. However, while working on da Costa's paraconsistent logic I started to think it was a good idea. Brazil is famous for carnival and soccer – I was not really interested in such things – but also for the Amazon forest and the contrasting futuristic project of Brasilia, capital of the land of the future, created out of nothing.

While working on my Master's I wrote a letter to da Costa, but did not succeed in getting in touch with him. This was before e-mails and standard mail did not work well, nor did I have the exact address. By chance da Costa showed up in Paris in January 1991, just few days before the application deadline for a 1-year Lavoisier grant from the French Ministry of Foreign Affairs to go to Brazil. Da Costa gave a series of talks and I attended one at IHPST. This was quite spectacular. I was introduced to him and I met him nearly every day for 1 week. The contact was very good. Da Costa asked me why I was interested in paraconsistent logic. He was curious, but also apprehensive and worried, because many people are attracted to paraconsistent logic because they venerate contradictions – a particular case being the Marxists. When I told him that my interest was about the foundations of logic he was relieved. We shared the same interest; he wrote a book entitled *Ensaio Sobre os Fundamentos da Lógica* (Essay on the Foundations of Logic) that I would later translate into French [7]. He wrote me a letter of invitation I was able to join with my application. Some weeks later I faced a jury of not less than ten very serious men at the Ministry of Foreign Affair in Paris. To go to study logic in Brazil was quite weird, moreover I did not speak Portuguese. But I was supposed to go to the best University of Brazil, USP (Universidade de São Paulo) and I additionally received a letter of support from Michel Paty of REHSEIS[12], a philosopher of physics and good friend of da Costa, who had been cooperating with USP for a couple of years. Shortly after the interview I received a positive reply to this grant. Before traveling to Brazil I wrote two papers based on my Master's thesis and went to Portugal to learn the language. I gave a talk at the University of Braga and visited Porto and Lisbon. I liked Lisbon very much, and it continues to be one of my favorite cities in the world. In August I flew from Paris to São Paulo.

My departure from Paris in 1991 was a turning point in my life. On the one hand, I received a grant, which in itself was enough for housing and daily expenses, and after that I always succeeded, not always easily, by earning my life doing research and/or teaching logic. On the other hand, I left France and never went back there to live permanently. During my studies in Paris I survived financially with a little money from my parents,

[12] REHSEIS was a CNRS research institute about history and philosophy of science founded by Paty, attached to Paris 7; it has now been incorporated into an institute called SPHERE, CNRS-Paris 7.

some grants and some paid work I did. At this time in France, there were some grants based on family income for undergraduate students, provided that the student progressed normally with his studies. Since my father had a quite low salary I received this kind of grant and was also housed in a low cost "cité universitaire" where priority was given to people with low income (for a limited duration of time). At the Master's level the grants were based only on the quality of the studies of the students; I received such kind of grants for 2 years. These grants were helpful but far from enough for daily living. To get more money I started to give private lessons. In Paris there is a good market for that. I first mainly gave lessons to secondary school students, helping them to do their homework in all fields. However, I progressively concentrated on giving philosophy lessons to high school students preparing for the "baccalauréat". I was quite good at that in the sense that my students succeeded in obtaining good marks at the exams. However, it was a bit tiring and time consuming because I had to go to the people's home in all districts of Paris. Nevertheless, it was very interesting in the sense I was in touch with a great variety people and saw their homes. I stopped doing this when at some stage I started to work at the "Lycée autogéré" of Marly le Roy.

We taught students in their homes or in our homes. I did this for about 3 years. I gave a class of philosophy twice a week. This was quite an interesting experience. The salary was low but it was a regular income. The students were interesting people and we had many discussions. In general, I think teaching is a nice activity and that we learn a lot by teaching. The final year I also worked in a marketing company. We worked mainly through using the telephone. We had to phone to a huge quantity of various people, asking them all kinds of questions. At some point we worked for Arthur Andersen. Consulting agencies were becoming important in France at this time. We had to phone to the most important companies and succeed in interviewing of one of the main directors; this was a detailed interview that lasted about 1 h and consisted in asking all about what he expected from a consulting agency. It was quite challenging. I learned a lot about the services of consulting agencies and about how to succeed to talk to the boss. At some point, I also applied to work in the new MacDonald's in the Latin quarter; they wanted students for part-time work. I thought it could be a funny experience and would give me a regular income, but I was not selected.

I left France when I was 26 years old. So, roughly speaking I have now spent half of my life in France and the other half outside of France. While I regularly return to France on visits, and although I like the country very much for a few days or a few weeks at a time, I am rather critical of the French intellectual world. Moreover, I am becoming doubtful about the evolution of the culture, society and politics in France. Concerning the French academic world, the people are often both arrogant and ignorant at one and the same time. They think they know and yet they do not know, which is the worst kind of ignorance (Blaise Pascal wrote a nice text about two kinds of ignorance in *Pensées*). I think that the division between universities and CNRS is not a good idea. In Switzerland and in Brazil, in contrast to France where CNRS researchers get a full salary and have no real obligation to teach, the science foundations (respectively SNF and CNPq) restrict themselves to giving grants in support of projects directed by university professors.

4 Research Around the World (1991–2002)

I did two PhDs in about 5 years while living in Brazil, Poland, France and the USA, and traveling to many other countries – in particular making two round-the-world trips. I defended a PhD dissertation on universal logic in the Department of Mathematics at the University of Paris 7 in July 1995 and a PhD on logical truth in the Department of Philosophy at the University of São Paulo in March 1996. At the same time that I was working on my PhDs I was writing papers, assisting conferences, presenting talks, meeting and having discussions with a lot of people around the world. I do not think it is a good idea to work on a PhD in complete isolation, writing a hundred pages, not connected with what is going on, that nobody will never read. I believe that research is a collective endeavor that flourishes by interaction between human beings. On the other hand, I think it is good to have moments of isolation to concentrate on writing. I wrote each of my PhDs in about 1 month after months of thinking and discussion. When I write a paper this is also the methodology that I use: after thinking about a topic and presenting it to conferences, I sit down and write the paper in 1, 2 or 3 weeks. I do not work on only one topic. I work on many topics that are in gestation and, depending on the circumstances, one or another will materialize in a paper. During this 10-year PhD and post-PhD period, I gave about 200 lectures and wrote about 50 papers.

4.1 Interaction with Newton da Costa in São Paulo

I arrived in Brazil in August 1991. I didn't really like the country at first sight, but only after a couple of years. It certainly changed my life because the Brazilian spirit is very different from many other countries. It is quite difficult to know exactly at which point I changed, because I do not remember exactly how I was before, but when I go to France I see how much people are different. In Brazil, in general, people smile and are kind. This is good for physical and psychic health. There is no depression here. It will maybe arrive when people are "rich": locked in their house watching TV and taking their car to go to the supermarket, a project of society promoted by politicians in Brazil as in many other countries. However, it not clear that they will succeed in Brazil. Other countries in Central and South America are like Brazil: Mexico, Costa Rica, Peru, Bolivia and Chile. Yet in Brazil it is stronger. Maybe this is a kind of indigenous spirit inherited from the natives to which outsiders have been mixed. Moreover, in Brazil it is reinforced by the Portuguese culture, which is quite different from the Spanish culture that predominates in other countries of Central and South America.[13]

In his famous book, *Brazil, Land of the Future*, Stefan Zweig wrote the following about Brazil: "Arriving in Rio, I received one of the most powerful impressions of my whole life.

[13] Some people use the expression "Latin America" to talk about Central and South America. I think this is misleading. Can we not say that French-speaking people in Canada are Latin? In Brazil "latinos" has a pejorative flavor, designating in a negative way "other" people of South America, those who speak Spanish.

I was fascinated, and at the same time deeply moved. For what lay before me here was not merely one of the most magnificent landscapes in the world, a unique combination of sea and mountain, city and tropical scenery but quite a new kind of civilisation. There were colour and movement which fascinated and never tired the eye; and wherever one looked there was a pleasant surprise. I was overwhelmed by a rush of joy and beauty... Brazil's importance for the coming generations cannot be assessed even by the most daring calculations. I knew I had looked into the future of our world." [51]

My first contact with Brazil was not easy because I arrived in São Paulo, which at the time was very dirty and polluted. The country was not yet open to mondialisation. There was very few imported products. The people were dressed with ugly clothes. The price of a home phone line was the same as a car. At first I lived on the campus of the University of São Paulo, far from the center with quite nothing inside. One main attraction was a collection of snakes at Butantã Institute.

I attended the seminar of da Costa. I soon realized that very few people, if not to say nobody, worked any longer in paraconsistent logic in Brazil. That was not a problem for me because my interest was to develop a general theory of logics and I was mainly interested in working on the theory of valuation, a general semantic theory developed by da Costa. I wrote two papers with da Costa on the theory of valuation and developed ideas on abstract logic, which I later wrote down in my paper "Recherches sur la logique abstraite", which are a central part of my PhD in mathematical logic. I was concentrating on "logical structures" in a Bourbachic spirit. Bourbaki had a strong influence in Brazil and in particular in São Paulo where André Weil, Jean Dieudonné and Alexander Grothendieck (about the visit of the latter see [2]) came for extended stays in the 1940s and 1950s. Da Costa was a former student of Edison Farah, a set-theoretist who had been a good friend of André Weil. He was retired and we used to visit him in his house; he told us many stories about the adventures of the Bourbachic tribe in Brazil.[14] Da Costa brought me also to the house of Miguel Reale, a person he had known for many years. Reale was an important jurist – the man responsible for Brazil's new civil code – who took a strong interest in philosophy and logic. He founded the *Instituto Brasiliero de Filosofia* (IBF), launched the *Revista Brasileira de Filosofia* (RBF), and the first series of conferences on philosophy in Brazil.

During my first year in São Paulo, I developed contacts with various students and colleagues of da Costa. Edelcio de Souza and his wife, Simone, helped me to discover São Paulo, in particular taking me to all the most famous restaurants in town. Andrea Loparic, who was working with da Costa on the theory of valuations and lived nearby the USP, invited me for lunches at her house – she speaks fluent French, having done her PhD in Belgium, and also takes an interest in psychoanalysis. I also had contact with Décio Krause, who lived in Curitiba, but regularly came to USP and Nelson Papavero, who worked in biology and put me in touch with a friend of his in Rio, who later on invited me to work with him. I met Otávio Bueno, who at this time was a young undergraduate student with long blond hair who stayed late at night at the library – we used to come back to town together by bus discussing many topics. I also had some contacts with a group of

[14] Farah proved a statement that Weil thought was false: the equivalence between the axiom of choice and infinite distributivity of intersection and union.

young French guys; I shared a flat with one of them and started to write a novel in French that was never published called *L'Oubli de Vivre* (Forgetting to Live).

In April 1992 I went on a 1-month trip outside of Brazil; I went to Chile and Argentina. This was quite interesting. Argentina is completely different to Brazil. I went to Buenos Aires and liked it very much. I had a colleague there whom I had met in Paris, Francisco Naishtat, who had lived in Paris during the military dictatorship and also had spent 1 year at Berkeley. He introduced me to colleagues and friends, in particular Mario Lipsitz, a former PhD student and friend of Michel Henri, who became a good friend of mine. I gave a talk at the Computer Department of the University of Buenos Aires, where I met Marcelo Conglio (who later moved to Brazil where he is presently the President of the Brazilian Society of Logic) and Carlos Alchourrón (known in particular for the AGM theory, the logic of theory change he developed with David Makinson and Peter Gärdenfors). In Chile I gave a talk to the logic group of Rolando Chuaqui at the Department of Mathematics of PUC (Pontifical Catholic University) in Santiago. I went North up to the desert of Atacama, and South to Ponte Arena and Tierra del Fuego via Chiloé Island. I climbed one of the most active volcanos near Puerto Monte and also visited Valparaiso.

In June 1992 I went to Rio de Janeiro for the second time (I had been there briefly in January). I visited Vera Vidal, a specialist of Quine whom I had met in Paris together with da Costa, who was at this time the director of the Philosophy Department of the Federal University of Rio de Janeiro (UFRJ), where I am presently working. I also met Luiz Carlos Pereira, a proof-theoretist, who did his PhD with Dag Prawitz in Sweden and who works at the PUC of Rio de Janeiro. This was the time of ECO 92, the very big United Nations Conference on Environment and Development that gathered more that 100 heads of government from the whole world. There were many events. I missed the concert of Philip Glass accompanying the projection of Godfrey Reggio's movie *Powaqqatsi* but I attended an Amazonian version of *A Midsummer's Night Dream* by Shakespeare directed by Werner Herzog with the telenovela actress Lucelia Santos, best known as Escrava Isaura.

In August 1992 I again went to Argentina, this time to take part to the 9th Latin American Symposium on Mathematical Logic, which took place in Bahía Blanca. This is a town about 700 km south of Buenos Aires where there is an important group of logicians founded by Antonio Monteiro, a Portuguese who first moved to Brazil and the settled down in Bahía Blanca. The people there work mainly on algebraic logic, with connections with the work of the Romanian logician Grigori Moisil, who developed De Morgan algebras. The congress took place between August 10–14, 1992. I was supposed to go back after 1 year in France, i.e. July 31, and my air ticket had 1 year's validity. It was not possible to extend this validity and I had to buy a new air ticket. I had little money and I went from São Paulo to Bahía Blanca by bus (2 days of travelling between São Paulo to Buenos Aires) and train (one night on the train from Buenos Aires to Bahía Blanca).

At the congress I presented my joint work with da Costa on the theory of valuation. There were celebrities like Jerome Keisler and the Polish logician Stan Surma, one of the major figures of logic of Poland after World War Two (WW II), who directed a group on logic in Krakow. On the night train back to Buenos Aires he was on the seat next to me and I told him that I was going to spend 1 year in Poland. He drew me a map of logic in Poland. He had left Poland during the communist time and now lived in New Zealand, but knew all the Polish logicians, most of them having been his students. He was traveling with his son, Charles. Later in China in 2007 during the 2nd World Congress on Universal Logic where he went together with his father who was an Invited Speaker, Charles told me about his adventures in Africa where his father had taught mathematics in the jungle upon escaping communism in Poland. In China I also made an interview/movie of Stan Surma relating his carrier, in particular how he was sent to Moscow and studied with Kolmogorov. He was not converted to communism but to constructivism.

4.2 In Wrocław, Poland Hosted by Jan Zygmunt

After 1 year I could have stayed in Brazil, incorporating myself in one way or another into the Brazilian academic system – da Costa supported me to stay longer. However, when I told him that I wanted to go to Poland he also strongly supported me in that. He had been several times in Poland himself and had invited Polish logicians to Brazil during the 1970s. For reasons that are even now not completely clear, logic flourished in Poland between the two world wars, and Poland became the strongest centre of activities of logic in the world; this was known as the Lvov-Warsaw school – see Woleński's books and papers. When WII started, Alfred Tarski was visiting the USA and was forced to stay there, settling in UC Berkeley after some years of uncertainty and developing what would be the main group of logic in the world for three decades. He visited Poland during the communist times only a few times. At the University of Wrocław I had an office where it was written: "Tarski was here in 195?" – I do not remember the exact date. Why did I want to go to Poland? Had not the center of logic moved from Poland to California? Not for what I was interested in. My main interest was the theory of the consequence operator, initially developed by Tarski at the end of the 1920s. He did not go on working on this theory when in California and developed model theory there. Nevertheless, people in Poland after WWII kept working on that and related subjects. In particular Jerzy Łoś and Roman Suszko, and their student Ryszard Wójcicki. I tried to get in touch with Wójcicki. He was a friend of da Costa and had spent some months in Brazil. However, it was not clear where he was located and whether he was still working on this subject – indeed at some point he focused on the philosophy of science. I ended up in Wrocław working with Jan Zygmunt, a former student of Stan Surma and at that time the Editor-in-Chief of *Studia Logica*.

My reasons for going to Poland were not exclusively logical. After the experience of 1 year of being immersed in a completely different culture, I wanted to repeat the experience with a different environment. I had always had good feelings about Poland and I was curious to live in a country of the second world.[15]

When I was young it was a mystery to us of what life was like in the Soviet empire. It was difficult to go there, those who went were communists and what they said was not taken seriously, which contrasted the anti-communist propaganda based in particular on stories of people who had escaped and were refugees. When I was 14 years old I

[15] The tripartition of the world into First World, Second World and Third World, was effective after Alfred Sauvy coined the expression "Tiers Monde" in a paper published August 14, 1952 in the French magazine *L'Observateur*.

went on an exchange trip with my school and stayed 2 weeks with a German family in Bayreuth. One tourist attraction there was to go to the border. The separation between East and West was much more than the symbolic wall separating Berlin into two parts. It was a metallic wall of grids and explosives of thousands of kilometres separating the communist world from the rest of the world. At the border we gazed at the communist world, a hundred metres away from where we were, looking at people we could not meet. The whole atmosphere looked different, old-fashioned, with old cars. West Germany, on the contrary, was hypermodern, much connected to the USA, especially in this Bavarian region that was controlled by the Americans.

The Berlin Wall was destroyed in 1989, and the Soviet Union came to an end in August 1991 when I was in Brazil. When I arrived in Poland it was a transition period where most people were still living in the old communist style. At the same time there was a very strong capitalist development, a kind of conquest by different kinds of people coming from all over the world. The Polish people were quite optimistic because many of them have strong sympathy for the USA and antipathy for the Russians who dominated the country for many centuries before and during communism. I arrived in Poland in Warsaw in September 1992 aboard the legendary Paris–Moscow train of communist times, which does not run anymore. The situation was strange and confusing.

Paris–Warsaw was a 24-hour trip. The train arrived at the end of the afternoon on a Friday. I was expecting some French officials at Warsaw train station but there was nobody there. I did not speak Polish, had very little money in my pocket, and did not know where to go. For my stay in Poland I had applied for another grant from the French Ministry of Foreign Affairs. It was quite easy to get this grant, there was no impressive interview like for the grant to Brazil. Very few people were attracted to come to Poland. This exchange program was still working as during the communist times, so, in particular, our grant was in local currency. My grant was less than 100 US dollars per month, which was supposedly enough to survive on in Poland (on the other hand, Poles who came to France would get their grants in French currency, which corresponded to an amount of more than 1000 US dollars). At Warsaw train station I asked the help of a friendly Polish man who had traveled with me on the train and I met a French girl who was in the same situation as me. The guy helped us to call the French Embassy. They said that they could do nothing right now because it was Friday afternoon but offered us to stay in an apartment for visitors in the French Embassy for the weekend. I stayed with that girl and we visited Warsaw. On Monday morning we had a meeting with some authorities who put us on a train to Wrocław. Coincidentally this was also the destination of this girl (the only other French grantee in all Poland) who was going to Wrocław's art school, which is quite important.

In Wrocław Jan Zygmunt took care of me, first lodging me for a few days in a student dormitory, and then getting me a small flat in the house of scientists – designed during the communist period for professors–on Maria Curie-Skłodowska street. At the University of Wrocław there is a Department of Logic and Methodology of Science, like in several other universities in Poland, of which Zygmunt was the director. It is traditionally one of the most important centres of logic in Poland. After WWII, the Soviets annexed large parts of Eastern Poland, including the city of Lvov (now this area is in the Ukraine), and gave a part of Germany, the Silesia, back to Poland; thus the city of Breslau was renamed with its original name, Wrocław. The University of Breslau was important during the German time; in particular, this is where Schrödinger worked and where in 1884 Frege's book *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl* (in English: *The Foundations of Arithmetic: the logical–Mathematical Investigation of the Concept of Number*) was published.

After WWII the university went on to be an important university in Poland, one of the main ones indeed, with a lot of students coming from all of Poland and in my time also from other countries. After WWII Wrocław was a kind of transposition of Lvov; many Polish people from Lvov moved to Wrocław, including academic people and the Ossolineum foundation (which is famous for its library and is also an academic publishing house, the original publisher of *Studia Logica*). Jerzy Łoś, one of the leading figures of logic in the 1940s and 1950s, worked there and in 1949 he published, with Wrocław University Press, his famous monograph about logical matrices [31]. The Department of Logic had several unpublished manuscripts by him, some in French, which Zygmunt showed me. Jan Zygmunt was much interested in the history of logic. He is the official editor of Tarski's works in Poland – Tarski's son Jan Tarski visited him when I was there. After seeing Zygmunt's paper "Life and Work of Mojżesz Presburger" [52] I had the idea of asking him to write a similar paper about Adolf Lindenbaum, who is considered to be one of the most prominent Polish logicians of the inter-war period. After many years, the paper is now finished and has just been published in *Logica Universalis* [53].

I was interested in the work of Lindenbaum because I was working on a generalisation of the maximalisation theorem, which is attributed to him. Zygmunt showed me a quote of a Polish logician making a nice metaphor comparing a maximal theory to a glass that is so full that adding only one more drop causes it to overflow. During that period I made good progress in my work. I fully developed the idea and project of universal logic – this was where I decided to choose the expression "universal logic" – and I wrote my PhD thesis about logical truth that I sent to my advisor Philippe de Rouilhan in Paris. I gave various talks at Polish universities, in particular in Łódź (invited by Grzegorz Malinowski, a many-valued logician directing a strong group there), Krakow (where I met Wroński), Kielce (where I met Prucnal), and went to the Czech Republic near Prague to take part in LOGICA'93 – one of the first meetings of the ongoing LOGICA series of events.

I very much enjoyed the centre of Wrocław, with its Gothic architecture and many canals. It was like time had stopped in the 1930s – and then there were the suburbs with communist buildings from the 1960s. I usually had lunch at the Ossolineum restaurant, which was next door to the Department of Logic. I liked Polish food very much and also the fact that people drank tea all the time. I prefer tea to coffee. Coffee in Poland was very expensive but tea very cheap and there was a huge variety, some very good types, imported from China; this was a positive consequence of the Soviet regime.[16] The atmosphere was interesting, a mixture of the old communist world with the rising of a new world with a lot of adventurers. The Russian army was on the leave. There were some Russian markets where everything was sold, from watches to planes. I bought some oil paintings and started

[16] I have never been a Marxist. I think the materialist philosophy of Marx is rather crude. On the other hand, the communist world as it emerged from Marxism–Leninism was interesting because it was preserved from consumerism, all people were highly educated and there were no beggars in the streets.

Fig. 4 Consciousness 3 – A painting that I realized during the Polish winter of 1993

to paint (Fig. 4). This was quite an inspiring atmosphere and I finished my novel *Forgetting to Live* and wrote a couple of short stories.

It was quite cold during the Polish winter, but there was beautiful, sunny snowy weather and a nice landscape with rivers and trees. I remember ducks landing on the ice and crowds of ravens. At first I was quite isolated, in particular because I did not speak Polish; I tried at the beginning but just gave up. In the end I knew a lot of people and there were parties nearly every day. The city was booming, the economy was strong, new bars and shops were opening all the time. Was elected as mayor of Wrocław a logician, former student of Zygmunt. Since my grant was low Zygmunt suggested that I give classes at *Alliance Française* where his daughter was learning French. I was engaged there and gave a 2-hour class per week. I received a salary similar to my monthly grant. It was an advanced class and I just had to talk in French about any subject I liked. That was fun. I also organized a cine-club at *Alliance Française*. I started with the projection of a French movie that was not difficult from the linguistic point of view, *La Guerre du feu* (Quest for Fire) by Jean-Jacques Annaud. In Poland, French language and culture are very popular, this is so for various reasons, one of them being that Napoleon created the Duchy of Warsaw in 1807, liberating Poland from Prussia. Even if this era had a very short life of 7 years, it was very important for Polish people. After that Poland was occupied by Russians and Austrians until 1918.

At the end of my 1-year stay my idea was to go to France to defend my philosophy PhD and go ahead to California. Nothing of that directly worked out. I applied for a Fulbright grant in Paris when I was in Poland and was selected for an interview but was not able to go because I did not have enough money to pay for the trip. When I was in Poland I was registered at the Swiss consulate, this is, in fact, mandatory for any Swiss citizen abroad and I received the magazine for Swiss abroad. In this magazine there was an article about the Swiss National Science Foundation (SNF) encouraging Swiss from abroad to apply for SNF grants. I decided to apply for an SNF grant for young researchers, to go back 1 more year to Brazil while waiting for the next deadline to apply for a Fulbright grant. The man responsible for SNF grants for Swiss abroad was the President of the Swiss Academy of Science, at that time Jean-Daniel Nicoud, Professor at the EPFL (Polytechnical Federal School of Lausanne), the father of the optical mouse. He told me that it would not be a problem for me to get the grant but that he wanted to meet me before making a decision. I planned to go there just after the end of my stay in Poland, which was extended for a couple of months as requested by Zygmunt to the Polish academic authorities.

Back in Paris at the end of 1993, after having spent the end of the year with my family, I went to Lausanne early in 1994, met Nicoud and gave a talk at his institute (EPFL Institute of Microinformatics). I received a positive reply for the grant but had to wait about 6 months before effectively getting it – including the air ticket. During this time I went to Paris, trying without success to defend my PhD but having a positive interview at the American Centre for a Fulbright grant that I would use the following year. The future was good but I had no money for the present. A good friend of mine, Pascal Duval, helped me for accomodation and I succeeded to survive by signing a contract with the publisher Masson to translate the book of da Costa into French and giving some private lessons on logic to the President of the *Cause Freudienne*, the Lacanian Association of Psychoanalysis. Her name was Liliane Majhoub-Trobas and she wanted to know more about logic because Jacques-Alain Miler was giving a seminar making many references to logic (the main interest of Lacan was topology, but Miller is more interested in logic). I enjoyed doing that. In June 1994 I came back to Poland to send to Brazil the many boxes of books and papers I had had to leave there. I also took the opportunity to go to the LOGICA meeting in the not so far Czech Republic for a second time. I presented my first talk about universal logic there. The corresponding paper, entitled simply "Universal logic", was subsequently published in the proceedings.

Let me now explain the situation with my philosophical PhD. In France, for a PhD it is just necessary to write a monograph, there are no classes to attend, no necessity, therefore, to be on location. Before leaving Paris in 1991 I matriculated for a PhD at the Philosophy Department of Paris 1 with de Rouilhan and for a PhD at the Mathematics Department of Paris 7. I sent my philosophy PhD to de Rouilhan from Poland and waited for his feedback. He did not like the work. When I met him in Paris he told me that this was a completely different kind/style of work from what he was acquainted with and that we would never reach an agreement. He told me he would help me to work with another colleague. That was quite fair, but I was not able to find anybody who was interested. The next in line was Dubucs and he also did not like the work either. So I left Paris for Brazil without solving the problem and decided to concentrate on my mathematical PhD. In Paris I also met my advisor for this PhD, Daniel Andler, who encouraged me to go ahead.

My philosophical PhD is called *Sur la Vérité Logique* (On Logical Truth). This is a new framework for logical truth for structures with one binary relation. I discuss the crucial notions of constants, variables, quantifiers, infinity, equality and function. I developed a theory based on an idea of Wittgenstein to have a (non-ambiguous) name for each object and applied this theory to classical logic but also to paraconsistent logics – the standard model theoretical approach does not properly work for such logics. Later, I presented this work for a PhD at USP in São Paulo. This is still one of my less known works.

4.3 São Paulo Again and Los Angeles

In July 1994 I arrived in São Paulo for the second time in my life. With a grant from the Swiss National Science Foundation, I thought I would have a comfortable financial situation, but this was not the case – I would have to wait another 8 years for that. The Brazilian government had just launched a new currency, the "Real". This completely stopped the inflation and that was good. However, quite unexpectedly the currency rose everyday for a couple of months. The Real was launched on the basis of a 1 to 1 correspondence with the US dollar, but after a few months, the value of 1 Real was about 1.20 dollars. This was a general situation vis-à-vis all currencies, including the Swiss Franc, and therefore I lost an important part of my grant. I found a room to rent close to Paulista Avenue in front of Maksoud Plaza, at the time the most prestigious hotel in the city, the Rolling Stones came here during my stay.

São Paulo turned out to be one of the most expensive cities in the world. Richard Sylvan and Graham Priest visited Brazil at this time, and I remember that they were quite surprised at the situation. This was the first time I met these two pseudo-Australian logicians and the last time in the case of Sylvan, who unfortunately died 2 years later on a trip to Bali. I gave Sylvan my paper "Universal logic" that I had written just after my talk at LOGICA'94 in the Czech Republic that I was finishing to prepare to send for publication. The next day he gave me back the paper; he had read it in detail and made some annotations (I still have this copy).

In August 1994 I wrote my PhD of mathematical logic in about 1 month. I was working hard day and night putting all my results together. The title of this work is *Recherches sur la Logique Universelle (Excessivité, Négation, Séquents)*. The first part contains some general results about completeness based on the concept of relatively maximal theory. Then there is a general theory of negation, followed by the study of some paraconsistent systems. In the third part I establish the connection between valuations and sequents and develop some general results about cut-elimination. My PhD has led to about five published papers. The main result is a theorem that establishes a close relation between sequents and bivaluations. I had been looking for that for 2 years and in August 1994, while writing my PhD, the solution naturally appeared. I had the basic idea by reading the first paper of Gentzen, which is about Hertz's *Satzsysteme*, his main inspiration for sequent calculus and the cut rule. This theorem is a nice, new and non-trivial result – a minimum condition I think to get a PhD in Mathematics. Having finished writing my PhD in September 1994 I was eager to send it to Andler, since I wanted to defend it by the end of the current French academic year, i.e. July 1995.

The process for the defense was quite long. At the Department of Mathematics of Paris 7, it was as follows: if your PhD advisor thought your thesis was good enough he submitted it to a commission of the department, of which he was not part. This commission sent your work to referees chosen by them who were to give their advice; on this basis the commission allowed or refuted the defense. Andler was very careful and read my PhD in detail; it took him about 3–4 months to do so. Then he presented it to the commission of the department. I was a bit nervous; I had to wait and relax. I gave copies of my PhD to different colleagues, in particular to Luiz Carlos in Rio who carefully read the part on sequent calculus and corrected some minor mistakes. The chosen referees gave approval for the defense in April 1995 and the defense was carried out in July 1995.

Early 1995 I went to California to enjoy my Fulbright grant. I was invited to the Department of Mathematics of UCLA (University of California at Los Angeles), by Herbert Enderton (1936–2010). I knew his books of introduction to logic and set theory, which are very nice. UCLA together with UC Berkeley was at this time one of the most important centres for logic. I arrived in Los Angeles and found a flat to rent in Venice. I was going by bus from Venice to the Westwood area where UCLA is located, via Santa Monica. I liked to see the sunset on Venice beach and sometimes I walked on the beach from Venice to Santa Monica.

At UCLA I attended seminars in the Department of Mathematics and in the Department of Philosophy. People from the Math Department attended seminars at the Philosophy Department (but not vice versa); in particular Tony Martin who had a double position in both departments, but also Yiannis Moschovakis, the greatest Greek logician since Aristotle, as he liked to say. In the Philosophy Department I attended the seminar of Joseph Almog and also that of David Kaplan. This was a small group of people who knew each other and used to go out to dinner together every week. Kaplan also invited people to parties at his house; I went there once. It was friendly and quite provincial. It was not clear at all that they knew what was going in the rest of the world, even outside California, or LA. Once during his seminar Almog told us that he had discovered a very nice paper by a totally unknown guy named John Corcoran. In the Math Department the situation was a bit strange with some weird people showing up. There were seminars of mathematic logic, in particular the Cabal seminar. No philosophers came but there were computer scientists at the main seminar. There was a wine and cheese party after this seminar, where I gave a talk.

Mathematical logicians at UCLA liked to play badminton. I was sharing an office with a nice guy called Gary, who had come back from Israel after having worked with Shelah. He taught me how to play badminton in a way I did not know, the idea being to hit the shuttlecock so that your partner cannot get it like the ball in tennis. When I was a child I used to play badminton in a way similar to what is called frescobol in Brazil and matktot in Israel, the objective being to keep the shuttlecock in the air. This is a good example of how by changing the teleological rule, we completely change the game. I wrote a paper on this topic later entitled “Rules of the game” (joint paper with Tarcisio Pequeno).

I had no car and only went outside of LA twice: one visit to Ojai and one to some natural parks near Palm Springs organized by the Fulbright program for all Fulbright fellows. The director of the program was a nice lady called Ann Kerr, widow of the former president of the American University of Beirut (AUB) who was shot on location–later, in 2012, I organized the 3rd World Congress on the Square of Opposition at AUB. During this tourist tour I remember that we stopped at a shop in an Indian reserve and an Indian woman told us a bit about the story of her tribe. At the end of her talk one of the Fulbright Fellows wanted to thank her and asked her how to say “thank you” in her native language. She replied that there was no word for that in her language. I think that this is a very interesting situation that can be interpreted in two opposite ways: these natives were too rude to express their thanks or they were advanced enough not to need to transform their feelings into a formality through a word. It is, indeed, very easy to say “thank you” all the time without feeling anything.

I went back to Paris at the beginning of July 1995 to defend my PhD in the Mathematics Department of Paris 7. LA airport was under threat by a person nicknamed the Unabomber (later on identified as Theodore Kaczynski, a former mathematician of UC Berkeley) and in Paris I escaped a bombing in the underground train RER, which killed many people. My PhD defense was not a big event. I had prepared some slides but the projector did not work. There was no party after the defense. In France it is a tradition to organize a “Pot de thèse”: drinks and food paid by the student and/or his family after the defense. I did not live in Paris, none of my family came to my defense, I had little money of my own, and what little I did have I was trying to save for the future. The members of the jury were Jean-Louis Krivine, Jan Zygmunt, René Guitart, Michel de Glas and Daniel Andler.

4.4 Copacabana

After my PhD I was not eager to have a permanent position as professor at a university. I wanted to go on doing research. I wanted to go to Rio de Janeiro, this was the place I wanted to live for a couple of years. I had made contact for this with Mauricio Kritz, whom I knew through his friend Nelson Papavero, himself a good friend of Newton da Costa. Kritz worked at the LNCC = Laboratório Nacional de Computação Científica (National Laboratory for Scientific Computing), a research lab located near the Sugar Loaf. One of the bosses there was his former PhD advisor Marco Antonio Raupp, later Minister of Science in Brazil. Kritz had some interest to apply logic and category theory to biology. I arrived in Rio in July 1995 and had some difficulties to receive a grant, in particular because my profile of post-doc from abroad was not part of the system, and Brazilian bureaucracy is quite tricky. I had to wait nearly 1 year without a salary. Also my laptop computer broke because of voltage variations and I had no money to buy another one. Nevertheless I did not give up.

At some point there was a permanent job for a logician with the Department of Philosophy at the University of Brasilia. I went there to meet the people and to discuss the conditions. I had already been in Brasilia in 1992 during my first stay in São Paulo with my friend Hilan. They were ready to contract me but I decided not to go. I preferred to stay in Rio without a job than to go to Brasilia. The reason is that I think the place for daily life is not a secondary thing, it is really important. Brasilia is an interesting town to visit for a couple of days, but I would never like to live there. One friend of mine, who did his PhD in Oxford, told me that Oxford was the worst place he had lived. I have always been very cautious about the place I live. In Rio I rented a room in a flat in Copacabana on Atlantic Avenue, facing the ocean. Every morning I went to swim and run on the beach. I started running more and more, and after a couple of months I was easily running the full lengths of the beaches of Copacabana, Ipanema and Leblon without difficulties. My friend Otávio is also a good runner, he likes to participate in ultra-marathons (100 miles).

In January 1996 I was invited by my friend Carlos Knudsen to give a summer course in logic at the Math Department of the Federal University of Pará in Belém, in the North of Brazil, some 2500 km from Rio. Belém is a nice town on the estuary of the Amazon river.

I visited the interesting Émil Goeldi museum and research centre on Amazonia. Since arriving in Rio I had been going to USP every two weeks, staying at Marcelo Coniglio's place, which was in the centre of São Paulo in a trendy area. I was going there to attend da Costa's seminar at USP. On March 27, 1996 I defended my PhD of Philosophy on logical truth at this University.[17]

When I arrived at USP in 1991 I did not matriculate as a PhD student, just as a visitor. However, at some point the people of the Philosophy Department thought it would be nice if I had a PhD from their department. I therefore matriculated and presented the work I originally intended to present at Paris 1 with de Rouilhan. Many years later it was quite useful for me to have this diploma to get a job in Rio. USP is considered the best University in Brazil and also the best in all of Central and South America. Many famous scientists have worked at USP, like the physicist David Bohm, who emigrated to Brazil after he was fired from Princeton University during the McCarthy era.

After about 1 year Kritz succeeded to get me a grant at the LNCC. It lasted 1.5 years. In 1998 my grant was cut. This was the first economical crisis of the "plano real" and the Brazilian government decided to cut the grants of foreigners. I again spent more than half a year without income. Petrucio, a friend of mine who was working at the Math Department of UFF, a university on the other side of the Guanabara Bay, then succeeded to get me a position as an Invited Professor, and I also got a research grant from FAPERJ – Rio de Janeiro Research Foundation – to go on working at LNCC, which had been relocated to Petrópolis ("City of Peter"), the former Imperial city, 60 km from Rio in the mountains. I circulated between these different locations, always based on Copacabana.

Despite my financial instability during this period, I went to many conferences in Brazil and in the world. This may appear quite paradoxical, but is explained by the fact that, on the one hand, I was member of some research projects with money for traveling to events and, on the other hand, I started to be an invited speaker. Here is the list of events in which I took part during my stay in Rio from 1995 to 1999. I have indicated the name of the event, dates, location and the title of the lecture.

- 9th Brazilian Meeting of Philosophy, September 3–7, 1995, São Paulo, Brazil, *Applications of Paraconsistent Logic to Justice and Law*
- Pratica'96 (Proof, Type and Category) April 24–25, 1996, Rio de Janeiro, Brazil, *Rules, Derived Rules, Permissible Rules and the Various Types of Systems of Deduction*
- 11th Brazilian Meeting of Logic May 6–10, 1996, Salvador, Brazil, *The Mathematical Structure of Logical Syntax*
- 3rd WoLLIC (Workshop on Logic, Language, Information and Computation) May 8–10, 1996, Salvador, Brazil, *Paraconsistent Model Theory*
- 3rd Bariloche Meeting of Philosophy August 29–31, 1996, San Carlos de Bariloche, Argentina, *Present Philosophical Tendencies*
- 2nd Annual Conference on Applications of Logic in Philosophy and Foundations of Mathematics April 24–27, 1997, Karpacz, Poland, *Universal Logic in Perspective*
- 27th International Symposium on Multiple-Valued Logic May 28–30, 1997, Antigonish, Canada, *What is Many-Valued logic?*

[17] Among the members of the Jury was David Miller, who was secretary and friend of Karl Popper. He fully approved my dissertation with a 10/10 mark – I guess because it is 100 % irrefutable.

- 1st World Congress on Paraconsistency July 30–August 2, 1997, Ghent, Belgium, *What is Paraconsistent Logic?*
- 4th Wollic (Workshop on Logic, Language, Information and Computation) August 20–22, 1997, Fortaleza, Brazil, *Classical Negation can be Expressed by One of its Halves*
- 20th Brazilian Congress of Applied and Computational Mathematics September 8–12, 1997, Gramado, Brazil, *Neojunction: between Conjunction and Disjunction*
- 4th Brazilian Meeting of Analytic Philosophy October 6–9, 1997, Florianópolis, Brazil, *Philosophical Aspects of Non Classical Logic*
- Logic, Proofs and Algorithms April 15–18, 1998, Campinas, Brazil, *Sequents and Bivaluations*
- Visions of Science – Meeting with Sokal and Bricmont April 27–28, 1998, São Paulo, Brazil, *Round table on Logic and Language*
- Pratica'98 (Proof, Type and Categories) June 18–19, 1998, Rio de Janeiro, Brazil, *Logical Rules and Logical Laws*
- 4th Bariloche Meeting of Philosophy June 24–27, 1998, San Carlos de Bariloche, Argentina, *Philosophy and Logic: Asset and Perspectives*
- Stanislaw Jaśkowski Memorial Symposium July 15–18, 1998, Torun, Poland, *The Paraconsistent Logic Z*
- 20th World Congress of Philosophy August 10–16, 1998, Boston USA, *Do Sentences Have Identity?*
- 8th National Brazilian Meeting of Philosophy – ANPOF September 26–29, 1998, Caxambu, Brazil, *Round table on Translations between Logics*
- 11th Brazilian Meeting of Logic May, 24–28, 1999, Itatiaia, Brazil, *Singular Terms in Mathematical Logic*
- 5th Wollic (Workshop on Logic, Language, Information and Computation) May, 24–28, 1999, Itatiaia, Brazil, *A Sequent Calculus for Lukasiewicz's Three-Valued Logic*
- 1st Principia International Symposium August 9–12, 1999 Florianópolis, Brazil, *Are there any Logical Principles?*
- 11th International Congress of Logic, Methodology and Philosophy of Science August 20–26, 1999, Kraków, Poland, *The Philosophical Import of Polish Logic*
- 1st Southern African Summer School and Workshop on Logic, Universal Algebra and Theoretical Computer Science December 1–10, 1999, Johannesburg, South Africa, *From Paraconsistent Logic to Universal Logic*

A series of papers correspond to these talks, which the reader can find in my complete bibliography, up to the present, at the end of this paper. During this period I made good progress in my logical research and had several new ideas. In particular I discovered an important connection between modal logic and paraconsistent logic, showing that S5 is a paraconsistent logic, this being connected with the square of opposition. Details of the development of these works are described in my 2007 paper "Adventures in the paraconsistent jungle" and my 2012 paper "The new rising of the square of opposition".

I wrote the two papers "What is paraconsistent logic?" and "What is many-valued logic?".[18] These papers are connected to clarification and understanding of basic notions

[18] Later I wrote a paper entitled "What is classical propositional logic?" I intend to go on writing other "What-is" papers – see the section "Future papers" in my list of writings.

of the fields of paraconsistent logic and many-valued logic. This can be called "logical philosophy" if we use this expression in a way similar to how Bertrand Russell used the expression "mathematical philosophy". Logical philosophy in this sense is neither mathematical logic, nor philosophy of logic.[19] The idea is to have a deeper understanding of the mathematical notions involved in connecting them to philosophical questions. This is not "philosophy of", because by so doing we are actively interacting with the object of study. Moreover, this is not just mathematics, because we are trying to understand the meaning of the constructions, not just to prove theorems. Modern logic is often rejected both by philosophers and mathematicians. Philosophers say that it is mathematics and they do not see the philosophical value of all this. Mathematicians say that this stuff may have philosophical interest but that this is not real mathematics, it is a lot of formal trash with no real mathematical value. Both are right to some extent. Modern logic is many times lost in a twilight zone between mathematics and philosophy, being neither one, nor the other. Yet it can be both and then it is beyond mathematics and philosophy, the mixture of the two leading to some real enlightenment.

During this period I worked on general philosophy. For the first time I attended a WCP: World Congress of Philosophy, the one in Boston in 1998. On this occasion I met Quine. I had been in touch with him when editing a special issue of the journal *Logique et Analyse* about logic in Brazil. He sent me a paper "Mission to Brazil" [37] recalling his 1-year stay in Brazil in the 1940s. Quine quite liked Brazil and learned to speak and write Brazilian Portuguese. He published an excellent book in Portuguese, which has not yet been translated into English *O Sentido da Nova Lógica* [36]. At this event in Boston I presented my paper "Do sentences have identity?" criticizing Quine's idea that sentences exist but not propositions because there are "no entities without identity" (one of Quine's most famous mottos).

I presented a talk about the present state of philosophy in Bariloche, Argentina, where every 2 years there is a big international congress of philosophy. In this talk I critically presented three tendencies of philosophy: traditional philosophy, analytic philosophy and new age philosophy. I emphasized in particular the weakness of academic philosophy in not being able to attract ordinary people. This was later published in my 2003 bilingual booklet *Tendances Actuelles de la Philosophie/Tendências Atuais da Filosofia*. In Bariloche I attended a talk by John Searle on money. I think he is a good example of what can done. Searle is one of the most famous philosophers alive, but nevertheless in the USA it is very difficult to find a book by Searle in a bookshop.

I faced this difficulty when at Stanford in 2000. I was looking for his book *The Mystery of Consciousness* [44] and was not able to find it in the bookshops of Palo Alto and surroundings areas. It was, in fact, easier to find it in bookshops in Brazil, translated into Portuguese. In Brazil this book even helped me to solve a question with the police. I had this book in my bag when I was flying from Rio de Janeiro to Johannesburg, South Africa, to attend a big school and conference organized by Val Goranko. When checking in at the airport the police was convinced that I was a drug trafficker because, as they explained to me later, I was well dressed and arrived at the last minute at the check-in. This was a typi-

[19] The expression "philosophical logic" is also used, cf. in particular *The Journal of Philosophical Logic*. Its meaning is not necessarily clear.

Fig. 5 With Arthur Buchsbaum and Newton da Costa at the 2nd World Congress on Paraconsistency in Juquehy, Brazil, 2000

cal behaviour of traffickers, who were numerous for this destination, Rio de Janeiro being a huge cocaine export hub in South America and South Africa one of the most important consumers on the market. When they asked me what I was going to do in South Africa and saw the book of Searle, they just let me go away without an advanced search.

During this period in Rio I reinforced my contact with Brazilian colleagues, on the one hand, the group of Campinas, on the other hand, the group of Fortaleza. Campinas is traditionally the most important centre for logic in Brazil. There is there, indeed, a Centre for Logic, it is called Centre for Logic, Epistemology and History of Science, *CLE = Centro de Lógica, Epistemologia e História da Ciência*, but it is mainly logic. Up to now they have escaped the danger of being transformed into a Centre of History and Philosophy of Science, as did happen with the LMPS congresses launched by Tarski, where logic has became marginal. CLE has, in fact, been visited by Tarski in the 1970s. This visit was organized by Itala D'Ottaviano, one of the two pillars of the CLE for nearly 30 years together with Walter Carnielli, both being ex PhD students of da Costa. This is a very active group with many visitors and publications, which is connected to the Brazilian Society of Logic (SBL).

After the 1st World Congress of Paraconsistency (WCP1) organized in Ghent, Belgium, in 1997, CLE logicians planned to organize WCP2 in 1999 to commemorate the 70th birthday of Newton da Costa. With João Marcos, a master student of Walter at the time, we travelled by car during 3 days on the coast of the state of São Paulo to find a good location. We found a nice hotel on a nice beach, named Juquehy (meaning "Rain Dance" in Puppy Guarana). WCP2 was held in 2000 (there was a delay) and it was a great success; the participants enjoyed it very much – the best WCP so far (Fig. 5). I became good friends with João Marcos, who is interested in many things and, like Walter, fluently speaks about 10 languages. I only speak three language fluently. I think it is nice to know lots of languages, it helps to open our minds. This was the idea of Schopenhauer, whose favorite language was Spanish – I think Portuguese, especially from Brazil, is better. However, I am not sure that knowing a lot of languages is a sufficient or necessary condition for good communication. Most of the time it is beyond words, and João Marcos and Walter sometimes have difficulties understanding each other despite being able to speak 10 languages.

My contact with Fortaleza was established through Arthur Buchsbaum. He did his PhD on paraconsistent and paracomplete logics. We were in touch via e-mail when I was in Poland. When I arrived in Rio he was working in Fortaleza, Ceara, and came to Rio for the defense of his PhD at PUC. This was the first time I met him and we became good friends. His advisor was Tarcisio Pequeno from Fortaleza. After spending many years in Rio working at IBM and then the Department of Informatics of PUC-Rio, the best in Brazil, Tarcisio came back to Ceara and started to develop a group of logic and artificial intelligence (LIA). He was interested in a mixture of paraconsistent logic and non-monotonic logic.

At this time, Tarcisio invited me to take part in his research group, and I started to go regularly to Fortaleza, about 2000 km north of Rio. As a good friend of Tarcisio and his colleagues and students (Ana Teresa, Marcelino, Ricardo) I enjoyed the climate and culture very much; it was quite different from Rio. Besides logic he had a strong interest in philosophy and with Vanderveken he organized a congress with Searle in Fortaleza. We had many deep discussions. What I learned with Tarcisio, and also other colleagues in Brazil, is to work all the time with uncertainty and flexibility and I think this is good.

4.5 Stanford, USA and Two Trips Around the World

I arrived in Stanford, California in January 2000, where I would stay 2 years. My situation in Rio was instable. Half of the time I was without income. I decided to apply for an advanced grant of the Swiss National Science Foundation to go somewhere. The idea of going to Stanford appeared naturally. I was interested in going to California for a longer period. There was a good connection between the group of Patrick Suppes at Stanford and the group of da Costa, in particular through the Chilean logician Rolando Chuaqui and, more recently, with Francisco Doria from Rio who had been visiting Pat, and his former student Acacio de Barros who was on location working with Pat. One idea I had about Suppes was his relation with Tarski. In 1957 together with Tarski and Henkin he organized a big meeting on axiomatic methods at Berkeley (cf. [25]), which was the first step for the launching of the LMPS (logic, methodology and philosophy of science) series of congresses, the first having been organized in 1960 at Stanford by Suppes himself. When I arrived at Stanford Suppes was mainly working on the brain and I would work with him on this topic but he let me develop my research as I liked. When I was at Stanford John Etchemendy, philosophical logician, became the Provost of the university, a function he still occupies. Sol Feferman and Grigori Mints were the leading logicians. Johan van Benthem came every year and he was becoming a strong person for the development of logic there. Richard Zach was finishing his PhD at Berkeley and was lecturer at Stanford.

Again my financial situation was quite a disaster because this was a time of a booming economy in Silicon Valley and the Swiss franc, as other currencies, was losing ground against the American dollar – although nowadays the Swiss franc is higher than the US dollar. I had to rent very expensive rooms (nearly 1000 USD a month) in some houses of "poor" people[20], and I was going around the "farm", as Stanford is known, by mountain bike, and from time to time going to San Francisco, taking my bike with me on the train. This was the time of the rising of Google, founded by the two Stanford students, Larry Page and Sergey Brin, recalling the exploits of the legendary Stanford duo Bill Hewlett and Dave Packard, who during my stay gave to Stanford the highest donation in the history of donations. Maybe in the future Page and Brin will surpass this record. A few years earlier, Bill Gates had also donated a large amount of money to the Computer Science Department of Stanford. Besides money the farm also boasted nice racoons, several statues by Rodin (a version of *The Thinker* and others), and a memorial installation about Eadweard Muybridge, author of the "motion picture" that proved Leland Stanford's contention, that all four hooves of a horse are off the ground at the same time while trotting).

When I was at Stanford I also travelled quite a lot taking part in congresses. During the summer of 2000 I took part in two big events of the Association for Symbolic Logic – the ASL wanted to do something special for the year 2000. They organized one event in Champaign, USA and one in Paris, France. That June in Champaign I saw Saunders MacLane for the first and only time. The ASL congress in Paris in July was much bigger; it took place at the Sorbonne and was the biggest congress in the history of logic up to now. I presented talks about universal logic at both events. In August I went to the ESSLLI in Birmingham (12th European Summer School in Logic, Language, and Information). I had been selected to give a course on paraconsistent logic. That was nice; among the students of that course I counted Alessio Moretti and Catarina Dutilh-Novaes.

The following year, 2001, for the first time I went to a meeting of the SEP (*Society of Exact Philosophy*), an annual meeting held alternately in Canada and the USA – this time it was in Montréal. I gave a talk titled "Sentence, proposition and identity", work I was doing in connection with Suppes's idea about congruence. Identity is a notion in which I have always been interested. It is a notion where typically logical philosophy is important. One has to know what a congruence relation is. On the other hand, one needs to have a good philosophical interpretation of this notion. Suppes has articulated the notion of congruence to study the relation between sentences and propositions, an articulation that can be used to defend the notion of proposition against Quine's rejection of it. During my stay at Stanford Quine died and there was a 1-day workshop with talks by Suppes, Mints, Feferman, etc. Pat's was quite provocative, sinking Quine in his grave: by an analysis of the bibliography of *Word and Object*, he showed that Quine knew quite nothing about the recent advances and experiments about language and psychology when he wrote that book, relying mainly on what his Harvard colleague and friend B. F. Skinner told him.

In May 2001 I went to Russia for the first time in my life for the Smirnov meeting to which I had been invited – a biennial logic meeting in Moscow. Mike Dunn, Paul Weingartner and Diderik Batens were also there. The event was organized in particular

[20] At some point I was renting a room in a house in Menlo Park, and the woman asked if I would take a shower in the morning or in the evening. My reply was both and her reply that this was not possible.

Fig. 6 With Mike Dunn, Alexander Karpenko and Vladimir Vasyukov at the Academy of Science in Moscow in 2001

by Vladimir Vasyukov, whom I knew from Poland. His father had been a Russian army officer and he had grown up in Poland. Another organizer was Alexander Karpenko, the director of the chair of logic of the Russian Academy of Sciences in Moscow; he was also a poet, with whom I became good friends (Fig. 6). After that I went to Russia several times, for some conferences in Moscow and St Petersburg, and I always enjoyed it.

In June 2001, together with Darko Sarenac who was doing his PhD at Stanford, I organized a workshop on paraconsistent logic in Las Vegas. In July 2001 I did my first trip around the world. It was a round trip following the direction of the rotation of the earth in the Southern Hemisphere: from San Francisco to Australia, then South Africa, then Brazil, then back to San Francisco at the end of August. In Australia I went to Hobart, Tasmania, for the Australasian Philosophical meeting. There I took part in a workshop on logical pluralism organized by JC Beall and Greg Restall. Then I gave a lecture at a 1-day workshop in Melbourne where I met Lloyd Humberstone. Next, I went to the far west of Australia, to Perth, where Slater worked, and I gave my first talk about the square of opposition, in relation with his paper "Paraconsistent logics?" claiming that there are no paraconsistent logics. From Perth I flew to Johannesburg, where Goranko still worked at Rand Afrikaans University and gave a talk on universal logic. Then in Brazil I took part in events in Fortaleza and Florianópolis. In particular, I gave a mini-course on paraconsistent logic at the Annual Brazilian Meeting on Computer Science.

The Swiss National Science Foundation (SNF) grant that I had been receiving at Stanford was due to end in July 2001, before my first trip around the world. Pat Suppes asked me what my future plans were and offered to engage me for a couple of more months at Stanford to work with him. This seemed good to me and I gave him the positive reply to work with him up to the end of November 2001. In the autumn of 2001 I met Pat nearly every day for discussions; this led to our joint paper "Semantic computation of truth based on associations already Learned" (published in 2004). Therein we have a preliminary philosophical and theoretical basis to explain how the brain work. At his "Brain Lab" Suppes conducted lots of experiments but he was not a naive empiricist and he knew that data make sense only through a theory. What I remember about all these experiments is that, surprisingly, a triangle as a picture, as a written word, and as a spoken word provokes the same effect on the brain.

My plan after Stanford was to go to Switzerland. There was a new program of the SNF, a grant of several years, a mix of research and teaching, for people who had had a PhD for a couple of years but did not yet have a permanent position. The SNF program was designed to prepare a wave of retirements of professors, avoiding the loss of good researchers and brain draining due to unemployment, providing a smooth transition between research and the permanent position of professor. The idea was to locate the winners of these grants in some Swiss universities, where there was a possibility for them to get a job later. My plan was to go to the University of Neuchâtel where there was an Institute of Logic. The director of this institute, Denis Miéville, had written to me, suggesting that I apply. I therefore applied for this grant at the beginning of 2001 and was selected for an interview at the beginning of January 2002 in Bern. My idea was to stop in Bern while doing a second quasi-complete around the world trip, earth rotation: leaving San Francisco for Brazil stopping in New Zealand, Asia, India and Europe. My intention was to go to Brazil, wait for the result, stay in Brazil if the result was negative and apply again the following year. The competition was very tough. About 30 people were selected every year in all Switzerland and all fields. I knew it would not be that easy to apply with a project in logic.

A the end of November I left California. I first stopped at the Australasian logical meeting in Wellington. From Wellington I went by train to Auckland. I did this trip of a whole day's duration with Stephen Read. In Auckland I took part in a philosophical meeting and then flew to Singapore. From there I went to Thailand, crossing Malaysia by bus, and then flying from Bangkok to Katmandu, Nepal. It was shortly after the Nepalese royal massacre, when the King and seven members of the royal family had been killed by his son, and in the countryside some Maoists were trying to get the power, attacking and murdering people. I nevertheless took a drive in the direction of Himalaya and visited Bakthapur, which I liked very much. From Katmandu I flew to Varanasi, India. It was my first visit to this country. I did a trip by boat on the Ganges and I went to visit the nearby city of Sarnath, where the Buddha made his first speech. I then went by train to Chennai and Mumbai.

From Mumbai I had a night flight to Zurich with a stop in Dubai on Emirates Airlines. My flight was to arrive in the morning, shortly before the interview in Bern. I tried several times to change this flight with the travel agency that had booked my around the world trip in San Francisco, but there was no way to change it. I went to Emirates's office in Mumbai, also without success. My flight was at 3:00 in the morning. Before going to the airport I was training in my hotel room in Mumbai for the interview. I was quite exhausted when boarding the plane after queueing and confusion at this big airport. Fortunately the atmosphere on board was pretty relaxing, Emirates at this time had been elected number one airline in the world. We landed at Dubai airport early in the morning, where I had to change planes. At this airport I bought a music tape to which I listened a lot during the coming months, especially when flying. It was by a woman singer whom I did not know yet, the Lebanese Nawal Al Zoghbi, and this was her seventh and most successful album, *El Layali* (The Nights), with a song of the same name, my favorite one. On arrival in Zurich I had to run to catch my train to Bern. Once on the train I fell asleep and nearly missed disembarking at Bern. After leaving my luggage in a locker at the station, I ran to the SNF building. I arrived only 2 or 3 minutes before the interview. I asked if it would be possible to grab a quick coffee. They told me that there was no time to do so, that the jury was waiting for me. I succeeded to rightly conduct the presentation of my project.

After that I took the train again to go to Neuchâtel, where I would sleep and meet my potential collaborators the next day. This is what I did. I remember that there was a lot of snow and that it was cold; this was my first time in Neuchâtel and I quite liked it. After that I went back to Brazil and waited for the reply. It came sooner than expected, at the beginning of February. My contract would start August 1st in Switzerland, I had time to relax a bit. I took this opportunity to visit Machu Pichu and went on the Inca Trail. I also crossed the region of the Titicaca lake up to La Paz in Bolivia – a wonderful trip. In April 2002 I organized a workshop with Décio Krause in Florianópolis[21] for the 80th birthday of Pat Suppes, who was visiting Brazil on his way to Chile.

5 Launching Projects from Switzerland (2002–2008)

5.1 By the Transparency and Depth of a Swiss Lake

I left Brazil at the end of July 2002. I spent a few days in Rio de Janeiro with Arthur Buchsbaum at his brother's house at the foot of the "Cristo redentor" – Christ the redeemer – and participated in the 8th WoLLIC (Workshop on Logic, Language, Information and Computation) that took place at PUC-Rio. From Rio I flew to Italy to take part in the 14th ESSLLI (European Summer School in Logic, Language and Information) in Trento in which João Marcos was organizing a workshop on paraconsistent logic, together with Diderik Batens (the big smurf of logic in Belgium). This was giving continuity to the 1st and 2nd World Congresses on Paraconsistent Logic (WCP1, Ghent, Belgium 1997; WCP2, Juquehy, Brazil, 2000).

At this time I started discussions with Walter Carnielli about the organisation of WCP3 in Toulouse at IRIT for 2003 (the idea being to hold WCP every 3 years). IRIT (*Institut de Recherche en Informatique de Toulouse*) is a major research centre in France, with 700 researchers and a good diversity of lines of research, including strong logic trends, in particular with the two fuzzy guys, Henri Prade and Didier Dubois (one of the main editors of the *Journal Fuzzy Sets and Systems*) of the ADRIA team (*Argumentation, Décision, Raisonnement, Incertitude et Apprentissage*), and with the LILaC team (*Logique, Interaction, Langue, et Calcul*) with people like Andreas Herzig, Philippe Besnard, Philippe Balbiani, Jonathan Ben-Naim and Luis Fariñas del Cerro. Luis was at that time the gen-

[21] Florianópolis is located on a very nice island in the state of Santa Catarina, south of Brazil. This is where my friend Arthur Buchsbaum moved after living for a couple of years in Fortaleza, because it was too hot for him. This is also the place where Newton da Costa moved with his family after his retirement from USP. Saint Exupéry used to stop on this island on his way to Chile. The hut where he stayed still exists near a small village whose central street is called "Avenida Pequeno Príncipe".

eral director of IRIT and gave us full support for organizing this event. Walter had been in touch with Luis and IRIT for a couple of years. IRIT had taken over the *Journal of Non-Classical Logics*, founded by Walter and others in Campinas, which on this occasion had been renamed as *Journal of Applied Non-Classical Logics*.

From Trento I went to Neuchâtel to start a new life. This would be the first time after I left Paris in 1991 that I would really have my own house and the first time I would have a non-precarious position. This was a 4-year contract with the Swiss National Science Foundation (SNF), with a possible 2-year extension, and at the end perhaps a permanent position. In Neuchâtel I was attached to the *Institute of Logic and Centre of Semiological Research* founded by Jean-Blaise Grize. Grize was retired when I arrived and the institute was directed by his former PhD student, Denis Miéville, who was mainly interested in the work of the Polish Logician Stanislaw Leśniewski, the PhD advisor of Alfred Tarksi. It was through this Polish connection that I arrived in Neuchâtel. Miéville wanted to develop his institute, having heard about the SNF program, reading some of my papers and knowing I had been working in Poland, he invited me to come there. When I arrived in Neuchâtel, Miéville was the Rector of the University of Neuchâtel, as had been also Jean-Blaise Grize between 1975–1979. Grize (1922–2013) was a former student of Piaget. The father of Jean Piaget, Arthur Piaget, (1865–1952) was the first Rector of the University of Neuchâtel (1909–1911). Jean Piaget (1896–1980) was born in Neuchâtel and his first research study was on the mollusks of the lake. He was then professor in Paris and founder of a research centre in Geneva, *Geneva International Centre of Genetic Epistemology*. Grize worked with him in Geneva and then settled down at the University of Neuchâtel.

Neuchâtel is a beautiful city in Switzerland, situated by a lake with transparent water from the mountains and thus good inspiration for philosophy. Schopenhauer wrote: “The real philosopher always looks for limpidity and precision, he will invariably try to resemble not a turbid, impetuous torrent, but instead a Swiss lake which by its calmness preserves transparency despite its great depth, a great depth revealing itself precisely through its great transparency.” (*On the fourfold root of the principle of sufficient reason*) [46]. Precision is famous in Switzerland through the watch industry. Neuchâtel is a few kilometers from Watch Valley, where the most famous watches in the world are made, at the border between Switzerland and France in two small towns in the mountains: Le Locle and La Chaux-de-Fonds. The latter is also famous because it is the town of Le Corbusier and, less known, of Louis Chevrolet (Chevrolet’s logo is a cross, remembering the Swiss flag). In La Chaux-de-Fonds the correspondence between Bertrand Russell and Louis Couturat [22], which was edited and published by Anne-Françoise Schmid [45], had been found recently.[23] A very interesting book where the who logicians talk about many different subjects.

[22] Louis Couturat (1868–1914) promoted the work of Leibniz on logic and completely unknown for several centuries. Couturat is the author of *La Logique de Leibniz* [8] and also *L’Algèbre de la logique* [10], *Histoire de la langue universelle* [9] and many other interesting books.

[23] Anne-Françoise is a good friend of mine who invited me to take part to a joint project jointly organized with Nicole Mathieu, “Modélisation et interdisciplinarité”. For 5 years (2007–2012) we invited, listened to and questioned 12 great French intellectuals in Paris. I have a strong memory of two of them: Maurice Godelier and Radyadour K. Zeytounian. This resulted in a book – I was responsible for the section on logic and linguistics featuring Jean-Pierre Desclés and Patrick Blackburn, see my paper [98].

The region of Neuchâtel is also where the legendary absinthe beverage was conceived at the end of the 18th century. The nickname of this beverage is the "green fairy" and it was very popular among artists in the 19th century. Van Gogh was a great consumer and it is suspected that the colours of his paintings are connected with the psychic effects of this beverage, which are considered hallucinogenic. At some point absinthe was outlawed in all countries – in Switzerland from October 7, 1910 until March 1, 2005. So when I arrived in Neuchâtel it was still illegal, but everybody had a clandestine bottle at home. After its re-legalisation (following some strict rules that lower the effects), it became a popular drink in bars and restaurants of the city. Neuchâtel, like many places in Switzerland, is an interesting mixture of provincial and cosmopolitan atmospheres. Switzerland is one of the countries in the world with the highest percentage of foreigners, in particular because of the strong traditional asylum policy of the country.

The basis of Switzerland's banking system was laid by French protestants seeking asylum in Switzerland from persecution – 16th to 18th centuries. When I lived in Neuchâtel the recent political refugees were mostly Albanians. There are also people settling in Switzerland for business due to the very stable political and economical situation. In Neuchâtel there is one of the biggest Philip Morris's factories in the world, which was the first factory of Philip Morris outside the USA. When I arrived in Neuchâtel there was the Swiss national exposition – nearly 40 years after the preceding one in 1964 in Lausanne – and it was really animated. I found a big flat to rent, with a nice view of the lake and the mountains and used to invite colleagues and friends for fondue parties.

When I arrived in Neuchâtel, Miéville was quite busy with his activities as rector, but he had meetings with his assistants and students almost every Monday morning. This was a friendly group. There was a diploma of logic "licence de logique" which disappeared through the Bologna process of "uniformisation"[24] of studies in Europe. At the beginning of October the annual meeting of the Swiss Society for Logic and Philosophy of Science was organized in Neuchâtel. The topic for 2002 was quantification. This was the opportunity to present the work of my PhD on logical truth. I presented this work again, in more detail, at a small workshop that we had in the winter of 2003 at a monastery near Lausanne. It led to my CQFD 2005 paper "Le Château de la Quantification et ses Fantômes Démasqués". This was directly connected with the work of my colleagues in Neuchâtel on Leśniewski's systems. Leśniewski had the idea that a variable can range over one, many or zero objects, contrary to the main trend, up to now dominating, where a variable range only over single objects. I told them that I was not interested into empty domains, because *if there is nothing, there is nothing to say*. For me that perfectly justifies the standard model theory developed by Tarski according to which the domain of a structure is always considered to be non-empty. As a consequence, "for all" implies "there exists", which is better formulated as "at least one", avoiding existential and scholastic problems.

I know the work of Fraïsse, French logician friend of Tarski, his tentative to develop a framework for model theory with empty domains, but I think philosophically it is ambiguous for the above reason. Although I like Fraïsse's idea of *zerology* (cf. [15]), considering that we can understand things through limit cases, I also think that there is a danger

[24] Giving the same form, unifying the multiplicity into one single form. There is no exact English translation of this word.

in focussing on exceptions. Obsession for exceptions and monstrous cases in philosophy can lead to sophistry. At some point in our discussions Pat Suppes emphasized that, saying that what is important first of all is the general situation. I agree with him: if you focus on awkward cases you will create an awkward theory. Of course, in mathematics one counter-example is enough to kill a theorem, but there is a difference between a theorem and a conceptual framework.

This question of empty domain and quantification is also related to the square of opposition, a topic that became one of my main topics during the beginning of my stay in Neuchâtel.[25] I already discussed the square of opposition in relation with paraconsistent logic visiting Slater in Perth in July 2001. I discussed it again at the 14th ESSLLI in Trento, leading to my 2005 paper "Parconsistent logic from a modal viewpoint". However, in my first months in Neuchâtel I made important progress, which led to the development of my research project on the square of opposition, where paraconsistent logic is only one aspect. Part of this work can be found in my 2003 paper "New light on the square of opposition and its nameless corner". The two central ideas that I had at this time are concerned with the extension of the theory of the square of opposition in two ways: to the third and other higher dimensions (polyhedra), to go to polytomy (this does or does not lead to polyhedra).

Going to the third dimension by itself was not a new idea, because people had already the idea to generalize the square of opposition to a cube of oppositions. However, I presented a different polyhedra than a cube, to which I was led by wanting to relate three hexagons of opposition dealing with interactions between modalities and negations. On the one hand, in doing that I used hexagons, which are two-dimensional generalisations of the square of opposition. On the other hand, my construction of polyhedra of oppositions was motivated by an intrinsic necessity and was not the product of an abstract generalisation in the air. Generalisation and abstraction are two very interesting features of mathematics and thought in general. To generalize is interesting, but most of the time it is trivial. There are some people who want to generalize everything all the time; this is quite superficial. Most of the time these people are not capable of studying the details of a particular thing, because this requires care and attention. The same can be said about abstraction, the two phenomena going hand to hand. To jump into abstraction, to fly in the sky of generalities allowing us to have a general vision of the phenomena, is a very difficult task to perform. If we fail, we sink into an ocean of trivialities or just splash in a swimming pool with artificial plastic tools. This difficulty manifests itself regarding the question of polytomy of oppositions.

The theory of oppositions started with dichotomy promoted by Pythagoras and Plato. It was generalized into a trichotomy by Aristotle introducing the notions of contrariety and indetermination. Blanché's generalisation of the square of opposition strongly emphasized trichotomy. His hexagon is obtained by putting together two triangles [26]. Kant in his logic book [27] argues that only dichotomy is *a priori*, that all other polytomies are empirical. However, Blanché's hexagon indeed shows that he is wrong. One may decide

[25] I will not present the details of this study here. I have described it in my 2012 overview paper "The new rising of the square of opposition".
[26] About this hexagon, see Blanché's masterpiece [4] and my 2012 paper "The power of the hexagon".

to generalize dichotomies of opposition to any kind of polytomies; this can easily be done mathematically. One can generalize, for example, the hexagon of Blanché based on trichotomy to an octagon of oppositions based on the mix of two quatritomic figures (one of contrariety, one of subcontrariety) looking like the octagons found in Islamic culture. The question is whether or not this is philosophically interesting. Similar discussions can be developed about many-valued logic. I do not necessarily agree with Quine saying that many-valued logic is not logic but algebra, but I think that it is important when working with many-valued logic to keep in mind the philosophical aspect. Truth and falsity are represented in modern logic by 1 and 0. This does not necessarily mean that all natural numbers or real numbers can be considered as truth values.

Back in the old world I was eager to develop contacts with people of different countries in Europe. It started first with France and Italy. At the 14th ESSLLI in Trento, which was attended also by Alessio Moretti a French-Italian guy living in Nice. It was the second time I met him after the 12th ESSLLI in Birmingham in 2000. We had good contact because, like me, he was interested in many things, including Smurfs. He was a philosopher with an open mind and a good knowledge of both analytic and continental philosophy. Moreover, he also had a strong interest in logic and mathematics. He lived in Nice and was doing a PhD under the direction of Jean-François Mattéi (1941–2014), a specialist in Pythagoras and Plato – he wrote the excellent book on Plato, *L'Etranger et le Simulacre* [33], which I read when working on Plato's cave.

At the end of August 2002 there was the ASPLF congress organized by Mattéi in Nice, and Alessio invited me to stay at his house. I went to Nice by plane from Geneva, a short (less than 1 h) and very beautiful flight, one of my favorites. There is a nice view of the lake and Mont Blanc on take-off in Geneva, and then you go south seeing the Alps up to the Mediterranean sea, with the plane flying above the sea before landing at Nice's airport. I like the region of Nice very much; it is geographically similar to Corsica and culturally it is an atmosphere that is a mix of France, Italy, England and Russia. The ASPLF congress was nice. I met a lot of new people. ASPLF means *Association des Sociétés de Philosophie de Langue Française*. It is an organisation that gathers French-speaking philosophers from all around the world and there have been international congresses regularly since 1938 – in Nice this was the 29th congress. This was the start of a long collaboration with Alessio. Later, I presented him with my idea about the square of opposition, which became his favorite topic up to now and he did a PhD with me (2005–2009) on the *Geometry of Logical Opposition* [34] at the University of Neuchâtel.

In October I went to Nice again, on my way to Sardinia. I had organized a series of talks on my new ideas on the square of opposition. From Nice I took a boat to Corsica – my first time in Corsica since I had left in 1975. I arrived in Ajaccio, went south to Bonifacio, going through Colomba's village near Sartène. I spent one night in Bonifacio and the next day took a boat to Sardignia and then crossed the whole island to arrive in Cagliari where I gave my first talk of this tour – I had contact with Francesco Paoli, whom I had met in Tasmania in 2001 and who worked there. From Cagliari I took a boat to Naples. The arrival in Naples by boat is very impressive, with a view of Vesuvius. In Naples I had contact with Nicola Grana, a long-time friend of Newton da Costa, who worked on paraconsistent logic and the theory of valuations (see [18]). I gave two talks and in 2 days Nicola showed me everything in Naples, a city I enjoyed very much. It is a bit like Brazil (there are many Italian descendants in Brazil, especially in São Paulo, and most of them are from Naples). I also visited Pompeii. From Naples I went to Rome, my first visit to this impressive city, and then go on up to Siena, where my friend Claudio Pizzi was, an Italian modal logician who has been adopted by the Brazilian community, and again gave a talk on the square. From Siena I went back to Nice and then Geneva.

I also made strong connections with a group of people in Nancy, where I met Fabien Schang and Katarzyna Gan. From some time we developed an N3-linkage (N3 = Nice-Neuchâtel-Nancy). In Nancy a group of logic and philosophy of science was developed by Gerhard Heizmann, in particular through the Archives Poincaré – the famous mathematician Henri Poincaré was from Nancy. This culminated with the organisation of the 14th LMPS in 2011. Fabien was at this time a PhD student of Heizmann, who later on invited me to take part in Fabien's defense. Heizmann was very talented in organizing events, in particular emphasizing very friendly gathering around food and drinks. He certainly inspired me for my later organisational activities. In the autumn of 2002 I took part in an event he was organizing named PILM = Philosophical Insights into Logic and Mathematics: The History and Outcome of Alternative Semantics and Syntax. Many important logicians were there, in particular van Benthem and Hintikka. I presented work I had been doing since Stanford, showing that in most of the cases possible worlds are superfluous for relational semantics.[27] The corresponding paper was subsequently published in the nice book resulting from PILM: *The Age of Alternative Logics* (see [3]). Later, there was also a congress on Polish logic in Nancy, organized by Roger Pouivet, a French philosopher married to a Polish woman, who translated the famous book on the principle of contradiction in Aristotle by Łukasiewicz into French.

On this occasion I met Katarzyna, a Polish student, who was going back and forth between Nancy and Poznań. This was the beginning of a long-lasting friendship and collaboration. I went to Poland in 2003 to take part in an event that Jan Zygmunt organizes every year in Karpacz, a nice village in the mountains in the south of Silesia, bringing with me two colleagues from Neuchâtel.

In November 2002 a position for a logician at the University of Lausanne was announced. I decided to apply, although I was not very motivated because I had just started my 4-year SNF fellowship, which in many senses was more interesting for me for the coming years. This position was in particular organized in the perspective of the development of the *Centre Romand de Logique et Philosophie de la Science*, a research centre to develop activities in logic and philosophy of sciences between Geneva, Lausanne and Neuchâtel.[28] I was selected with two other peoples (one of them was Marcus Kracht) for a lecture and an interview. I remember that during the interview, since I had nothing to lose, I clearly told the committee that I was not interested in teaching logic just by presenting truth tables, translation exercises and so on. The committee did not retain any of us. They decided to summon other candidates among those not selected for the first

[27] This is related to my paper "Possible worlds: a fashionable non-sense?", which I wrote when at Stanford, stressing the fact that most of the time people do not know what they are talking about when they talk about possible worlds – this paper has not yet been published, it was rejected several times in the actual world. Late in 2010 I published a paper entitled "What is a possible world?"

[28] The creation of this centre, an idea due to Jean-Claude Pont, was aborted due to discord between people.

round, among them Achille Varzi from Columbia, who was chosen after the interviews of this second round. But Achille gave up. People suspected that he applied a very common strategy in the USA: since the salaries in Switzerland are very high, and his salary at Columbia was low, in the end he got an increase of his salary to stay in New York, which he likes very much, after a free trip to visit Switzerland for the interview.

In the second place, the committee had listed Jacques Duparc, who therefore got the job. Jacques was my former friend from the Lycée Gabriel Fauré in Annecy. After entering the math sup "prépa" he quit because he liked mathematics very much and this was too much oriented towards engineering. For some reason he then entered the faculty of medicine and then quit again to do philosophy. After a couple of years I met him by chance in Paris and I told him I was studying logic and explained to him all about it and the logic cursus at Paris 1 and Paris 7. He had been working on Heidegger and Wittgenstein but then started to study logic and liked it very much, especially set theory, getting very good results and ending up defending a logic PhD at Paris 7, after a stay at UCLA. At the time of the open position in Lausanne he was teaching logic in Germany.

Miéville was sad not to have Varzi because Varzi had been interested in the mereology of Leśniewski. At this stage I think Miéville did not support me because in any case I would be around at least 4 years and the idea was to enlarge the logic community with one more person. On the other hand, the president of the committee was a mathematician, Dominique Arlettaz, from the University of Lausanne (now Rector of this university), with whom I had previously had a peculiar issue. In 1999 before going to Stanford, there was a position of logic in Lausanne, a kind of tenure track position. I applied and was chosen by Dominique Arlettaz. Since I also got the grant to go to Stanford, I decided not to go to Lausanne. I explained this to Arlettaz but he insisted very much on me taking this position, saying that I was exactly the kind of person they were looking for. If I had taken this position in 1999, I probably would have been nominated full Professor of Logic in 2003 at the University of Lausanne, the time this position was attributed to Duparc.

I do not regret anything. It was not easy to choose but I think this was the right choice, both going to Stanford in 2000 and not making specific efforts to get the position in 2003. Before the interview in Lausanne I clearly told Arlettaz in a personal meeting when he was visiting the University of Neuchâtel to give a talk that I was not interested in the position right now, but he told me they could not wait.[29]

With my SNF project I had a 4-year grant for a student doing a PhD with me. In Switzerland there was no one in the position to do a PhD in logic. I decided to invite a Brazilian student, Alexandre Costa-Leite, whom I had met a couple of years perviously at a congress in Florianópolis. At the time he was an undergraduate student at the Federal University of Goias, near Brasilia, the region he is from. We were keeping in touch and he was asking me for advice. I advised him in particular to do his Master's at UNICAMP. This what he did: a Master's thesis on the Fitch paradox, directed by Walter Carnielli. He

[29] I have always thought it strange that at universities everywhere in the world whole life positions are assigned in a few weeks. Also, most of the time the procedure is highly dubious due to the way the jury is constituted. The ones who choose are not rightly chosen and/or self chosen, so on this basis how can the final choice be good? André Weil stated an interesting law of university hiring: first-rate people hire other first-rate people. Second-rate people hire third-rate people. Third-rate people hire fifth-rate people, . . .

had just finished, so that was good timing. He was glad to be invited to Switzerland, he adapted quickly and easily and enjoyed Neuchâtel very much and also the nearby cities of Bienne and Fribourg. He also traveled around, benefiting from the Erasmus program to do research visits of a couple of months each in Amsterdam, Paris and Lisbon during his PhD. The title of his PhD is *Interactions of Metaphysical and Epistemic Concepts*. This is a work in logical philosophy related with combination of logic. The defense was in 2007 with Paul Gochet, Pascal Engel and Arnold Koslow as members of the jury.[30] Alexandre had a strong interest in the combination of logic; we both went to CombLog'2004 in Lisbon, organized by Amilcar Sernadas and Walter Carnielli at IST (*Instituto Superior Técnico*), and the day after his defense we organized the workshop CombLog'07 in Neuchâtel. At CombLog'2004 I presented the "copulation paradox": the fact that when combining conjunction and disjunction, we get more: distributivity. Gabbay liked the expression. Later I developed this work further with Marcelo Coniglio, who was also present at this event among many others: Razvan Diaconescu, Jospeh Goguen, Carlos Caleiro, Don Pigozzi, etc. There was a really nice atmosphere in Lisbon surrounded by Fado and Bifadas.

I attended CombLog'2004 on my way back to Switzerland after a third around the world trip, this time in reverse rotation. It was the shortest, less than 30 days, and the craziest around the world trip I have done. From Geneva I went to St. Petersburg, then Hong-Kong, Bali, Australia, Tahiti, Easter Island, Santiago de Chile, Buenos Aires, São Paulo, Lisbon and back to Geneva. In Australia I took part to the Australian Congress of Philosophy on South Molle Island in the Great Barrier Reef Marine Park. I presented the talk "Three definitions of human beings", comparing Plato, Aristotle and Desmond Morris's definitions, respectively, bipeds without feathers, rational animals and naked apes. In his famous essay on the axiomatic method [35], which strongly inspired Alfred Tarski, Blaise Pascal emphasises the trilogy axioms/definitions/demonstrations. Nevertheless, Pascal claims that there are some notions that are useless to define and gives as an example "human beings". On the contrary, I think that this notion is one of the most interesting to define. I am working on this subject and have presented several time talks on it but have not yet written a paper. Besides this around the world trip I also went to series of events I had already taken part in before: LMPS, WCP, SEP. The 12th LMPS took place in Oviedo, Spain, the 21st WCP in Istanbul, Turkey and the 31st SEP in Montréal (all in 2003). I also kept in touch with Brazil and South and Central America, taking part in several meetings there, in particular going to the 12th SLALM (Latin-American Symposium on Mathematical Logic) in San Jose, Costa Rica in January 2004. I enjoyed Costa Rica very much, visiting both the Pacific and the Caribbean sides.

In Geneva in October 2004 I took part in the first congress on Louis Rougier (1889–1982) organized by Jean-Claude Pont. Before this event, like many people, I knew nothing about Rougier, who was a very important French philosopher of the 20th century, a friend of Schlick and promoter of the Vienna Circle. He married Lucy Friedman, Schlick's sec-

[30] Alexandre was my first PhD student. Alessio Moretti, the second, also defending his PhD in Neuchâtel. Tzu-Keng Fu from Taiwan started a PhD with me in Switzerland but after my departure he finished it with Oliver Kutz at the University of Bremen in Germany, the title is *Universal Logic and the Geography of Thought*.

retary, and adopted her daughter from a previous marriage – this woman came to the meeting in Geneva. Rougier has written many interesting books and was very active. The reason why he remained unknown is due to his constant support of Marshal Pétain during and after the Second World War. For this and other reasons nobody wanted to talk to or about him. Rougier was neither pro-nazi, nor antisemitic but he was anti-communist and also anti-democrat. He defended neo-liberalism, a position that was not welcome among the intellectuals in France after WWII. He was also anti-rationalist and anti-Christian. Although I disagree with most of his ideas, I think he is certainly one of the most brilliant thinkers of the period. His criticism of rationalism is quite interesting, his way of questioning basic principles considered as obvious by rationalists, indeed not obvious at all and sometimes wrong. Rougier was the first to strongly defend relativism in logic. He had a good knowledge of what was happening in logic and the appearance of different non-classical logics. He was a conventionalist and had the idea that there are different logics for different situations. He wrote the following: "Avec la logique, l'homme a brûlé sa dernière idole" (with logic, man has burnt his last idol). I later wrote the paper "Rougier: logique et métaphysique" (2011), a critical presentation of his main ideas, and asked Mathieu Marion, a good specialist on Rougier, to write an introduction on Rougier's essay "On the relativity of logic" for the *Anthology of Universal Logic* I edited in 2012. (A good friend of mine, Louis Allix, surprisingly told me that Rougier was one of his uncles.)

In 2004 I also took part in an interesting meeting: the *Third Philosophy Day*, at UNESCO in Paris, on November 18. I was invited to take part in a panel on globalisation organized by my friend Francisco Naishtat. On this occasion I presented the idea of a world university. At the end of the day there was a very nice music show by Herbie Hancock, after which the happy few were invited to a cocktail party on the roof of the UNESCO building (UNESCO has good location in Paris, near Napoleon's gravestone at *Invalides*). Herbie had a nice pendant with coloured feathers and I asked him what it was. He told me that was a gift from Carlos Castaneda.

5.2 *Universal Logic's Take-Off*

In the autumn of 2004 I started to organise the 1st UNILOG, 1st World Congress and School on Universal Logic projected to take place in April 2005 (Fig. 7). In October 2003 I had organized a small workshop on universal logic with Alexandre Costa-Leite, with the participation of Arthur Buchsbaum, who was visiting me in Switzerland at the time, and also Petrucio, Paulo and Sheila Velso, Darko Sarenac, Jacques Duparc, Ramon Jansana and a few others. However, organizing a big event was a completely different story. This was my first experience of doing that and I worked very hard to be sure it would be a success. It was important to find a good location and good timing. I chose the city of Montreux, which is beautifully situated on the banks of Lake Geneva, and for the timing I thought that Easter would be nice because many important events are organized in the summer and in Spring around Montreux there is still a lot of snow, which is quite nice. During the event we organized a trip to the Marmot paradise which is at 2000 m above Montreux; we went there by the typical cog railway and this was a very pleasant

Fig. 7 Poster of the 1st UNILOG

promenade at the middle of the snow. I decided to organize a combination of a school and a congress. For the congress I had the idea to have a contest and a secret speaker (a speaker whose identity is revealed only at the time of his talk). This format was successful and was repeated at the following editions of the event. Alexandre continuously helped me to organize this event and suggested to have Kripke, one of his favorite philosophers, as the secret speaker. We had no idea if he would accept, but he did. He liked the universal logic project and had never been to Switzerland, he wanted to get to know this country. We succeeded in organizing his coming without much trouble and it was a success (Fig. 8). Indeed, the whole event was a great success for various reasons. The hotel was a charming family hotel. Montreux is really relaxing. We had lots of famous logicians and also a very good book exhibition. Some Chinese logicians came and suggested organizing the second UNILOG in China and we started to discuss this possibility in Montreux.

The book *Logica Universalis – Towards a General Theory of logic*, published by Birkhäuser, a collection of papers related to universal logic was launched at the 1st UNILOG. In November 2004 I entered into contact with Birkhäuser to propose to them to launch a new journal: *Logica Universalis*. Birkhäuser is the publisher of the journal *Algebra Universalis* and many other important mathematical journals and books, in particular the collected papers of Alfred Tarski. For me that was the natural editor for a journal on universal logic. They very quickly replied to me. Birkhäuser is based in Basel, about 90 min by car or train from Neuchâtel. Two representatives visited me in the next days and I also went to Basel. They told me that launching a journal is a very serious enterprise, that it has to be done carefully so that the journal will not disappear in a few years. They thought a book series would be safer. Anyway, we agreed to produce as a kind of experiment a book with the title *Logica Universalis*. Thomas Hempfling, now the present director of Birkhäuser, came to the 1st UNILOG in Montreux to talk with the people, to feel if the atmosphere was propitious for a new journal of logic. Birkhäuser agreed to

Fig. 8 With Saul Kripke on the roof of Hotel Helvetie

sponsor the contest by offering a 500-Swiss franc book coupon to the winner. In the end all was positive and Birkhäuser decided to launch the journal *Logica Universalis* and the book series *Studies in Universal Logic*. This was carefully prepared and both started in 2007. Up to know I have managed the editing of this journal and the book series – very interesting work. *Logica Universalis* rapidly became one of the most important journals of logic and in the book series we have published many interesting titles, the latest being a book on Leon Henkin: *The Life and Work of Leon Henkin* [25].

After the great success of the 1st UNILOG – many people told me it has been the best conference of their life – I was quite optimistic about my future. But success rhymes with difficulties. It is like mountain climbing: the higher we go, the more beautiful view we have, the more risks we face, and the more likely it becomes that a lethal fall may happen. In September 2005 I was traveling in the West part of USA. I gave a talk at the Santa Fe Institute in New Mexico, an interesting interdisciplinary institute where a friend of mine, David Krakauer, was working and visited the nearby artistic town Taos where Julia Roberts lives. I did not go there to meet her, but I was interested to visit Taos because this is where the novel of Fredric Brown, *The Far Cry*, takes place. At the Sante Fe Institute I met the writer Eric McCormack who dedicated me one of his books. He is living in Santa Fe and had some interest for logic. In California I visited Pat Suppes at Stanford and drove through Los Angeles by the highway number 1 through legendary places like Carmel and Big Sur, a trip I had no opportunity to do when at UCLA and Stanford since I had no car at this time. I visited Herb Enderton at UCLA and gave a talk there at the cheese and wine seminar.

After my talk at UCLA I was relaxing in Ojai, a nice town in the mountains near LA, on my way back to San Francisco. I then received some surprising news: the Rector of the University of Neuchâtel told me it would not be possible for me to have my SNF grant at the University of Neuchâtel prolonged. I had been working at this university for 3 years, I had 1 year left but already had to ask for the 2-year prolongation at the SNF.

Such a prolongation would, in principle, be nearly automatic, since the idea of the SNF was to avoid brain drains, keeping the good researchers in Switzerland until they found a permanent job, I had not yet found one. I had to ask for a prolongation by writing a report of what I had done during the 3 years, and I had done a lot of things, that was not the problem. The problem was that I needed the formal agreement of the university that was hosting me.

Since I arrived in Neuchâtel the rector had changed. It was no longer Denis Miéville. Miéville had become quite unpopular and he had to leave his position. Since the direction of the university was confused, the state of Neuchâtel, who finances the university, decided to change the way to choose the rector. They decided that the rector will be chosen by the state (the same situation, indeed, as in France and Brazil), not by an internal decision of the university. They decided to choose Alfred Strohmeier, a computer scientist, former director of the big Department of Informatics of the Federal Polytechnical School of Lausanne (EPFL). The e-mail I received when in Ojai was from Strohmeier, but he told me that it was not possible for him as a rector to give me his approval for the prolongation of my grant since Miéville was against it and the approval of the rector has to be in conformity with the decision of the host institute. What was happening was that Miéville did not want me to stay for two intertwined reasons. Due to his failure as a rector he was quite bitter, he wanted to come back to his activities of logician, and from this point of view the success of my activities was overshadowing. However, it was difficult for him to tell me that he did not want me to stay, so he tried to make me believe that he had supported my prolongation of the SNF grant but that Strohmeier was against it, because he was against him as a former rector. However, the real state of affairs was revealed.[31]

Faced this intricate situation I tried to find a solution. Strohmeier told me he would support my prolongation if another institute of the University of Neuchâtel would host me. I talked with Jean-Jacques Aubert, new Dean of the Faculty of Letters and Human Sciences. He wanted me to stay and did his best to support me. I ended up at the Institute of Psychology, directed by Anne-Nelly Perret-Clermont, who was in connection with Piaget school. Because I had the invitation of this institute, Strohmeier gave me his approval and received a positive reply from the SNF for a 2-year prolongation of my project. I move to her institute at the end of my 4-year period, i.e. in August 2006. I worked there for 2 years and it was very nice. Students of psychology were quite interesting and I took the opportunity to give courses on various topics, in particular on imagination.

During this period I organized an interdisciplinary congress on imagination at the University of Neuchâtel. I really like interdisciplinarity and Piaget, a symbolic figure of Neuchâtel, who was also a promoter of interdisciplinarity; he coined the word "transdisciplinarity", which he thought was better. So I thought that the circumstances were good for organizing interdisciplinary events. In 2005, still at the Institute of Logic, I organized the First Interdisciplinary Congress on the Symbol – the main topic of which was a critical appraisal of Ferdinand de Saussure's claim that arbitrary signs (by opposition to symbols that are considered as non-arbitrary signs) are very important. My idea of such an interdisciplinary event was to try to gather colleagues of one university from as different fields as possible. The University of Neuchâtel was good for that because all the main fields that

[31] When Miéville retired in 2012, the Institute of Logic of the University of Neuchâtel was closed.

can be found in a university are there, and the university is small so that it is not so difficult to get in touch with everybody. The University of Neuchâtel has different buildings in different places in the town, but the town is also small, everything is at a walking distance. I had an office at the Department of Mathematics in the Faculty of Science. From this faculty, I succeeded to convince not only a mathematical colleague (Alain Robert, the author of a very good book on non-standard analysis [38]) but also a chemist, a geologist and a physicist to join the event. Colleagues from the Faculty of Theology (Lytta Basset), the Department of Linguistics (Louis de Saussure), the Department of Philosophy (Daniel Schulthess), etc., joined. The event was quite a success. I also invited a couple of colleagues from outside, in particular Claudine Tiercelin, who gave a talk on Peirce. This gave birth to the book *La Pointure du Symbole*. By doing this congress and book I discovered many things, in particular the excellent works on signs by the typeface designer Adrian Frutiger (see [17]). For the book I received additional papers by Robert Dewar on traffic signs, Elisabeth Nemeth on Neurath, and Jeremy Narby on life signs.

The second event was easier to organize because I had already organized one and the theme imagination was naturally attractive. For this second event I invited Catherine Chantilly, my present wife, to film the event. We had become acquainted through an artistic event she has been organizing every year since 2005 with Alessio. They took place in castles in the Bourbonnais, the centre of France, her region of origin, which is full of castles. I took part in this event in 2005, 2006 and 2007, and enjoyed it very much. In 2005 I had a photo exhibition there about the *Do Not Enter* traffic sign (in French: *Sens Interdit*. In 2003 I started to work seriously on photography and decided to start a parallel career in photography, which I found as interesting as filming. My idea to work on the *Do Not Enter* traffic sign suddenly appeared to me in Moscow in 2003 when I took the first photo of such a sign. My interest in this sign was connected with symbolism. I presented this exhibition in several locations. In French the name of the exhibition was *Le Monde en Sens Interdit*, in English *Do not Enter in the World*. This is a funny name for this exhibition that I had in Chengdu, China, while I was on a trip in 2006 to prepare the organisation of the 2nd UNILOG (Fig. 9).

After my visit in the winter of 2006 in China, we decided to organize the 2nd UNILOG in Xi'an, the ancient capital of China, rather than in Beijing where the 23rd LMPS was projected. We scheduled UNILOG for August 16–22, 2007, just after the LMPS meeting so that participants of LMPS could join us and know more about China. Xi'an is a city which is, indeed, much more interesting than Beijing. It is more authentic with a strong cultural variety and the famous *Terracota Warriors*. Huacan He the main Chinese organizer was working at The Northwestern Polytechnical University in Xi'an at the time. The 2nd UNILOG was a nice event, which allowed people to know more about China.

Before the 2nd UNILOG I organized the First World Congress on the Square of Opposition in Montreux, June 1–3, 2007. I was not in favour of again organizing UNILOG in Montreux because I thought it better to have this event circulating around the world. However, since the conditions in Montreux were very good, I thought it would be nice to take the opportunity to organize a second event there. The SQUARE was easier to organise because it was smaller (no school). Yet, it was nonetheless quite challenging. In particular I wanted to develop interdisciplinarity. I thought the square was a good basis for that because it is a very simple theory that everybody can understand. I also wanted

Fig. 9 Do not Enter in the World, Chengdu, China, February 2006

to have an artistic part. Since we were in Montreux, music naturally came to my mind. I asked Michael Frauchiger to help me with that. Michael is Swiss and married a Brazilian girl, the Manager of the Lauener Foundation.

Henri Lauener (1933–2002) was a Swiss philosopher who before his death gave some money to organize a prize, which is awarded on a regular basis. I was at the first award ceremony in Bern in 2004 – the prize was awarded to Pat Suppes – and I noticed that there was an important music part, jazz music. Michael told me that it was so, because Lauener liked music very much. Michael asked jazz musicians to prepare pieces of music based on the square of opposition and a show for our event in Montreux and we got a nice result. For this event we also produced a movie, a remake of the biblical story of Salomé, where we used the square to articulate the relations between the four main characters of the story. Our version of this story is quite different from the original and Flaubert's and Oscar Wilde's versions. The movie was shot in Morocco in April 2007 with Catherine acting as Herodias, Alessio as Herod, Joana Medeiros, a Brazilian actress, as Salomé and myself as St John the Baptist. The four of us filmed and directed the movie at the same time, it was quite an interesting experience. Another event I organized when in Switzerland was a workshop on possible worlds with Saul Kripke in Neuchâtel in June 2008. I had invited Kripke for the SQUARE in Montreux but at the last moment he had to cancel his trip due to health problems, but I was able to reschedule his air ticket so that he could come to Switzerland in the near future.

During the period 2005–2008, besides organizing seven events (two UNILOGs, two interdisciplinary congresses in Neuchâtel, the first SQUARE, CombLog'07 and the workshop with Kripke), I also took part in several events and did some lecture tours in USA

Fig. 10 With Alexandre Costa-Leite and Arnold Koslow, New York, April 2008

and India. In Europe I took part in *The Impact of Categories – 60 Years of Category Theory in Historical and Philosophical Retrospect*, October 10–14, 2005 (organized by Andrei Rodin at ENS in Paris, France); *Paradox: Logical, Cognitive and Communicative Aspects* November 4–6, 2005 (Organized by Jurgis Skilters in Riga, Latvia); *Applications of Algebra to Logic and Informatics X*, March 6–12, 2006 (Organized by Joanna Grygiel in Zakopane, Poland); *the 15th Annual Meeting of the European Society for Philosophy and Psychology* July 9–12, 2007 (organized by Kevin Mulligan in Geneva, Switzerland); *Identity and Structure*, December 7–8, 2007 (Organized by Karin Verelest in Brussels, Belgium); *International Workshop on Truth Values*, March 29–June 2, 2008 (Organized by Heinrich Wansing in Dresden, Germany); *ECAP08 – 6th European Conference on Computing and Philosophy*, June 16–18, 2008 (Organized by Jean Sallantin in Montpellier, France).

In Central and South America I took part in the 13th SLALM (Latin American Symposium on Mathematical Logic), August 7–12, 2006, in Oaxaca, Mexico and in the 14th SLALM, May 11–17, 2008, in Paraty, Brazil. This was my first time in Mexico and I really enjoyed Oaxaca and Puerto Escondido on the nearby coast. In Oaxaca I presented a talk on absolute maximality, a concept that I introduced. The idea is the following: Emil Post proved the maximality of classical propositional logic (CPL) in the sense that the connectives of this logic cannot be strengthened, but it is known that we can extend CPL by adding non-definable connectives like modal operators. CPL is post-maximal, but it is not "absolutely maximal". A logic that would be absolutely maximal is a logic to which no

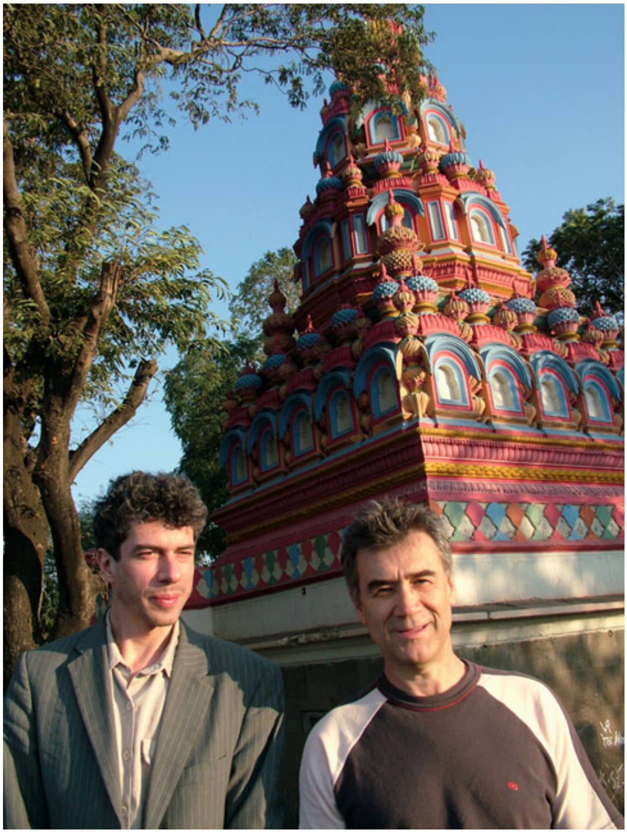


Fig. 11 With Walter Carnielli in Poona, India, December 2005

additional connectives can be added without trivialisation, i.e., transforming this logic into a trivial logic in which anything is derivable from anything. Up to now I have not written yet a paper on that topic. After SLALM in Mexico I flew to Madrid in 2006 to attend the World Congress of Mathematics. On this occasion the Field medals were attributed to Terence Tao, Grigori Perelman, Andrei Okounkov and Wendelin Werner. I presented a poster on universal logic in the logic section and won the prize for posters in this section. I attended an interesting lecture by Leo Corry about Hilbert and Einstein. After Madrid I went to Budapest for the 31st Meeting of ASPLF, the topic of which was *Le Même et l'Autre* where I took part in a panel on identity and logic. In July 2008 there was the 4th World Congress on Paraconsistency (5 years after WCP3 in Toulouse) in Melbourne, Australia.

In 2008 I also did a tour of lectures in the East of USA, going from New York to Buffalo and back. Alexandre was doing a post-doc at CUNY with Arnold Koslow (Fig. 10). I also went for 1 month on a mission of the Swiss Secretary of Education and Research in India. I gave talks in Kanpur, Mumbai (where my friend Raja Natarajan works at the Tata Institute of Fundamental Research), Chennai and stopped on the way to Varanasi, one of my favourite towns in India. This was my third visit to India after the first one on the way to my SNF interview in Bern in 2001, and a second at a congress of artificial intelligence in Pune in December 2005 (Fig. 11). After the congress in Pune I drove about 3000 km up to Goa then Hyderabad and back to Mumbai, spending about 2 weeks with Peter Arndt walking and discussing on Goa's beaches.

6 Back to the Land of the Future (2008–2014)

6.1 Honeylips Beach

My contract with SNF was ending July 31st 2008 and I did not know exactly what I would do in the near future. In Switzerland there was no job. At some point in 2006, I started to make some formal administration to be able to be a candidate in France. In this country, if you want to apply for a job at a university you must first be "qualified" by a national committee. This is a pre-selection that will tell you if you can apply for positions as assistant professor or full professor, or for no positions at all. It is not easy and and you have to apply in some specific sections. I applied to the philosophy section and mathematical section for full professorship and was qualified for both. However, there were no interesting positions open, so I did not apply. It is, in fact, not easy to find a position that corresponds to what we are doing and I never wanted to apply to a permanent one in something outside of my main research interest, i.e. logic.

I then received an invitation from my friend Tarcisio Pequeno for a 3-year research position to work in Fortaleza, Brazil (Fig. 12). I think this was a nice proposal that would allow me to go on with my research and apply for some jobs. Moreover, I like the state of Ceara, of which Fortaleza is the capital. Due to some delay typical of Brazilian bureaucracy – in particular there was a problem in getting the visa to work there – I started to work officially in Fortaleza only on October 1st, 2008. I had given up renting my flat in Neuchâtel and with Catherine I spent some days in a camping near Neuchâtel, discovering another aspect of Switzerland that I had not imagined: ordinary people living on campsite. I did not succeed getting my visa at the Brazilian consulate in Geneva. At the beginning of September we decided to go Portugal; this was the first step of 2000 km in the direction of Brazil, and I had some contacts at the Brazilian consulate in Lisbon. I succeeded in getting my visa there but I had to wait a couple of days. We decided to go to the Algarve, South of Portugal.

At this time I started to develop my project of an anthology of universal logic. The idea was to prepare a collection of logical papers of the 20th century in the spirit of universal logic, with each paper commented by a specialist. I chose 15 papers and asked colleagues to work on them. The book was finally released in 2012. The full title is *Universal Logic: An Anthology – From Paul Herz to Dov Gabbay*. There was some delay due to copyright issues. The preface of about 10 pages is a general presentation of what universal logic is, explaining the choice of these 15 papers. Among them, there are five papers that were translated into English for the first time: two that were originally in German by Paul Hertz

Fig. 12 With Tarcisio Pequeno in Guaramiranga, Ceara, Brazil

and Paul Bernays and three that were originally in French by Alfred Tarski, Haskell Curry and Jean Porte.

We finally moved to Fortaleza. There I started to work again in paraconsistent logic and also took part in the international congress commemorating the 80th birthday of Newton da Costa in Campinas in September 2009. The work on paraconsistent logic that I developed in Ceara is philosophical. I had always been unsatisfied with the existing philosophical discussion surrounding paraconsistent logic, ambiguous praise or/and apology of contradictions, remembering Mao Tse-tung's 1937 essay, *On contradiction*. I now have a clear idea of a serious philosophical approach that can justify paraconsistent logic. I developed these ideas when I was in Fortaleza but up to now I have not had any time to write this down, it will be developed in a projected paper entitled: "Cats, tigers and stones". The idea is that tigers are cats that are not cats. Big Cats, as we say in English, are different from small domestic cats, which are the real cats. This is a typical example interestingly expressed here by a scriptural variation of a capital letter, a "differance" à la Derrida. For most concepts we can make a difference between "real things" and border cases. A general perspective of such border cases corresponding to "to be and not to be" can be found through representation. A typical symbolic example is Magritte's 1919 painting, *The Treachery of Images*, known also through what is written on it "Ceci n'est pas une pipe" (This is not a pipe): we have a pipe which is not a pipe.

This approach combines well with the dual notion, that of paracomplete negation, which is better known and more accepted. The idea is that, given a concept C, it is possible for something to be neither C, nor non-C, because it is out of the scope of the concept. For example, we can say that animals such as wolves, snakes, birds, etc., are not cats – but does it make sense to say that a stone is not a cat? The same with many notions such as the number 4, a cell phone, the Soviet Union, etc., which are all in an incomplete zone dual of the inconsistent zone. The idea is to construct a logic with a negation which is both paraconsistent and paracomplete. Up to now it is not clear what a good mathematical construction for that is.[32]

While in Fortaleza I went on to work on universal logic. I “crystalized”[33] the similarity between Garrettt Birkhoff’s approach to universal algebra and universal logic by the expression “axiomatic emptiness”, and also developed the notion of anti-classical logic as a typical example and motivation for axiomatic emptiness. I presented a talk about that at the 6th Smirnov’s Readings in June 2009 in Moscow. I also worked hard to prepare the 3rd UNILOG in Portugal. The event took place between April, 18–25, 2010 in Estoril, near Lisbon. Carlos Caleiro and his colleagues from IST in Lisbon did a great job, which was nearly annihilated by the eruption of Eyjafjallajökull in Iceland. The eruption started on April 14 and led to the strongest ever paralysation of air traffic in Europe. We were afraid that we would have to cancel the event, to which we were expecting about 300 participants. We succeeded, nevertheless, to organize the event with about 200 participants. On the one hand, the eruption stopped and, on the other hand, many people succeeded to come in one way or another. Some people came by train, bus or car.

Among them, the famous proof-theorist Gerhard Jaeger, who was one of our keynote speakers. Facing this situation he decided to come by car from Bern, which is a driving distance of about 2200 km – to compare: the distance between New York and Miami is about 2100 km. I did not see his car, but as a German I guess he has a BMW, Mercedes or Porsche, a comfortable and speedy car; nevertheless in France and Spain the speed on the highway is limited, a situation which is different from Germany. Moreover, the eruption did not affect direct flights from Brazil or USA to Portugal. Hintikka was able to join as a secret speaker. It was nice to have him at UNILOG despite the fact that his project is clearly opposed to universal logic, since he believes, like Ross Brady and Huacan He, in the idea of a universal system of logic, which according to him is its own system IF.

During this period I also kept working on the square of opposition, in particular on developing further applications. In October 2009 in Buenos Aires I presented applications of the square and the hexagon to economy. I emphasized that, like in many other situations, the theory of opposition behind the square permits us to refine our thinking and theories, going beyond dichotomies. I discussed the trichotomies buy-sell-rent and save-spend-invest. In November 2009 I went to Corsica to prepare the Second World Congress on the Square of Opposition, which took place June 17–20 in 2010 at the University of Corsica Pasquale Paoli in Corte. This was probably the first time that a real international congress was organized there with people from all over the world who had never been to Corsica before, in particular Brazilians, like my friends Arthur Buchsbaum and Luiz Carlos Pereira. They were very impressed by Corsica, and I was happy to introduce them

[32] The expressions “paracomplete negation” and “paracomplete logic” were introduced by Newton da Costa, by analogy to the expression “paraconsistent negation” and “paraconsistent logic” suggested to him by Miró Quesada. Quesada suggested to him “non-alethic” for something which is both “paraconsistent” and “paracomplete”. I have suggested “paranormal” for a good para–llelism.

[33] This is a Frenchism: it means that various ideas or thoughts nicely take a specific form.

to the place where I grew up. There are many tourists in Corsica during the summer, but they are mostly from Europe.

Corte is a small town in the middle of the mountains and the atmosphere of the congress was very relaxing. Michel Barat, the Rector of the Academy of Corsica, came and delivered an introductory speech. It was not just a few formal words but a whole lecture. He is a philosopher and also a freemason who has been the great master of the Grande Loge de France (GLDF). We tried to organize an artistic counterpart like at the first congress in Montreux but we just ended up with some Corsican students singing some traditional Corsican polyphonic music that was not inspired by the square. During this stay I discovered by chance a Corsican band with a Portuguese/Spanish name – "A Primavera" (the Spring) – mixing Corsican music and music from the Andes, which I liked very much. This revealed to me a strong connection between Corsica and Peru.

During this time in Fortaleza I also worked with Catherine on the project of a movie on living philosophers. My original idea was to make interviews of the most important living philosophers (between 10 to 15 people). Catherine likes to work on films and joined me on this project. In Fortaleza she was contracted by the FUNCAP (Cearense Science Foundation) to do such work. We had already started this before arriving in Ceara, filming in particular Jaakko Hintikka in Paraty. By developing the project, we changed the general setting and decided to rather produce a series of movie by countries, a project that can be called *World Philosophy* or/and *Philosophy in the World*. There are different reasons for proceeding this way. One is that it is not at all clear who the 10 most important living philosophers are, if there are any. The other reason is to present less known philosophers from different countries and cultures. The project is developing slowly but I think at the end it will be nice.

During this period in Fortaleza we produced a first movie about France, filming Alain Badiou, François Laruelle and Jean-Luc Marion in their private homes during our visits in Paris. In 2010 we went on a 1-month trip to Canada, from West to East, filming John Woods in Vancouver and Thomas De Koninck in Quebec City. De Koninck is known to have inspired Saint Exupéry for *The Little Prince*. The aviator was hosted by De Koninck family when Thomas was a child and he was asking a lot of questions. Later, Thomas De Koninck became Professor of Philosophy at the University of Laval in Quebec City. Saint Exupéry is certainly very philosophical. I think the case of De Koninck perfectly justifies the new direction of our project. De Koninck is not one of the 10 most important philosophers in the world, but he is an important philosopher in Canada, certainly not as well known abroad as he should be.

When I was in Fortaleza, a position opened up at the Federal University of Rio de Janeiro in 2009. After reflection I decided to apply seriously. I was 44 years old and up to this age had not had a permanent position. I knew that waiting more would make it more difficult each time, but I was in some sense ready to face the challenge. In my life I always have refused to choose a job for comfort. The reason I decided to apply to this position in Rio is related to three criteria:

- Nice place to live
- Good position in a good university
- Interesting intellectual environment.

Rio is one of best towns to live: it is geographically spectacular, a big international town, one of the most famous cities in the world, in the middle of the sea, the forest and the mountains.[34] I had been working in different departments, but my preference was to be a professor of logic in a philosophy department. This was exactly the position that was offered, at the University of Brazil, a.k.a. the Federal University of Rio de Janeiro, or UFRJ, the oldest and largest university in Brazil, which is considered one of the top three universities in Brazil, and in all of Central and South America, along with USP in São Paulo and UNICAMP in Campinas. The intellectual environment in Brazil is good for me because this is one of the countries in the world where logic is the most developed, with a variety of people from mathematical, philosophical and computer science backgrounds. There is a Centre of Logic in Campinas (CLE), a Brazilian Society of Logic (SBL) and a regular Brazilian Congress of Logic (EBL). I, therefore, applied for this position and was selected in May 2009, but it took more than 1 year to be contracted and to start my work in Rio due to some bureaucratic difficulties. Going back to Fortaleza from Rio de Janeiro after the interview I had a serious problem swimming in the sea. In Fortaleza, we lived in a flat by the sea, in "Praia de Iracema" (Honeylips Beach – Iracema is a legendary Indian girl from the Tabajara tribe), with a 180-degree panoramic view. It was like we were at the middle of the sea where dolphins were bathing. I went on swimming and was driven by strong waves to some sharp rocks to which I tried to grip on, but which cut my skin. I succeeded to get back to the beach bleeding. I still have some (not so dramatic) scars.

6.2 Flowing in the River of January

I started to work at the University of Brazil – UFRJ – in Rio de Janeiro in August 2010, the second semester of the academic year. In Brazil the academic year is divided into two semesters. The first one starts at the end of summer, i.e. after Carnival, sometime in February, and goes up to early July, then there is a break of about 1 month, and then the second semester goes until December. With a permanent position I started to give some regular courses. In the Department of Philosophy of UFRJ there is an introduction to logic class which is mandatory for undergraduate students and several optional classes of logic.

We are supposed to give two classes per semester. Generally, the first semester I give "Introduction to Logic" and in the second semester an optional class which is a follow up. Additionally, there is an optional class on logic and/or a graduate seminar each semester. I really like giving the class on the introduction to logic. For me it is important to give a very general vision of logic not limiting this class to some truth-tables, basic first-order logic and translation of sentences of natural language into these logical systems. In this perspective I wrote a paper "Logic is not logic" giving a general idea of what is logic and emphasizing the difference between reasoning and the theory of reasoning that I have expressed by the a capital variation, one more "differance". I also orient Master's and PhD theses. Brazilian students are friendly, they have a variety of interests, are open to new

[34] For me the ten top cities in the world are: Rio de Janeiro, Geneva, Paris, St Petersburg, Rome, Lisbon, San Francisco, London, Montréal and Athens.

Fig. 13 With Peter Arndt in Regensburg, Germany, January 2012

ideas and are eager to learn. A couple of years ago the Brazilian government has launched a very nice project called "Scientific Initiation" to encourage research already among undergraduate students. The students can receive a grant to develop a research project, and in all universities in Brazil there is an annual week of 'Scientific Initiation", where students present their work and are evaluated by a jury. At Brazilian universities research is also strongly encouraged among professors, who are not considered just teachers. The Brazilian research council (CNPq) gives grants to this end.

Since 2010 I have been taking part in many conferences and have organized many events. I was invited to different events, in particular: the Journées Alain Badiou in Paris, the 70th birthday congress of Istvan Németi in Budapest, the 60th birthday congress of Arnon Avron in Tel Aviv and the Vasiliev memorial congress in Moscow. In the winter of 2012 I took a tour of Germany, supported by the Humboldt Foundation, and gave more than ten lectures throughout the whole country and also neighboring countries like Liechtenstein and Poland. I gave talks to philosophy departments, to mathematics departments and to computer science departments. I met some old friends (like André Fuhrmann in Frankfurt and Peter Arndt in Regensburg – Fig. 13) and made new ones – I was particularly happy to give a talk in Darmstadt, a place related to Ernst Schröder.

Before going back to Brazil I went to Lebanon to prepare the 3rd World Congress on the Square of Opposition at the American University of Beirut (AUB). I had contact there with Ray Brassier, at the time Director of the Department of Philosophy of AUB (where David Makinson had worked previously) and Wafic Sabra, the Director of the Center for

Fig. 14 With Catherine Chantilly and João Ricardo Moderno, Rio de Janeiro in 2013

Advanced Mathematical Sciences of AUB, who had worked with David Bohm in London. The event in Beirut was a great success. We received the support of the embassies of Switzerland, France, Italy and Brazil. The Swiss Ambassador in Lebanon, Ruth Flint, kindly offered a very nice Swiss cheese and wine cocktail party for the opening of the event and the Brazilian Ambassador, Paulo Roberto Campos Tarrisse da Fontoura, also offered a cocktail party for all participants at the Brazilian cultural center in Beirut[35]). He was happy to learn that logic was an important field of research in Brazil.

In the autumn of 2012 I started to organize the 4th UNILOG, scheduled to take place March–April, 2013, in Rio de Janeiro. This was an important step for the development of UNILOG. It was the most successful of all UNILOGs, with nearly 500 participants, 50 invited speakers, many tutorials, including one about logic for the blind given the by late Laurence Goldstein. And of course there was also a contest and a secret speaker (the late Grisha Mints – he was not able to travel, but he presented an online lecture). For the organisation we benefited from the full support and dedication of Katarzyna and her husband Przemysław Krzywoszyński from Adam Mickiewicz University. The school started with a general discussion about reasons for studying logic and ended with a discussion about publication, to encourage young researchers to publish. There was also a good book exhibition. The event was organized in a marvelous location at the foot of the Sugar Loaf. We succeed in organizing it in a strategic military school there – ECEME – thanks to the president of the *Brazilian Academy of Philosophy* (ABF), João Ricardo Moderno (Fig. 14).

I was elected titular member of this Academy in 2012 and shortly after I was named Director of International Relations of ABF. In this position I presented the candidacy of Rio de Janeiro to host the World Congress of Philosophy. I did that during the summer of 2013 at the 23rd World Congress of Philosophy in Athens. I was supported in particular by Itala d'Ottaviano, representing the IBF (*Brazilian Institute of Philosophy*). We lost to Beijing. The Chinese had already put in a bid once before and lost, so this was their turn. During the assembly the members admitted that Rio would be nice… but next time. The WCP take place in 2018 in Beijing (the frequency of WPC is once every 5 years) and

[35] In Brazil there are many people from Lebanon and Syria. One of my Brazilian colleagues, Fabio Tfouni, came to this square event with his parents and met relatives living in Syria.

then most probably in 2023 in Rio. During the 23rd WCP in Athens, I also discussed with the Honorary President of ASPLF (*Association des Sociétés de Philosophie de Langue Française*), Jean Ferrari, who suggested that I organize a congress of ASPLF in Rio. At this time the next ASPLF was already scheduled for August 2014 in Rabat, capital of the Kingdom of Morocco. I went there to present the candidacy of Rio and I also presented a plenary talk there, the topic being "Le possible et l'impossible". At the assembly of ASPLF in Rabat we discussed the candidacy of Rio de Janeiro and the idea is to organize the congress of ASPLF in Rio in 2018.

While in Rio I launched, on the one hand, a new series of events, an annual 2-day workshop at the Sorbonne called *Logic in Question/La Logique en Question*, and on the other hand, a seminar of logic in Rio de Janeiro, *Logica Carioca*, trying to gather logicians spread over the various universities in Rio (there are about ten universities in this city). I also launched four new editorial projects: *Cadernos de Lógica e Filosofia*, *SAJL = South American Journal of Logic*, *Logic PhDs* and the *Encyclopaedia of Logic*. *Cadernos* is a book series published by College Publication, a publishing house of Dov Gabbay. The idea is to translate some important works related to logic for undergraduate students and/or a wide audience.

We will soon launch the translation of *The Game of Logic* by Lewis Carroll, a collection of papers by Peirce and the translation of Tarski's classical book *Introduction to Logic and the Methodology of Deductive Sciences*. The idea of *SAJL* is to promote interaction between logicians of all kinds in South America and to make their works known in the rest of the world. This is a joint project with Marcelo Coniglio. We are preparing a special issue dedicated to the Argentinian logician, Carlos Alchourrón. *Logic PhDs* is another series with College Publication, were PhDs of famous logicians will be published.

The PhD of Haskell Curry presented by Jonathan Seldin (Curry wrote his PhD in German and was the last student of Hilbert) and the "classical" PhD of Krister Segerbeg on modal logic presented by Patrick Blackburn will be published soon. We will also publish the PhD of Saunders MacLane (presented by Peter Arndt), which was on logic, as few people know, and was also written in German and defended at Göttingen. In this series we will additionally publish the best recent PhDs. The *Encylopaedia of Logic* is a joint project with IEP = *Internet Encyclopedia of Philosophy*. Since August 2013 I have been the logic area editor of IEP and with the support of the general editors of this web Encyclopedia, James Fiser and Bradley Dowden, and the general editor of College Publication, Dov Gabbay, I also decided to produce a printed encyclopedic logic book.

In February 2014, together with Mihir Chakraborty, I organized the 5th World Congress on Paraconsistency at the Indian Statistical Institute in Kolkata, in India. People in India are starting to have a strong interest in paraconsistent logic. We decided to organize three tutorials preceding the congress so as to give a better understanding of what paraconsistent logic is. In May 2014, the 4th SQUARE took place at the Pontifical Laterian University (PUL) in the Vatican. Raffaela Giovagnoli, professor at PUL, had made the proposal to organize the event there during the 3rd SQUARE in Beirut, and we received the full support of Gianfranco Basti, Dean of the Faculty of Philosophy of PUL. Moreover, the Bishop Enrico dal Covol, Rector of the Pontifical Lateran University, gave an introductory lecture (in French) at the beginning of the event. Katarzyna Gan succeeded in organizing a cocktail party at the Polish Embassy in Vatican with the Ambassador Piotr Nowina-Konopka, and Juliette Lemaire (with the support of Anne Hénault) organized a cocktail party at the

Embassy of France in the Vatican with the Ambassador Bruno Joubert – this Embassy is located in Villa Pauline Bonaparte, named after the sister of Napoleon.

During the world soccer cup in June 2014 our university in Rio was closed, so I took the opportunity to make an extended trip to Europe through the Marie Curie exchange program GeTFun (Generalizing Truth-Functionality) set up by Carlos Caleiro and João Marcos. I did my first trip to Romania, where I visited my friend and colleague, also a GeTFunian, who lives in the very nice village of Sinaia. This is where the second International Mathematical Olympiad was organized in 1960, the first also having been organized in Romania, in 1959 in Brasov. Razvan won the Romanian Mathematical Olympiad when he was young and is now researcher at Simion Stoilow Institute of Mathematics of the Romanian Academy (IMAR) in Bucharest where I gave a talk and had the opportunity to meet an old friend of Grigori Moisil. I went to the Vienna Summer of Logic where we had the workshop GeTFun 2.0 and then to Greece where I was keynote speaker at a workshop on Logic and Utopia on Andros Island organized by Petros Stefaneas and Thalia Magioglou. I spent a few days in Athens, in particular to develop a good synergy between UNILOG'2015 which is scheduled for June 20–30, 2015 in Istanbul and the 10th Panhellenic Logic Symposium scheduled June 11–15, 2015 on Samos Island, the Greek Island where Pythagoras was born.

7 Projects

7.1 Research and Publications

I will go on working on universal logic, the square of opposition and non-classical logics, and also on various philosophical topics. To get an idea of what I plan to do, the reader may look at my projected future papers and books that can be found in my list of writings at the end of this chapter.

7.2 Organisation of Events

I think that workshops and congresses are very important. These are where people meet, discuss ideas, and plant the seeds of future collaborations. I have taken part in many events, and I constantly work to promote these kinds of exchanges.

To date I have launched three series of congresses:

- UNILOG *World Congress and School on Universal Logic* – a biennial or triennial peripatetic event – Montreux 2005, Xi'an 2007, Lisbon 2010, Rio de Janeiro 2013, Istanbul 2015
- SQUARE *World Congress on the Square of Opposition* – a biennial or triennial peripatetic event – Montreux 2007, Corsica 2010, Beirut 2012, Vatican 2014, Easter Island 2016
- LIQ *Logic in Question / La Logique en Question* – an annual 2-day workshop in the spring at the Sorbonne, Paris, France

Fig. 15 Square of Opposition event in Rapa Nui, aka Easter Island planned for 2016

The idea of UNILOG is to have a school followed by a congress. Moreover, during the congress there is a contest and a secret speaker. UNILOG is an event that promotes logic in all its aspects: philosophical, mathematical, computational, historical, and more. For the next gathering, in Istanbul in 2015, we will have three types of tutorials: tutorials about the history of logic (logic according to Aristotle, Leibniz, Kant, etc.), tutorials connecting logic with other topics/fields (logic and music, logic and information, logic and colours, etc.), and tutorials presenting an important theorem (incompleteness, completeness, compactness, etc.) in a universal perspective, i.e. examining the logical basis of this theorem. The contest will be about the future of logic and the secret speaker is still secret.

The idea of SQUARE is to have a really interdisciplinary event centred on a simple but rich and fruitful logical theory, the theory of opposition, also known as "the square of opposition". The expression "square of opposition" may refer to a particular stage of this theory, in particular Boethius's formulation of it with an explicit square diagram. However, the theory of opposition started with Pythagoras, Plato and Aristotle before any notion of a square entered into it, and it has been further developed on the basis of other diagrams: cubes, hexagons, polyhedra … The SQUARE is open to semiotics, linguistics, psychology, psychoanalysis, art, architecture, sociology, etc. After starting the first SQUARE in Montreux like the first UNILOG, we have organized the following square events in some eccentric/unsual localities: Corsica, Beirut, the Vatican. Our idea is to go on in that direction, and the next event is planned for Rapa Nui/Easter Island (Fig. 15). The SQUARE may end on the moon.

LIQ is a bilingual workshop in Paris about logic in its diversity, to promote interaction between anyone interested in logic in a wide sense, trying to answer some basic questions about the nature and importance of logic. I developed the idea of this workshop together with Jean-Pierre Desclés, who for many years was the director of the Department of Mathematics and Informatics at the University of Paris Sorbonne (Paris 4), Anca Pascu, a former student of his, and Amirouche Moktefi, one of the best specialists in the logic of Lewis Carroll. We had one workshop where we discussed the teaching of logic in high school. However, LIQ can also be topical, in the sense that it acknowledges and addresses current events – such as, for example, the workshop in 2014, where we had a special panel about Peirce, as it was the centenary of his death.

I am presently working with my colleague Ricardo Silvestre on the organisation of the *1st World Congress on Logic and Religion*, to be held April 1–5, 2015 in João Pessoa, Brazil. The connection between logic and religion is very strong, in many different ways. The notion of *logos* figures in the Bible, where it is identified with God: “In the beginning was the *logos*, and the *logos* was with God, and the *logos* was God”(John 1:1). In Christian culture, many “proofs” of God were proposed from Anselm to Gödel. In the other main religions, reasoning is also important. I had planned to organize such an event when I was in Switzerland. I wanted to organize it in *Monte Verità*, the “Hill of Truth”, located near Ascona in Ticino, where a peculiar colony attracting many famous people was created at the beginning of the 19th century. It is now a centre for congresses of all kinds administrated by ETZ Zurich (the Swiss Federal Institute of Technology). This event in João Pessoa will probably be the first in another successful series of itinerant international events.

I am actively taking part in the organisation of two other series of events. The Brazilian Congress of Logic (EBL) is organized by the Brazilian Society of Logic (SBL), of which I am presently the vice-president. The 17th EBL was held in Petrópolis in 2014. The next EBL is scheduled to take place in Pirenópolis, a nice bucolic village near Brasilia, in 2017. The World Congress of Paraconsistency (WCP) was staged in Ghent, Belgium in 1997 (WCP1); in Juquehy, Brazil in 2010 (WCP2); in Toulouse, France in 2003 (WCP3); in Melbourne, Australia in 2008 (WCP4); and in Kolkata, India in 2014 (WCP5). I was one of the principal organizers of WCP2, WCP3 and WCP5. The plan is to stage the next one, WCP6, in Vancouver, Canada in 2016.

In Athens in 2013, during the 23rd World Congress of Philosophy, as the Director of International Relations of the Brazilian Academy of Philosophy, I put forward the candidacy of Rio de Janeiro to host a future staging of this event. The 24th congress will be held in Beijing in 2018, and we are working on organizing the 25th edition in Rio in 2023. Before that we intend to organize the 37th congress of ASPLF (Association des Sociétés de Philosophie de Langue Française) in Rio de Janeiro in 2018.

7.3 Editorial Work

Up until now I have launched two logic journals: *Logica Universalis* (LU) and the *South American Journal of Logic* (SAJL); and three book series: *Studies in Universal Logic*, *Logic PhDs*, and *Cadernos de Lógica e Filosofia*. I will go on to develop these editorial projects and develop new ones. I have the idea of developing a book series dedicated to some specific logical systems: *All about the Modal Logic S5*, *All about First-Order Logic* and *All about Łukasiewicz's Three-Valued Logic L3*, etc. These books will include historical, philosophical and mathematical accounts of these systems. I also intend in the future to launch the *World Journal of Pictorial Philosophy* (WJPP), a journal of philosophy mainly based on the use of images. The cover will be a picture of Plato's cave.

In the coming months I will devote a lot of energy to editing entries in the area of logic for the Internet Encyclopedia of Philosophy, of which I am the present editor. In addition to being on the web, these entries will also be gathered into a printed book to be titled *Encyclopaedia of Logic*, the first ever encyclopedia of logic. To develop this project I have divided the entries into three categories: history (main figures, schools and books), tools (theorems, concepts, logical systems), and notions/topics (including the relation of logic to other fields). Hundreds of entries are on the way...

I will also go on to develop the project of "World Philosophy" with Catherine Chantilly, producing a series of DVDs, each dedicated to philosophers of a given country.

7.4 University of the World

Beyond my research in logic, I have the idea to promote the creation and development of a world research center, which I prefer calling a "world university", or better, a "university of the world". On the one hand, I firmly believe that research and teaching have to be linked; so calling it a "research center" would not sufficiently emphasize the teaching aspect, which is an essential component in my view. On the other hand, the word "university" fits nicely because it is connected with universality and the universe.

Human beings, including academics, still behave in a very primitive way, close to tribalism. Groups fighting against groups, an "evolution" of cannibalism. I think the creation of a world university could get us out of that. How, where and when? Those are elementary but fundamental questions, which will be answered by actions promoted and undertaken by people who are conscious of the limitations of tribe-centered activities. We can be inspired by people who have already worked in that direction, like for example Julian Huxley, brother of Aldous Huxley, who was the first director of UNESCO and also the creator of the World Wide Fund for Nature (WWF). I already gave some preliminary hints in my paper "Les universités face à la globalisation: vers une université mondiale?", which I presented at UNESCO in 2004.

8 Workshops and Events Organized

1. *Logic in Rio de Janeiro*, Rio de Janeiro, Brazil, February 17–18, 2000
2. *2nd World Congress on Paraconsistency*, Juquehy, Brazil, May 8–10, 2000
3. *Workshop on Paraconsistent Logic – Part of International Conference on Artificial Intelligence IC-AI'2001*, Las Vegas, USA, June 25–28, 2001
4. *Foundations of Science Workshop dedicated to the 80th Birthday of Patrick Suppes*, Florianópolis, Brazil, April 22–23, 2002
5. *3rd World Congress on Paraconsistency*, Toulouse, France, July 28–31, 2003
6. *International Workshop on Universal Logic*, Neuchâtel, Switzerland, October 6–8, 2003
7. *1st World Congress and School on Universal Logic UNILOG'05*, Montreux, Switzerland, March 26–April 3, 2005

8. *Symbolic language – Interdisciplinary workshop on logic, semiotics, linguistics, chemistry, physics, psychology, mathematics, philosophy, theology and art*, Neuchâtel, Switzerland, December 7–8, 2005
9. *The Square of Opposition – A general framework for cognition*, Montreux, Switzerland, June, 1–3, 2007
10. *CombLog'07 – International Workshop on Combinations of Logics*, Neuchâtel, Switzerland, July 4th, 2007
11. *2nd World Congress and School on Universal Logic – UNILOG'07*, Xian, China, August, 16–22, 2007
12. *Imagination – Interdisciplinary workshop on logic, semiotics, linguistics, chemistry, physics, psychology, mathematics, philosophy, theology and art*, Neuchâtel, Switzerland, December 17–18, 2007
13. *Workshop with Saul Kripke*, Neuchâtel, Switzerland, June 5, 2010
14. *3rd World Congress and School on Universal Logic – UNILOG'2010*, Lisbon, Portugal, April, 18–25, 2010
15. *2nd World Congress on the Square of Opposition*, Corte, Corsica, June, 17–20, 2010
16. *Workshop LIQ1 – Logic in Question 1*, Sorbonne, Paris, May 2–3, 2011
17. *Workshop LIQ2 – Logic in Question 2*, Sorbonne, Paris, May 2–3, 2012
18. *3rd World Congress on the Square of Opposition*, Beirut, Lebanon, July, 17–20, 2012
19. *4th World Congress and School on Universal Logic – UNILOG'2013*, Rio de Janeiro, Brazil, April, 18–25, 2013
20. *Workshop LIQ3 – Logic in Question 3*, Sorbonne, Paris, May 2–3, 2013
21. *Workshop on Imagination* at the *23th World Congress of Philosophy*, Athens, August 4–10, 2013
22. *5th World Congress on Paraconsistency*, Kolkata, India, February 13–17, 2014
23. *4th World Congress on the Square of Opposition*, Vatican, May 5–9, 2014
24. *Workshop LIQ4 – Logic in Question 4*, Sorbonne, Paris, France, May 12–13, 2014
25. *Workshop The Logic of Lewis Carroll*, Federal University of Rio de Janeiro, Brazil, November 28, 2014

9 Writings and Publications

9.1 University Writings

1. Beziau, J.-Y.: L'holomouvement selon David Bohm, Maîtrise de Philosophie, Bernard d'Espagnat (Advisor), 150 p., Department of Philosophy, University Panthéon-Sorbonne, Paris (1986). This work includes a discussion I had with David Bohm in London in July 1986.
2. Beziau, J.-Y.: Quels sont les moyens par lesquels on peut se convaincre soi-même (et les autres) de la vérité des assertions concernant les ensembles?, 13p., Homework. Maîtrise de Logique, Michel Eytan (Advisor), Department of Philosophy, University Panthéon-Sorbonne (Paris 1), Paris (1988)

3. Beziau, J.-Y.: D'une caverne à l'autre, 60 p., Diplôme d'Etudes Approfondies de Logique et Fondements de l'Informatique, Sarah Kofman (Advisor), Department of Philosophy, University Panthéon-Sorbonne (Paris 1), Paris (1988)
4. Beziau, J.-Y.: La logique paraconsistante C1 de Newton da Costa, 51 p., Diplôme d'Etudes Approfondies de Logique et Fondements de l'Informatique, Daniel Andler (Advisor), Department of Mathematics, University Denis Diderot (Paris 7), Paris (1990)
5. Beziau, J.-Y.: Recherches sur la logique universelle (Excessivité, Négation, Séquents), 176 p., Doctorat de Logique et Fondements de l'Informatique, Daniel Andler (Advisor), Department of Mathematics, University Denis Diderot (Paris 7), Paris (1995)
6. Beziau, J.-Y.: Sobre a verdade lógica, 200 p., Doutorado de Filosofia, Newton da Costa (Advisor), Department of Philosophy, University of São Paulo (1996)

9.2 Papers

1. Beziau, J.-Y.: Calcul des séquents pour logique non-aléthique. Log. Anal. **125–126**, 143–155 (1989)
2. Beziau, J.-Y.: Logiques construites suivant les méthodes de da Costa. Log. Anal. **131–132**, 259–272 (1990)
3. Beziau, J.-Y.: Au sujet d'une preuve du principe de contradiction, 3 p., Vincennes, France, unpublished (1991)
4. Beziau, J.-Y.: O princípio de razão suficiente e a lógica segundo Arthur Schopenhauer. In: Évora, F.R.R. (ed.) Século XIX : O Nascimento da Ciência Contemporânea, pp. 35–39. CLE-Unicamp, Campinas (1992)
5. da Costa, N.C.A., Beziau, J.-Y.: Carnot's logic. Bull. Sect. Log. **22**, 98–105 (1993)
6. Beziau, J.-Y.: La critique Schopenhauerienne de l'usage de la logique en mathématiques. O Que Nos Faz Pensar **7**, 81–88 (1993)
7. Beziau, J.-Y.: Nouveaux résultats et nouveau regard sur la logique paraconsistante C1. Log. Anal. **141–142**, 45–58 (1993)
8. Beziau, J.-Y.: La logique abstraite au sein de la mathématique moderne. Ruch Filoz. **50**, 289–293 (1993)
9. Beziau, J.-Y.: Sémantique universelle, 18 p., Champagne sur Seine, France, unpublished (1994)
10. Beziau, J.-Y.: Théorie législative de la négation pure. Log. Anal. **147–148**, 209–225 (1994)
11. Beziau, J.-Y.: Universal logic. In: Childers, T., Majer, O. (eds.) Logica'94 – Proceedings of the 8th International Symposium, Prague, pp. 73–93 (1994)
12. da Costa, N.C.A., Beziau, J.-Y.: La théorie de la valuation en question. In: Abad, M. (ed.) Proceedings of the Ninth Latin American Symposium on Mathematical Logic, pp. 95–104. Universidad del Sur, Baíha Blanca (1994)
13. da Costa, N.C.A., Beziau, J.-Y.: Théorie de la valuation. Log. Anal. **146**, 95–117 (1994)

14. Beziau, J.-Y.: De la logique formelle à la logique abstraite. Bol. Soc. Parana. Mat. **14**, 41–50 (1994)
15. Beziau, J.-Y.: Du Pont's paradox and the problem of intensional logic. In: Kolar, P., Svodoba, V. (eds.) Logica'93 – Proceedings of the 7th International Symposium, Prague, 1994, pp. 62–65 (1994)
16. da Costa, N.C.A., Beziau, J.-Y., Bueno, O.A.S.: Paraconsistent logic in a historical perspective. Log. Anal. **150–152**, 111–125 (1995)
17. da Costa, N.C.A., Bueno, O.A.S., Beziau, J.-Y.: What is semantics?. Sorites **3**, 43–47 (1995)
18. da Costa, N.C.A., Beziau, J.-Y., Bueno, O.A.S.: Aspects of paraconsistent logic. Bull. Interest Group Pure Appl. Log. **4**, 597–614 (1995)
19. Beziau, J.-Y.: Negation: what it is and what it is not. Bol. Soc. Parana. Mat. **15**, 37–43 (1995)
20. da Costa, N.C.A., Beziau, J.-Y.: Théorie paraconsistante des ensembles. Log. Anal. **153–154**, 51–67 (1996)
21. da Costa, N.C.A., Beziau, J.-Y., Bueno, O.A.S.: Malinowski and Suszko on many-valued logics: On the reduction of many-valuedness to two-valuedness. Mod. Log. **6**, 272–299 (1996)
22. Beziau, J.-Y.: Identity, logic and structure. Bull. Sect. Log. **25**, 89–94 (1996)
23. Beziau, J.-Y.: Logic may be simple. Log. Log. Philos. **5**, 129–147 (1997)
24. da Costa, N.C.A., Beziau, J.-Y.: Overclassical logic. Log. Anal. **157**, 31–44 (1997)
25. Beziau, J.-Y.: What is many-valued logic?. In: Proceedings of the 27th International Symposium on Multiple-Valued Logic, pp. 117–121. IEEE Computer Society, Los Alamitos (1997)
26. Krause, D., Beziau, J.-Y.: Relativizations of the principle of identity. Log. J. Interest Group Pure Appl. Log. **5**, 327–338 (1997)
27. Beziau, J.-Y.: Do sentences have identity?. In: Proceedings of the XXth World Congress of Philosophy. The Paideia Project, Boston (1998)
28. da Costa, N.C.A., Beziau, J.-Y.: Définitions, théories des objets et paraconsistance. Theoria **32**, 367–379 (1998)
29. Beziau, J.-Y.: Idempotent full paraconsistent negations are not algebraizable. Notre Dame J. Form. Log. **39**, 135–139 (1998)
30. Beziau, J.-Y.: Recherches sur la logique abstraite: les logiques normales. *Acta Univ. Wratislav.* **18**, 105–114 (1998)
31. Beziau, J.-Y.: Ruth Barcan Marcus est-elle la mère du fils de Wittgenstein? (Considérations existentialistes sur la formule de Barcan). Manuscrito **22**, 11–27 (1999)
32. Beziau, J.-Y.: A sequent calculus for Lukasiewicz's three-valued logic based on Suszko's bivalent semantics. Bull. Sect. Log. **28**, 89–97 (1999)
33. Beziau, J.-Y.: The future of paraconsistent logic. Log. Stud. **2**, 1–28 (1999). Romanian translation in Lucica, I., et al. (eds.) Ex falso qodlibet, pp. 159–181. Tehnica, Bucarest (2004)
34. Beziau, J.-Y.: A logical analysis of singular terms. *Sorites* **10**, 6–14 (1999)
35. Beziau, J.-Y.: Rules, derived rules, permissible rules and the various types of systems of deduction. In: Haeusler, E.H., Pereira, L.C. (eds.) Proof, Types and Categories, pp. 159–184. PUC, Rio de Janeiro (1999)

36. Beziau, J.-Y.: The mathematical structure of logical syntax. In: Carnielli, W.A., D'Ottaviano, I.M.L. (eds.) Advances in Contemporary Logic and Computer Science, pp. 1–17. American Mathematical Society, Providence (1999)
37. da Costa, N.C.A., Beziau, J.-Y.: La logique paraconsistante. In: Sallantin, J., Szczeciniarz, J.J. (eds.) La preuve à la lumière de l'intelligence artificielle, pp. 107–115. Presses Universitaires de France, Paris (1999)
38. Beziau, J.-Y.: Was Frege wrong when identifying reference with truth-value?. Sorites **11**, 15–23 (1999)
39. Beziau, J.-Y.: Classical negation can be expressed by one of its halves. Log. J. Interest Group Pure Appl. Log. **7**, 145–151 (1999)
40. Beziau, J.-Y.: La véritable portée du théorème de Lindenbaum-Asser. Log. Anal. **167–168**, 341–359 (1999)
41. Beziau, J.-Y.: Y a-t-il des principes logiques?. In: Dutra, L.H., Mortari, C.A. (eds.) Princípios Seu papel na filosofia e nas ciências, pp. 47–54. NEL, Federal University of Santa Catarina, Florianópolis (2000)
42. 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)
43. Beziau, J.-Y.: Sequents and bivaluations. Log. Anal. **44**, 373–394 (2001)
44. Beziau, J.-Y.: The logic of confusion. In: Arabnia, H.R. (ed.) Proceedings of the International Conference of Artificial Intelligence IC-AI2002, pp. 821–826. CSREA Press, Las Vegas (2001)
45. Beziau, J.-Y.: From paraconsistent to universal logic. Sorites **12**, 5–32 (2001)
46. Beziau, J.-Y.: What is classical propositional logic?. Log. Investig. **8**, 266–277 (2001)
47. Beziau, J.-Y.: Possible worlds: a fashionable non-sense?. Unpublished (2001)
48. Beziau, J.-Y.: La théorie des ensembles et la théorie des catégories: présentation de deux soeurs ennemies du point de vue de leurs relations avec les fondements des mathématiques. Bol. Asoc. Mat. Venez. **9**, 45–53 (2002)
49. Beziau, J.-Y.: The philosophical import of Polish logic. In: Talasiewicz, M. (ed.) Methodology and Philosophy of Science at Warsaw University, pp. 109–124. Semper, Warsaw (2002)
50. Beziau, J.-Y.: S5 is a paraconsistent logic and so is first-order classical logic. Log. Investig. **9**, 301–309 (2002)
51. Beziau, J.-Y.: Are paraconsistent negations negations?. In: Carnielli, W., et al. (eds.) Paraconsistency: The Logical Way to the Inconsistent, pp. 465–486. Marcel Dekker, New-York (2002)
52. Beziau, J.-Y.: New light on the square of oppositions and its nameless corner. Log. Investig. **10**, 218–232 (2003)
53. Beziau, J.-Y.: Quine on identity. Principia **7**, 1–15 (2003)
54. Beziau, J.-Y.: Bivalence, exluded middle and non contradiction. In: Behounek, L. (ed.) The Logica Yearbook 2003, pp. 73–84. Academy of Sciences, Prague (2003)
55. Beziau, J.-Y.: A paradox in the combination of logics. In: Carnielli, W.A., Dionisio, F.M., Mateus, P. (eds.) Workshop on Combination of Logics: Theory and Applications, pp. 75–78. IST, Lisbon (2004)

56. Suppes, P., Beziau, J.-Y.: Semantic computation of truth based on associations already learned. J. Appl. Log. **2**, 457–467 (2004)
57. da Costa, N.C.A., Beziau, J.-Y., Bueno, O.A.S.: On the usefulness of paraconsistent logic. In: Vanderveken, D. (ed.) Logic, Thought and Action. Logic, Epistemology, and the Unity of Science, pp. 465–478. Springer (2005)
58. Beziau, J.-Y.: Le château de la quantification et ses fantômes démasqués. In: Joray, P. (ed.) La quantification dans la logique moderne, pp. 211–232. L'Harmattan, Paris (2005)
59. Beziau, J.-Y.: Paraconsistent logic from a modal viewpoint. J. Appl. Log. **3**, 7–14 (2005)
60. Beziau, J.-Y.: Les universités face à la globalisation : vers une université mondiale ? In: UNESCO Philosophical Day 2004, pp. 207–222. UNESCO, Paris (2006)
61. Beziau, J.-Y.: Universal Logic in 13 questions. Bull. Sect. Log. **25**, 133–150 (2006)
62. Beziau, J.-Y.: The paraconsistent logic Z. A possible solution to Jaskowski's problem. Log. Log. Philos. **15**, 199–211 (2006)
63. Beziau, J.-Y.: Transitivity and paradoxes. In: Skilters, J. (ed.) The Baltic International Yearbook of Cognition, Logic and Communication, pp. 207–211. University of Riga, Riga (2006)
64. Beziau, J.-Y.: Les axiomes de Tarski. In: Pouivet, R., Rebuschi, M. (eds.) La philosophie en Pologne 1918–1939, pp. 135–149. Vrin, Paris (2006)
65. Beziau, J.-Y.: Many-valued and Kripke semantics. In: van Benthem, J., et al. (eds.) The age of alternative logics, pp. 89–101. Springer (2006)
66. Beziau, J.-Y.: Sentence, Proposition and Identity. Synthese **154**, 371–38 (2007)
67. Beziau, J.-Y.: Adventures in the paraconsistent jungle. In: Handbook of Paraconsistency. College Publication, London (2007)
68. Beziau, J.-Y.: Mystérieuse identitée. Le même et l'autre, identité et différence – Actes du XXXIe Congrès International de l'ASPLF, Eotvos, Budapest, 2009, pp. 159–162 (2009)
69. Beziau, J.-Y.: What is 'formal logic'?. In: Myung-Hyun-Lee (ed.) Proceedings of the XXII World Congress of Philosophy, vol. 13, pp. 9–22. Korean Philosophical Association, Seoul (2008)
70. Beziau, J.-Y.: Bivalent semantics for De Morgan logic (the uselessness of fourvaluedness). In: Carnielli, W.A., Coniglio, M.E., D'Ottaviano, I.M.L. (eds.) The Many Sides of Logic, pp. 391–402. College, London (2009)
71. Beziau, J.-Y.: Synopsis of Robert Blanché, 'Sur le système des connecteurs interpropositionnels' ['On the System of Interpropositional Connectors']. http://cahiers.kingston.ac.uk/synopses/syn10.7.html
72. Beziau, J.-Y.: Biconditional drive to paradox. Rev. Bras. Filos. **233**, 196–201 (2009)
73. Beziau, J.-Y.: What is a logic? – Towards axiomatic emptiness. Log. Investig. **16**, 272–279 (2010)
74. Beziau, J.-Y., Chantilly, C.: Salomé multi-screen. In: Sobieczszanski, M., Masoni-Lacroix, C. From Split-screen to Multi-screen, pp. 319–326. Peter Lang, Bern (2010)
75. Beziau, J.-Y.: What is a possible world?. In: Imaguire, G., Jacquette, D. (eds.) Possible Worlds. Philosophia Verlag, Munich (2010)

76. Beziau, J.-Y.: Rougier: Logique et Métaphysique. In: Murray, D.G. (ed.) 4th World Conference on Metaphysics, pp. 464–472. Fondation Ortega y Gasset, Madrid (2010)
77. Beziau, J.-Y.: Truth as a mathematical object. Principia **14**, 31–6 (2010)
78. Beziau, J.-Y.: Logic is not logic. Abstracta **6**, 73–102 (2010)
79. Beziau, J.-Y., Kritz, M.V.: Théorie et Modèle I: Point de vue général et abstrait. Cadernos UFS Filos. **6**, 9–17 (2010)
80. Beziau, J.-Y., Coniglio, M.E.: To distribute or not to distribute?. Log. J. Interest Group Pure Appl. Log. **19**, 566–583 (2011)
81. Pequeno, T., Beziau, J.-Y.: Rules of the game. In: Beziau, J.-Y., Coniglio, M.E. (eds.) Logic Without Frontiers, pp. 131–144. College Publication, London (2011)
82. Beziau, J.-Y.: Badiou et les modèles. In: Vodoz, I., Tarby, F. (eds.) Autour d'Alain Badiou. Germina, Paris (2011)
83. Beziau, J.-Y.: A new four-valued approach to modal logic. Log. Anal. **54**, 18–33 (2011)
84. Beziau, J.-Y.: Pure alethic modal logic. Coginitio **13**, 25–36 (2012)
85. Beziau, J.-Y.: The power of the hexagon. Log. Univers. **6**, 1–43 (2012)
86. Kritz, M.V., Beziau, J.-Y.: Théorie et Modèle II: Sciences empiriques. Cadernos UFS Filos. **8**, (2012)
87. Beziau, J.-Y.: La logique universelle. De la logique moderne à la logique post-moderne. In: Schmid, A.-F. Epistémologie des frontières, pp. 30–59. Pťra, Paris (2012)
88. Beziau, J.-Y.: The new rising of the square of opposition. In: Beziau, J.-Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition, pp. 6–24. Birkhäuser, Basel (2012)
89. Beziau, J.-Y.: History of truth-values. In: Gabbay, D.M., Woods, J. (eds.) Handbook of the History of Logic, Vol. 11 – Logic: a history of its central concepts, pp. 233–305. Elsevier, Amsterdam (2012)
90. Beziau, J.-Y.: Paralogics and the theory of valuation. In: Universal Logic: An Anthology – From Paul Hertz to Dov Gabbay, pp. 361–372. Birkhäuser, Basel (2012)
91. Beziau, J.-Y.: Les modèles selon Alain Badiou. Al Mukhatabat **1/3**, 251–305 (2012)
92. Beziau, J.-Y.: Preface of Universal Logic: An Anthology – From Paul Hertz to Dov Gabbay, pp. v–xii. Birkhäuser, Basel (2012)
93. Beziau, J.-Y.: The metalogical hexagon of opposition. Argumentos **10**, 111–122 (2013)
94. Beziau, J.-Y.: Three Sisters: Philosophy, Mathematics and Logic. In: Nabais, N., Pombo, O. (eds.) O lugar da Filosofia da Ciência na Universidade de Lisboa, pp. 171–191. CFCUL, University of Lisbon (2013)
95. Beziau, J.-Y.: Opposition and order. In: Beziau, J.-Y., Gan-Krzywoszynska, K. (eds.) New Dimensions of the Square of Opposition, pp. 321–336. Philosophia Verlag, Munich (2014)
96. Beziau, J.-Y.: The relativity and universality of logic. Synthese – Special Issue Istvan Németi 70th Birthday, on-line first March 2014 (2014)
97. Beziau, J.-Y.: Paraconsistent logic and contradictory viewpoints. Rev. Bras. Filos. **241** (2014)

98. Beziau, J.-Y.: Linguistica et Logica Acta. In: Mathieu, N., Schmid, A.-F. (eds.) Modélisation et Interdisciplinarité – Six Disciplines en Quête d'Épistémologie. Quae, Versailles (2014)
99. 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 Publication, London (2015)
100. Beziau, J.-Y.: Is modern logic non-Aristotelian?. In: Zaitsev, D. (ed.) Nikolai Vasiliev's Logical Legacy and Modern Logic. Springer, Heidelberg (2015)
101. Beziau, J.-Y.: Le possible et l'impossible, au-delà de la dichotomie. In: Ferrari, J., et al. (eds.) Actes du 35 ème Congrès de l'ASPLF – Association des Sociétés de Philosophie de Langue Francçaise. Vrin, Paris (2015)
102. Beziau, J.-Y., Franceschetto, A.: Strong three-valued paraconsistent logics. In: Beziau, J.-Y., Chakraborty, M., Dutta, S. (eds.) New Directions in Paraconsistent Logic. Springer, New Delhi (2015)
103. Beziau, J.-Y.: Modelization of causality. In: Beziau, J.-Y., Krause, D., Arenhart, J. (eds.) Conceptual Clarifications. Festschrift for Patrick Suppes on the Occasion of his 90th Birthday. College Publication, London (2015)
104. Beziau, J.-Y.: Possible worlds: a fashionable nonsense?, 10 p., unpublished, Stanford (2001)
105. Beziau, J.-Y.: Identification of identity, 8 p., unpublished, Neuchâtel (2008)

9.3 Edited Books

1. Arabnia, H.R., Beziau, J.Y., et al. (eds.) Proceedings of IC-AI'2001 (International Conference on Artificial Intelligence). CSREA Press, Las Vegas (2001)
2. Beziau, J.-Y., Costa Leite, A., Facchini, A. (eds.) Aspects of Universal Logic. University of Neuchâtel, Neuchâtel (2004)
3. J.-Y. Beziau (ed.): Logica Universalis. Birkhäuser, Basel (2005), second edition (2007)
4. Beziau, J.-Y., Costa Leite, A. (eds.): Perspectives on Universal Logic, Polimetrica, Monza (2006)
5. Beziau, J.-Y., Carnielli, W.A., Gabbay, D.M. (eds.): Handbook of Paraconsistency. King's College, London (2007)
6. Beziau, J.-Y., Costa Leite, A. (eds.): Dimensions of Logical Concepts. CLE, Campinas (2009)
7. Beziau, J.-Y., Payette, G. (eds.): New Perspectives on the Square of Opposition. Peter Lang, Bern (2011)
8. Beziau, J.-Y. (ed.): Anthology of Universal Logic – From Paul Hertz to Dov Gabbay. Birkhäuser, Basel (2012)
9. Beziau, J.-Y., Coniglio, M.E. (eds.): Logic without Frontiers. Festschrift for Walter Alexandre Carnielli on the Occasion of his 60th Birthday. College Publication (2011)

10. Beziau, J.-Y., Jacquette, D. (eds.): Beyond and Around the Square of Opposition. Springer Lang, Bern (2012)
11. Beziau, J.-Y., Gan-Krzywoszynska, K. (eds.): New Dimensions of the Square of Opposition. Philosophia Verlag, Munich (2014)
12. Beziau, J.-Y. (ed.): La Pointure du Symbole. Petra, Paris (2014)
13. Beziau, J.-Y., Krause, D., Arenhart, J. (eds.): Conceptual Clarifications. Festschrift for Patrick Suppes on the Occasion of his 90th Birthday. College Publication, London (2015)
14. Beziau, J.-Y., Chakraborty, M., Dutta, S. (eds.): New Directions in Paraconsistent Logic. Springer, New Delhi (2015)
15. J.-Y. Beziau (ed.): Encyclopaedia of Logic. College Publication, London (2015)

9.4 Edited Special Issues of Journals

1. Beziau, J.-Y., Doria, F.A. (eds.): Contemporary Brazilian Research in Logic Part I. Log. Anal. **153–154** (1996)
2. Beziau, J.-Y., Tsuji, M. (eds.): Contemporary Brazilian Research in Logic Part II. Log. Anal. **157** (1997)
3. Beziau, J.-Y.: The Challenge of combining logics. Log. J. Interest Group Pure Appl. Log. **19** (2011)
4. Beziau, J.-Y., Krause, D. (eds.): New trends in the foundations of science. Synthese **154**(3) (2007)
5. Beziau, J.-Y., Costa Leite, A. (eds): Uses of non-classical logics: foundational issues. J. Apppl. Non-Class. Log. **21** (2011)
6. Beziau, J.-Y., Read, S. (eds.): The square of opposition in historical perspective. Hist. Philos. Log. (2014)

9.5 Books

1. da Costa, N.C.A., Beziau, J.Y., Bueno, O.A.S.: Elementos de Teoria Paraconsistente dos conjuntos. CLE, Campinas (1998)
2. Beziau, J.-Y.: Tendances actuelles de la philosophie – Tendências Atuais da Filosofia. Nefelibata, Florianópolis (2003)

9.6 Translation

1. Translation of the book *Ensaio sobre os Fundamentos da Lógica* by Newton da Costa in French: Logiques Classiques et non Classiques, with a Preface and two Appendices by the translator. Masson, Paris (1997)

9.7 Reviews

1. Anellis, I.H.: Logic and its history in the work and writings of J. van Heijenoort, published in Mod. Log. **8**, 105–117 (2000)
2. Peirce, C.S.: Chance, Love and Logic. Bison Books (1998). Review published in: Mitcham, C. (ed.) Metaphysics, Epistemology and Technology, pp. 395–397. Elsevier, New York (2000)
3. Grattan-Guinness, I.: The search for mathematical roots, published in Rev. Mod. Log. **10**, 135–138 (2005)
4. Brady, G.: From Peirce to Skolem. A neglected chapter in the history of logic, published in Rev. Mod. Log. **11**, 155–161 (2007)
5. Odintsov, S.: Constructive Negation and Paraconsistency, published in Stud. Log. **100**, 653–65 (2012)
6. Shramko, Y., Wansing, H.: Truth and Falsehood. An inquiry into generalized logical values, published in Stud. Log. **102**, 1079–1085 (2014)

I have also written about 100 reviews for *Mathematical Reviews*.

9.8 Future Papers

All these papers are at a stage of gestation, the birth of each will depend on circumstances. Some will perhaps never be born, and others, not in this list, may appear.

9.8.1 Square of Opposition

1. Beziau, J.-Y.: "The two dualities *a priori/a posteriori* and *synthetic/analytic* in a hexagonal perspective"
2. Beziau, J.-Y.: "A semiotic hexagon"
3. Beziau, J.-Y.: "Incompatibility"
4. Beziau, J.-Y.: "Beyond dichotomy"
5. Beziau, J.-Y.: "The logic of traffic sign"
6. Beziau, J.-Y.: "Square and hexagon of causality"
7. Beziau, J.-Y.: "A triangle of ways of reasoning: induction, deduction, abduction"

9.8.2 Paraconsistency and negation

1. Beziau, J.-Y.: "Round squares are no contradictions"
2. Beziau, J.-Y.: "Abstract theory of negation"
3. Beziau, J.-Y.: "Two formulations/formalizations of the principle of non-contraction"
4. Beziau, J.-Y.: "What is paracomplete logic?"
5. Beziau, J.-Y.: "Cats, tigers and stones"

6. Beziau, J.-Y.: “Absurdity, triviality, nonsense, contradiction”
7. Beziau, J.-Y.: “Identity and contradiction”
8. Beziau, J.-Y.: “Disjunctive syllogism and paraconsistency”
9. Beziau, J.-Y.: “Conditional negation”
10. da Costa, N.C.A. and Beziau, J.-Y.: “Is god paraconsistent?”
11. Beziau, J.-Y.: “Philosophical aspects of the paraconsistent logic Z”

9.8.3 Order Logic

1. Beziau, J.-Y.: “What is zero-order logic?”
2. Beziau, J.-Y.: “The philosophy of first-order logic”
3. Beziau, J.-Y., Viana, P.: “Third-order logic”
4. Beziau, J.-Y.: “Undetermined constants, variables, parameters”

9.8.4 Philosophy of Logic/Logical Philosophy

1. Beziau, J.-Y.: “How to define logic”
2. Beziau, J.-Y.: “Logic: past, present and future”
3. Beziau, J.-Y., Carnielli, W.A.: “Names of logic”
4. Beziau, J.-Y.: “The identity sign”
5. Beziau, J.-Y.: “What is modal logic?”
6. Beziau, J.-Y.: “What is constructivism?”
7. Beziau, J.-Y.: “Non-transitive logics”
8. Beziau, J.-Y.: “What is a propositional function?”
9. Beziau, J.-Y.: “What is a propositional variable”
10. Beziau, J.-Y.: “Philosophy of logic, philosophical logic and logical philosophy”
11. Beziau, J.-Y.: “Logic, reasoning and rationality”
12. Beziau, J.-Y.: “Three notions of contingency”
13. Beziau, J.-Y.: “Peirce’s law: its meaning and behaviour”
14. Beziau, J.-Y.: “The ambiguity of quotation marks”
15. Beziau, J.-Y.: “Panorama de l’identité”
16. Beziau, J.-Y., Marcos, J.: “What is non truth-functional logic?”
17. Beziau, J.-Y.: “What is positive propositional logic?”

9.8.5 History of Logic

1. Beziau, J.-Y., Buchsbaum, A.: “Adventures of the turnstyle”
2. Beziau, J.-Y., Hudry, J.L.: “All men are white”
3. Beziau, J.-Y.: “The origin of classical logic”

9.8.6 Universal Logic

1. Beziau, J.-Y.: “The completeness theorem in a universal logic perspective”
2. Beziau, J.-Y.: “The incompleteness theorem in a universal logic perspective”
3. Beziau, J.-Y.: “Logics and theories”
4. Beziau, J.-Y.: “Absolute maximality”
5. Beziau, J.-Y., Sernadas, A.: “Combining disjunction with negation”
6. Beziau, J.-Y, Caleiro. C.: “Non truth-functional extensions of truth-functional logics”
7. Beziau, J.-Y.: “Bivalence and modality”

9.8.7 Language, Thought and Reality

1. Beziau, J.-Y.: “Exemplifying an idea”
2. Beziau, J.-Y.: “Word, idea and reality”
3. Beziau, J.-Y.: “Descriptive and normative”
4. Beziau, J.-Y.: “Islands and clouds”
5. Beziau, J.-Y.: “What is a category?”
6. Beziau, J.-Y.: “Imagination, conceptualization and possibility”
7. Beziau, J.-Y.: “Rationality and representation”
8. Beziau, J.-Y., Chantilly, C., Lihoreau, F.: “Sensation, feeling and emotion”
9. Beziau, J.-Y.: “Dices, hazardous symbol of chance”
10. Beziau, J.-Y.: “Symbolically typical”
11. Beziau, J.-Y.: “Aspects of structuralism”

9.8.8 General Philosophy

1. Beziau, J.-Y.: “Three definitions of human beings”
2. Beziau, J.-Y.: “Rodin’s thinker: a symbol for philosophy?”
3. Beziau, J.-Y.: “Rational animals”
4. Beziau, J.-Y.: “Control Z”
5. Beziau, J.-Y.: “Death in 5 lessons”

9.9 Future Books

9.9.1 Edited Books

1. Beziau, J.-Y. (ed.): *Anthology of Paracaconstent Logic*
2. Beziau, J.-Y. (ed.): *Anthology of Modal Logic*
3. Beziau, J.-Y. (ed.): *Anthology of Many-valued Logic*
4. Beziau, J.-Y., Moktefi, A. (eds.): *Conceptions of Logic through History*

9.9.2 Monographs

1. Beziau, J.-Y.: *Logic Compendium*
2. Beziau, J.-Y.: *An Introduction to Universal Logic*
3. Beziau, J.-Y.: *Beyond Dichotomy*
4. Beziau, J.-Y., Costa-Leite, A.: *A Panoramic Introduction to Paraconsistent Logic*
5. da Costa, N.C.A., Beziau, J.-Y.: *The World of Possible Logics*

References

1. Andler, D.: Models of uncountable theories categorical in power. PhD, University of California, Berkeley (1973)
2. de Azevedo, A.: Grothendieck no Brasil. Rev. Mat. Univ. **44**, 39–49 (2008)
3. van Benthem, J.F.K., et al. (eds.): The Age of Alternative Logics. Springer, Dordrecht (2006)
4. Blanché, R.: Structures intellectuelles. Essai sur l'Organisation Systématique des Concepts. Vrin, Paris (1966)
5. Bohm, D.: Wholeness and the Implicate Order. Routlege, London (1980)
6. Cori, R., Lascar, D.: Logique Mathématique, Volume I et II. Masson, Paris, (1993). English translation: Mathematical Logic. A course with Exercices, Volume I and II, Oxford University Press, Oxford (2000)
7. da Costa, N.C.A.: Ensaio sobre os Fundamentos da Lógica. Hucitec, São Paulo, (1980). French Translation: Logiques Classiques et Non Classiques. Masson, Paris (1997)
8. Couturat, L.: La Logique de Leibniz. D'après des Documents Inédits. Alcan, Paris (1901)
9. Couturat, L.: Histoire de la langue universelle. Hachette, Paris (1903)
10. Couturat, L.: L'Algèbre de la logique. Gauthier-Villars, Paris (1905)
11. Eisenberg, M.: Axiomatic Theory of Sets and Classes. Holt, Rinehart and Winston, New York (1971)
12. Feferman, A.B.: Politics Logic and Love. The Life of Jean Van Heijenoort. Jones and Bartlett, Boston (1993)

13. Feferman, A.B.: Jean van Heijenoort: Kaleidoscope. Log. Univers. Special Issue: Perspectives on the History and Philosophy of Modern Logic: Van Heijenoort Centenary, ed. by I.H. Anellis, **6**, 277–291 (2013)
14. Feferman, S., Feferman, A.B.: Tarski: Life and Logic. Cambridge University Press, Cambridge (2004)
15. Fraisse, R.: La zérologie: une recherche aux frontières de la logique et de l'art: applications à la logique des relations de base vide. Int. Log. Rev. **26**, 67–29 (1982)
16. Frege, G. (ed.): Philosophical and Mathematical Correspondence, Gabriel Hermes, Kambartek, Thiel and Veraart, University of Chicago Press, Chicago (1980)
17. Frutiger, A.: Signs and Symbols, Their Design and Meaning. Van Nostrand Reinhold, New York (1989)
18. Grana, N.: Sulla teoria delle valuazioni di N. C. A. da Costa. Liguori, Naples (1990)
19. Grothendieck, A.: Récoltes et Semailles – Réflexions et témoignage sur un passé de mathématicien. Unpublished
20. Grothendieck, A.: La nouvelle église universelle. Pourquoi les mathématiques, pp. 11–35. UGE, Paris (1974)
21. Guillaume, M.: Regard en arrière sur quinze années de coopération douce avec l'école brésilienne de logique paraconsistante. Log. Anal. **39**, 6–14 (1996)
22. Halmos, P.R.: I Want To Be a Mathematician. An Automathography. Springer, Berlin (1985)
23. Halmos, P.R.: How to write Mathematics. L'Enseignement Mathématique **16**, 123–152 (1970)
24. Heisenberg, W.: Der Teil und das Ganze. Gespräche im Umkreis der Atomphysik. R. Piper, Munich (1969)
25. Henkin, L., Suppes, P. and Tarski, A. (eds.).: The axiomatic method with special reference to geometry and physics. Proceedings of an international symposium held at the University of California, Berkeley, December 16, 1957–January 4, 1958. North-Holland, Amsterdam (1958)
26. Jung, C.G.: Memories, Dreams, Reflections. Pantheon book, New York (1963)
27. Kant, I.: Logik. (1800)
28. Kofman, S.: Comment s'en Sortir? Galilée, Paris (1983)
29. Le Lionnais, F.: Les Grands Courants de la Pensée Mathématique. Cahiers du Sud, Marseilles (1948)
30. Lévi-Strauss, C.: Tristes Tropiques. Plon, Paris (1955)
31. Łoś, J.: O matrycach logicznych. Travaux de la Société des Sciences et des Lettres de Wrocław, Série B, 19 (1949)
32. Manzano, M., Sain., Alonso, E. (eds.): The Life and Work of Leon Henkin. Birkhäuser, Basel (2014)
33. Mattei, J.-F.: L'Étranger et le Simulacre. Essai sur la fondation de l'ontologie platonicienne. PUF, Paris (1983)
34. Moretti, A.: The Geometry of Logical Opposition. PhD Thesis, University of Neuchâtel (2009)
35. Pascal, B.: De l'esprit Géométrique et de l'Art de Persuader (1657). English translation: On the Geometrical Spirit and the Art of Persuasion
36. Quine, W.V.O.: O Sentido da Nova Lógica. Martins, São Paulo (1948)
37. Quine, W.V.O.: Mission to Brazil. Log. Anal. **40**, 5–8 (1997)
38. Robert, A.M.: Nonstandard Analysis. Dover, New York (2003)
39. Rota, G.C.: Indiscrete thoughts. Birkhäuser, Boston (1997)
40. Rougier, L.: Traité de la Connaissance. Gauthiers-Villars, Paris (1955)
41. Rousseau, J.J.: Volume 1, Genève 1782, Volume 2, Neuchâtel, 1789 (posthumous). English translation: Rousseau: Confessions. Cambridge University Press, Cambridge (1987)
42. Rousseau, J.J.: Projet de constitution pour la Corse, 1765, posthumous publication 1825
43. Sakarovitch, J.: De la Coupe des Pierres a la Géométrie Descriptive XVI–XIX siècles. Birkhäuser, Basel (1997)
44. Searle, J.: The Mystery of Consciousness. New York Reviews of Books, New York (1997)
45. Schmid, A.F. (ed): Bertrand Russell et Louis Couturat. Correspondance sur la Philosophie, la Logique et la Politique (1897–1913). Editions Kimé, Paris (2001)
46. Schopenhauer, A.: Über die vierfache Wurzel des Satzes vom zureichenden Grunde, 1813, revised edition 1847. English translation: On the Fourfold Root of the Principle of Sufficient Reason. Open Court Publishing, Chicago (1974)

47. Slater, B.H.: Paraconsistent logics. J. Philos. Log. **24**, 451–454 (1995)
48. Sokal, A., Bricmont, J.: Impostures intellectuelles. Odile Jacob, Paris (1997). English translation: Fashionable nonsense. Picador, New York (1998)
49. Suppes, P.: Intellectual autobiography, Part I, 1922–1978. http://web.stanford.edu/~psuppes/autobio1.html
50. Weil, A.: Souvenis d'apprentissages. Birkhäuser, Basel, 1991. English Translation The Apprenticeship of a Mathematician. Birkhäuser, Basel, 1992
51. Zweig, S.: Brasilien. Ein Land der Zukunft. Bermann-Fischer, Stockholm (1941). English translation: Brazil, Land of the Future. Viking Press, New York (1941)
52. Zygmunt, J.: Life and Work of Mojżesz Presburger. Hist. Philos. Log. **12**, 211–223 (1991)
53. Zygmunt, J., Purdy, R.: Adolf Lindenbaum: Notes on his life, with bibliography and selected references. Log. Univers. **8**, 285–320 (2014)

J.-Y. Béziau (✉)
University of Brazil, Rio de Janeiro, Brazil
e-mail: jean-yves.beziau@logica-universalis.org

A Quantitative-Informational Approach to Logical Consequence

Marcos Antonio Alves and Itala M. Loffredo D'Ottaviano

Abstract In this chapter, we propose a definition of logical consequence based on the relation between the quantity of information present in a particular set of formulae and a particular formula. As a starting point, we use Shannon's quantitative notion of information, founded on the concepts of logarithmic function and probability value. We first consider some of the basic elements of an axiomatic probability theory, and then construct a probabilistic semantics for languages of classical propositional logic. We define the quantity of information for the formulae of these languages and introduce the concept of informational logical consequence, identifying some important results; among them certain arguments that have traditionally been considered valid, such as *modus ponens*, are not valid from the informational perspective; the logic underlying informational logical consequence is not classical, and is at the least paraconsistent *sensu lato*; informational logical consequence is not a Tarskian logical consequence.

Keywords Logical consequence, Information, Probability, Semantics, Informational logical consequence, Nonclassical logics, Paraconsistent logic

Mathematics Subject Classification (2010) Primary 03B22 · Secondary 03B53, 03B48

1 Introduction

Logical consequence can be considered in different ways. In intuitive terms, it is often understood as a relation established between a given set of statements of a language and a statement of the same language. In logic, it is common to define it in a way that is close to intuitive, for example, as a type of relation between elements of the power-set of a nonempty set, as done by Feitosa and D'Ottaviano (2004), among others. In this sense, logical consequência is as a relation between a given set (which could be empty or even infinite) of formulae of a language (usually formal), and a formula of the same language. The statements or formulae belonging to the given set are termed *premises*, and the other statement or formula is termed the *conclusion*. The role of the premises is to found, sustain, and support the conclusion.

In the 1930s, Tarski [19–22] introduced and improved his definition of consequence operator, or logical consequence, and proved the fundamental properties of it (cf. [2]).

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_3

In Tarskian terms, given the language L of the (classical) propositional calculus, a *consequence operator* on the set of formulae of L, Form(L), is a function $C\colon \wp(\text{Form}(L)) \to \wp(\text{Form}(L))$ such that, for all $\Gamma, \Delta \subseteq \text{Form}(L)$, it satisfies the following properties:

(T1) $\Gamma \subseteq C(\Gamma)$;
(T2) If $\Gamma \subseteq \Delta$, then $C(\Gamma) \subseteq C(\Delta)$;
(T3) $C(C(\Gamma)) \subseteq C(\Gamma)$.

The first property above is named reflexivity; the second is named monotonicity, and the third property is named transitivity. A logical consequence operator that satisfies these properties is usually named a *Tarskian logical consequence*.

When suitably combined in a sequence, the set of premises and the conclusion constitute an *argument*. We can say that a conclusion is logical consequence of a set of premises if, and only if, the argument constituted by the union is logically valid, or simply valid.

In semantic terms, the relation of consequence is usually defined starting from the degree of truth of the premises and conclusion: a formula is a logical consequence of a given set of formulae if, and only if, it is true under all circumstances (e.g., valuation, structure, interpretation, and model), such that all the premises are true.

In this chapter, we propose a definition of logical consequence based on the quantity of information present in the set of premises and in the conclusion. As a starting point, we use the usual languages of classical propositional logic (CPL), as constructed by Shoenfield [18], for example.

In our theoretical approach to logical consequence, we do not consider the qualitative semantic aspects of information, considered in works such as those of Dretske [7] and Gonzalez [9]. In our case, the informational value of a message or a formula depends only on its probability of occurrence, established using probability space. In the next section, we consider some elements of a usual axiomatic theory of probabilities, indicating some of its definitions and basic results, which will be used later; these concepts include the notions of event and random experiment.

In Sect. 3, we construct a probabilistic semantics for CPL; we establish a relation between the formulae of the CPL language and the events of a random experiment, from which we define a probability value for each formula of a given language. At the end of the section, we introduce the definitions of probabilistically valid and probabilistically equivalent formulae, and of probabilistic logical consequence.

In Sect. 4, we discuss the notion of quantity of information present in a formula of a CPL language; we propose a quantitative-informational definition of logical consequence, which we call *informational logical consequence*, and we demonstrate some of the results and properties that follow from this definition. In particular, we show the existence of arguments which are considered valid according to the classical perspective, but which are invalid from the informational perspective. For example, *modus ponens* is informationally invalid, given the possibility that the conclusion of this argument could possess a greater quantity of information than its set of premises. Furthermore, we show that the logic underlying informational logical consequence is not classical, but is, at the least, paraconsistent *sensu lato*. In addition, we demonstrate that although it might satisfy the property of transitivity, informational logical consequence is neither reflexive nor monotonic; in other words, it is not a Tarskian logical consequence.

In the final considerations, we summarize and analyze the main properties and results of informational logical consequence.

Our approach is based on the quantitative concept of information, developed in the *Mathematical Theory of Communication*. One of the pioneers in studies of the quantification, storage, and transmission of information was Hartley [12]. He describes the quantity of information present in a source in terms of its number of possible messages. Later, Shannon [16] further developed this idea, including new factors such as the effect of noise in the channel, possible economy in the transmission of information, and the possibility that messages might possess distinct quantities of information.

From the quantitative perspective, there can only be information where there is doubt; this, in turn, requires the existence of alternatives, which presupposes the presence of choice, selection, and discrimination. For Hartley [12], the information in a message is measured by the freedom of choice that someone has in selecting it, based on a source. According to Shannon and Weaver [17, pp. 8–9], "... information in communications theory relates not so much to what you *do* say, as to what you *could* say. That is, information is a measure of one's freedom of choice when one selects a message."

In an unbiased toss of a coin, for example, there are two equally probable possibilities: heads or tails. In an unbiased throw of a dice, there are six possibilities. The degree of freedom of choice in the first case is less than in the second. In the case of the dice, we could say many more things than would be possible in the case of the coin. Hence, from the present perspective, the quantity of information present in the throw of the dice is greater than that present in the toss of the coin.

For Hershberger [13], information can be defined as a measure of the reduction of uncertainty. It is related to the unpredictability in a message or a source, with the emergence of an element that was absent prior to its occurrence. In the toss of the coin, the reduction of uncertainty is less than in the throw of the dice. The occurrence of an event in a source such as the example of the toss of a coin only eliminates one alternative, while in the throw of a dice five equally probable alternatives are eliminated. Information is also associated with notions such as those of order (or organization, in the terms of Bresciani Filho and D'Ottaviano [3]) and entropy, as discussed by Alves [1].

The central interest in quantitative studies of information generally lies in measurement of the quantity of information in a source, rather than in particular messages. In a broad sense, a *source* can be characterized as a process that generates information. Its constituent elements can be understood as a finite set of events, of messages, of symbols that have a certain probability of occurrence. *Discrete ergodic sources* are those in which every element produced, in addition to being discrete, has the same statistical properties as any other, and its properties remain unaltered with time. Once the probabilities of occurrence of the elements are discovered, it becomes possible to predict its probability of occurrence at any moment, as occurs in the toss of a coin or the throw of a dice.

According to Shannon and Weaver [17], the *quantity of information of the i-th message* of a source F, denoted by $I_{i\,(\mathrm{F})}$, is the numerical value defined by

$$I_{i\,(\mathrm{F})} =_{\mathrm{df}} -\log_2 p_{i\,(\mathrm{F})}\,,$$

where "$p_{i\,(\mathrm{F})}$" denotes the probability of occurrence of the ith message of F.

The *quantity of information in a source* F with n elements, denoted by $\mathrm{H_F}$, is defined by

$$\mathrm{H_F} =_{\mathrm{df}} \sum_{i=1}^{n} p_{i\,(\mathrm{F})} \times I_{i\,(\mathrm{F})}\,.$$

If $\mathrm{H_C}$ and $\mathrm{H_D}$ denote the quantities of information in the throw of the coin and dice, respectively, then $\mathrm{H_C} = 1$ and $\mathrm{H_D} \approx 2.58$. Since, in both cases, the events corresponding to the messages are equiprobable, it can be demonstrated that the quantity of information in each message in the source is equal to the quantity of information in the source itself.

The greater the freedom of choice and reduction in uncertainty in a source, the more informative it is. Information reaches its maximum value in a source when all its messages have an equal chance of being selected. The informational value is zero when only one of them can occur. It is within this framework that in the next sections we propose our informational perspective of logical consequence.

2 Elements of an Axiomatic Theory of Probability

We employ probability in cases where two or more different results can occur in a given circumstance. This means that the result is not predictable (or is indeterminate), in the sense that it is not possible to previously determine which result might occur at a given moment.

Probability theory, henceforth denoted $\mathbb{P}$, studies random experiments. As a basis for $\mathbb{P}$, we shall use the Zermelo–Fraenkel set theory with the *choice axiom* (ZFC), with the usual elementary arithmetic theory (cf. [8]). Hence, the language (alphabet and definitions) and theorems of ZFC will also be considered elements of $\mathbb{P}$. The only symbols belonging to the alphabet of $\mathbb{P}$ are the symbols A_i, for $0 \leq i \leq n$ and $i \in \mathrm{N}$, where "A" represents the primitive concept of $\mathbb{P}$ known as *happening*, *result*, or *occurrence*, used for the definition of random experiment.

Definition 2.1 (Random Experiment)

(a) A *random experiment* or *random phenomenon*, denoted by "Σ," is one that, repeated various times, presents different results or occurrences, called *results of* Σ or *occurrences of* Σ, denoted by "$A_i(\Sigma)$."
(b) The *sample space* of a random experiment Σ, denoted by "$U(\Sigma)$," is the set of all possible results of Σ.
(c) The *number of elements* of the sample space, denoted by "$n(U(\Sigma))$," in which $n(U(\Sigma)) > 0$ is finite, is the quantity of elements of "$U(\Sigma)$."
(d) A sample space is *equiprobable* when all its elements have the same chance of occurring.

The toss of a coin, the throw of a dice, and the removal of a card from a pack are examples of random experiments. Their results would be the fall of the coin with one or other face upwards, the fall of the dice with one of the numbers from one to six upwards, and the withdrawal of one of the cards of the pack, respectively.

Definition 2.2 (Event of a Random Experiment)

(a) An *event* of a random experiment Σ, denoted by "$E(\Sigma)$," is any subset of the sample space $U(\Sigma)$, in other words, $E(\Sigma) \subseteq U(\Sigma)$.
(b) The *number of elements* of an event $E(\Sigma)$, denoted by "$n(E(\Sigma))$," is the quantity of elements of $U(\Sigma)$ belonging to $E(\Sigma)$.
(c) An *elementary event* E *of* Σ is that where $n(E(\Sigma)) = 1$.
(d) The *correct event* E *of* Σ is that where $n(E(\Sigma)) = n(U(\Sigma))$.
(e) The *impossible event* E *of* Σ, denoted by "$\emptyset$," is that where $n(E(\Sigma)) = 0$.
(f) A *contingent event* E *of* Σ is that where $n(\emptyset) < n(E(\Sigma)) < n(U(\Sigma))$.
(g) The *event* E_i *of* Σ, *complementary to* $E(\Sigma)$, denoted by "$\bar{E}(\Sigma)$," is defined by

$$\bar{E}(\Sigma) =_{\mathrm{df}} \{A \in U(\Sigma) | A \notin E(\Sigma)\}.$$

(h) The *event* E *of* Σ, *the union of* $E_i(\Sigma)$ *and* $E_j(\Sigma)$, denoted by "$(E_i(\Sigma) \cup E_j(\Sigma))$," is defined by

$$(E_i(\Sigma) \cup E_j(\Sigma)) =_{\mathrm{df}} \{A \in U(\Sigma) | A \in E_i(\Sigma) \text{ou}\, A \in E_j(\Sigma)\}.$$

(i) The *event* E *of* Σ, *the intersection of* $E_i(\Sigma)$ and $E_j(\Sigma)$, denoted by "$(E_i(\Sigma) \cap E_j(\Sigma))$," is defined by

$$(E_i(\Sigma) \cap E_j(\Sigma)) =_{\mathrm{df}} \{A | A\} \in E_i(\Sigma) \text{e} A \in E_j(\Sigma)\}.$$

Henceforth, when there is no risk of ambiguity, we shall remove the references to Σ between parentheses of the notations. Hence, instead of $A(\Sigma)$, $U(\Sigma)$, or $n(U(\Sigma))$, we shall use only A, U, or $n(U)$, respectively.

Example 2.3 (Random Experiments)

Σ_i	$U(\Sigma_i)$	$E(\Sigma_i)$	$n(E(\Sigma_i))$
Σ_1 (Toss of coin)	{H, T}	$E_1(\Sigma_1)$: {H}(Fall head) $E_2(\Sigma_1)$: {T}(Fall tail)	1 1
Σ_2 (Toss of biased coin)	$\{H_1, H_2, H_3, T\}$	$E_1(\Sigma_2)$: $\{H_1, H_2, H_3\}$(Fall head) $E_2(\Sigma_2)$: {T}(Fall tail)	3 1
Σ_3 (Throw of dice)	{1, 2, 3, 4, 5, 6}	$E_1(\Sigma_3)$: {2, 4, 6}(Fall even) $E_2(\Sigma_3)$: $\emptyset$ (Fall head)	3 0

In the above example, a "model" is proposed for each random experiment, its sample space, and some of its events, associating them with entities of a "world." In the third column, in parentheses, we give the common name for each event in order to express the results that comprise it.

Although our theoretical approach does not consider qualitative elements of information, the examples suggested throughout this chapter involve content, in order to aid understanding. Meanwhile, the same results could have been obtained using a purely formal mode of construction.

The notion of random experiment can be compared with that of a discrete ergodic source of information, as indicated in Sect. 1. In a way that is similar to this type of source, the elements belonging to a random experiment must be previously defined with precision. Furthermore, every sample space considered, in addition to being finite, must be equiprobable, and the probability values of the events must be given and fixed. This will enable us to determine with precision the existence of the relation of logical consequence between formulae (as shown in Sect. 4).

Definition 2.4 (Probability of an Event) The *probability of occurrence* of an event E in the random experiment Σ with an equiprobable sample space $U(\Sigma)$, denoted by "$p(E(\Sigma))$," is the numerical value defined by

$$p(E(\Sigma)) =_{\text{df}} \frac{n(E(\Sigma))}{n(U(\Sigma))}.$$

The *probability function*, p, provides events of a random experiment with values between 0 and 1; in other words, p: $E(\Sigma) \to [0, 1] \subseteq \mathbb{Q}$. The *probability value* of an event $E(\Sigma)$ is given by $p(E(\Sigma))$.

Example 2.5 (Probability of the Events of Example 2.3)

E	$E_1(\Sigma_1)$	$E_2(\Sigma_1)$	$E_1(\Sigma_2)$	$E_2(\Sigma_2)$	$E_1(\Sigma_3)$	$E_2(\Sigma_3)$	$E_1(\Sigma_3) \cup E_2(\Sigma_3)$	$E_1(\Sigma_2) \cap E_2(\Sigma_2)$	$\bar{E}_1(\Sigma_2)$
$p(E)$	1/2	1/2	3/4	1/4	1/2	0	1/2	0	1/4

Having constituted the basic elements of the language of $\mathbb{P}$, we now describe its axioms and some of its elementary results.

The axioms for $\mathbb{P}$ are as follows

(Ax$\mathbb{P}$1): $p(E) \geq 0$, for every $E \subseteq U$
(Ax$\mathbb{P}$2): $p(E_i \cup E_j) = p(E_i) + p(E_j) - p(E_i \cap E_j)$
(Ax$\mathbb{P}$3): $p(E \cup \bar{E}) = 1$.

In the next theorem, some results concerning the probability of events are stated, being important to the continuation of this chapter. For simplicity, we do not present the proof of this, and some other theorems. Their proofs can be found in Alves [2].

Theorem 2.6

(a) $p(E) \leq 1$, for every $E \subseteq U$.
(b) $p(U) = 1$.

(c) $p(\emptyset) = 0$.
(d) $n(E_i \cap E_j) = 0 \Rightarrow p(E_i \cup E_j) = p(E_i) + p(E_j)$.
(e) $p(E \cap \bar{E}) = 0$.
(f) $p(\bar{E}) = 1 - p(E)$.
(g) $\sum_{i=1}^{m} p(E_i) = 1$, for $E_i = \{A_i\}$, for $1 \leq i \leq m$, with $U = \{A_1, \ldots, A_m\}$.

Based on $\mathbb{P}$, in what follows we develop a probabilistic semantics for CPL.

3 A Probabilistic Semantics for Languages of CPL

We call this perspective the *probabilistic semantics for CPL* (henceforth, $S_{\mathbb{P}}$). As shown by Alves [2], the behavior of $S_{\mathbb{P}}$ is not strictly equivalent to the behavior of the usual classical veritative-functional semantics (henceforth, S_{V}), as developed by Tarski [19], Mendelson [15], Shoenfield [18], and Mates [14], among others.

We associate the formulae of a CPL language, denoted by "L," with the events of a random experiment, Σ. We define some of the notions that are fundamental to the objectives of this chapter, and describe some results characteristic of $S_{\mathbb{P}}$.

The expression "Form(L)" denotes the set of formulae of a language L, the letters "φ," "ψ," and "γ" are metalinguistic variables that represent elements of Form(L); "P_0," "P_1," "P_2," etc., are the atomic formulae of L, and "Γ" represents any finite subset of Form(L). We adopt negation and disjunction as primitive logical connectives of L.

Definition 3.1 (Situation for a Language *L*) A function f is a *Σ-situation for L*, or simply a *situation*, denoted by "$f(\Sigma)$," if $f(\Sigma)$: Form(L) $\rightarrow \wp(U(\Sigma))$, such that

(a) If φ is atomic, then $f(\Sigma)(\varphi) = E(\Sigma)$, defined by f itself;
(b) If φ is of the form $\neg\psi$, then $f(\Sigma)(\varphi) = f(\Sigma)(\psi)$;
(c) If φ is of the form $\psi \vee \gamma$, then $f(\Sigma)(\varphi) = f(\Sigma)(\psi) \cup f(\Sigma)(\gamma)$.

A *situation* for L consists of an attribution of a single event of a given random experiment to each well-formed formula of L, according to a function f. The fact that a situation is defined using a function enables us to avoid ambiguities in the next definitions, especially in the case of informational logical consequence.

To say that *f is an Σ-situation for L* is the same as saying that the random experiment *Σ is an f-structure for L.*

Although each formula of L is associated with a single event in a given situation $f(\Sigma)$, distinct formulae can be associated with the same event in $f(\Sigma)$. This always occurs, given that the set of formulae of any L is infinite, in contrast to the number of events of a random experiment, which is always finite.

Definition 3.2 (Probability of Formulae According to a Situation) The *probability function of a formula* φ according to $f(\Sigma)$, denoted by "$P(f(\Sigma))$," is defined as follows, where "p" is the probability function concerning events as defined in $\mathbb{P}$:

(a) If φ is atomic, $P(f(\Sigma))(\varphi) =_{\text{df}} p(f(\Sigma)(\varphi))$;

(b) If φ is of the form $\neg\psi$, $P(f(\Sigma))(\varphi) =_{\mathrm{df}} p(f(\Sigma)(\psi))$;
(c) If φ is of the form $\psi \vee \gamma$, $P(f(\Sigma))(\varphi) =_{\mathrm{df}} p(f(\Sigma)(\psi) \cup f(\Sigma)(\gamma))$.

It can be seen that $P(f(\Sigma))$: Form$(L) \rightarrow [0, 1] \subseteq \mathbb{Q}$, such that $p(f(\Sigma))(\varphi)$ is the image of φ according to $P(f(\Sigma))$. The probability value of the formula φ according to $f(\Sigma)$ is given by $P(f(\Sigma))(\varphi)$, a rational number between 0 and 1. It is the probability value of the event of Σ corresponding to φ, according to $f(\Sigma)$.

It can be shown that P($f(\Sigma)$) satisfies the properties described in Theorem 2.6, interpreted in the light of $S_{\mathbb{P}}$.

Henceforth, when we say "for every f," we mean "for every situation $f(\Sigma)$, given Σ." In addition, when there is no possibility of ambiguity or imprecision, we shall use "$P(\varphi)$" as an abbreviation for "$P(f(\Sigma))(\varphi)$," "$p(\varphi)$" as an abbreviation for "$p(f(\Sigma)(\varphi)$," "$f(\varphi)$" as an abbreviation for "$f(\Sigma)(\varphi)$," and "f" as an abbreviation for "$f(\Sigma)$."

In accordance with the usual definitions in L, we have

(a) $f(\varphi \wedge \psi) = f(\neg(\neg\varphi \vee \neg\psi))$
(b) $f(\varphi \rightarrow \psi) = f(\neg\varphi \vee \psi)$
(c) $f(\varphi \leftrightarrow \psi) = f((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$
(d) $P(\varphi \wedge \psi) = P(\neg(\neg\varphi \vee \neg\psi))$
(e) $P(\varphi \rightarrow \psi) = P(\neg\varphi \vee \psi)$
(f) $P(\varphi \leftrightarrow \psi) = P((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$.

Example 3.3 (Probability of Formulae in Σ_1 and Σ_2, for $U(\Sigma_1) = \{1, 2, 3, 4, 5, 6\}$ and $U(\Sigma_2) = \{\mathrm{H}, \mathrm{T}\}$)

φ	$f(\Sigma_1)(\varphi)$	$P(f(\Sigma_1))(\varphi)$	$f(\Sigma_2)(\varphi)$	$P(f(\Sigma_2))(\varphi)$
P_1	$\{2, 4, 6\}$	$1/2$	$\{\mathrm{H}\}$	$1/2$
P_2	$\{1, 3, 5\}$	$1/2$	$\{\mathrm{T}\}$	$1/2$
P_3	$\{1, 2, 3, 5\}$	$2/3$	$\emptyset$	0
P_4	$\{1\}$	$1/6$	$\emptyset$	0
P_5	$\emptyset$	0	$\emptyset$	0
$P_1 \wedge P_2$	$\emptyset$	0	$\emptyset$	0
$P_1 \wedge P_3$	$\{2\}$	$1/6$	$\emptyset$	0
$P_2 \wedge P_3$	$\{1, 3, 5\}$	$1/2$	$\emptyset$	0
$P_1 \vee P_3$	$\{U\}$	1	$\{\mathrm{H}\}$	$1/2$
$\neg(P_1 \vee P_3)$	$\emptyset$	0	$\{\mathrm{T}\}$	$1/2$
$P_1 \rightarrow P_2$	$\{1, 3, 5\}$	$1/2$	$\{\mathrm{T}\}$	$1/2$

Theorem 3.4 For every f, we have

(a) $P(\neg\varphi) = 1 - P(\varphi)$.
(b) $f(\varphi \wedge \psi) = f(\varphi) \cap f(\psi)$.
(c) $P(\varphi \wedge \psi) = p(f(\varphi) \cap f(\psi))$.
(d) $P(\varphi \rightarrow \psi) = p(f(\varphi) \cup f(\psi))$.

Definition 3.5 (Inconsistency, Validity, and Contingency of Formulae of L in $S_{\mathbb{P}}$)

(a) A formula φ is *probabilistically inconsistent* or *probabilistically contradictory*, which is denoted by "$\bot_{\mathbb{P}}$," if, for every $f(\Sigma)$, $P(f(\Sigma))(\varphi) = 0$.
(b) A formula φ is *probabilistically valid* or *probabilistically tautological*, which is denoted by "$\top_{\mathbb{P}}$," if, for every $f(\Sigma)$, $P(f(\Sigma))(\varphi) = 1$.
(c) A formula φ is *probabilistically contingent* if it is neither probabilistically contradictory nor probabilistically valid.

Definition 3.6 (Probabilistically Equivalent Formulae) Two formulae, φ and ψ, are *probabilistically equivalent*, which is denoted by "$\varphi \equiv \psi$," if, for every situation $f(\Sigma)$, $P(f(\Sigma))(\varphi) = P(f(\Sigma))(\psi)$.

Definition 3.7 (Probability Value of a Set of Formulae) Let $\Gamma = \{\varphi_1, \dots, \varphi_n\} \subseteq \text{Form}(L)$. The *probability value of Γ according to $f(\Sigma)$*, denoted by "$P^*(f(\Sigma))(\Gamma)$," is defined by

$$P^*(f(\Sigma))(\Gamma) =_{\text{df}} P(f(\Sigma))(\varphi_1 \wedge \dots \wedge \varphi_n)\,.$$

If $\Gamma = \emptyset$, then $P^*(f(\Sigma))(\Gamma) = 1$. When Γ is infinite, $P^*(f(\Sigma))(\Gamma)$ is indefinite.

It can be seen that $P^*(f(\Sigma))$: $\wp(\text{Form}(L)) \to [0, 1] \subseteq \mathbb{Q}$, such that $P^*(f(\Sigma))(\Gamma)$ is the image of Γ according to $P^*(f(\Sigma))$. The *probability value of the set of formulae Γ according to $f(\Sigma)$* is given by $P^*(f(\Sigma))(\Gamma)$, a rational number between 0 and 1. When there is no risk of ambiguity or imprecision, we shall write $P(\Gamma)$ instead of $P^*(f(\Sigma))(\Gamma)$.

According to Definition 3.7, the probability value of a given set of formulae of L in a situation f is, in the final analysis, defined from the probability value of the intersection of the events associated with the elements of Γ. In other words, by Definition 3.7, and Theorem 3.4c, we have that $P^*(f(\Sigma))(\Gamma) = P(f(\Sigma))(\varphi_1 \wedge \dots \wedge \varphi_n) = p(f(\Sigma)(\varphi_1) \cap \dots \cap f(\Sigma)(\varphi_n))$. Given the definitions in question, we can say that $f(\Sigma)(\Gamma) = f(\Sigma)(\varphi_1 \wedge \dots \wedge \varphi_n)$.

Strictly speaking, it is not possible to define the probability value of an empty set of formulae. Definition 3.2, which forms the basis of Definition 3.7, only applies to the probability values of formulae of language L. But $\emptyset$ is not a formula of L. Although a formula might be associated with the empty event, with its probability value being defined as zero, it makes no sense to say that an empty formula can have any probability value.

In order not to leave the empty set without a definition of the probability value, it was decided to define it arbitrarily in the way described above. This decision can be justified as follows: in this case, to say that $P(\Gamma) \neq 1$ would signify the existence of some $\varphi_i \in \Gamma$ such that $P(\varphi_i) \neq 1$. Since there are no formulae in Γ with this value, given the inexistence of formulae in Γ, then $P(\Gamma) = 1$.

Definition 3.8 (Probabilistic Logical Consequence) A formula φ is *probabilistic logical consequence* of a set Γ of formulae, which is denoted by "$\Gamma \models_{\mathbb{P}} \varphi$," if, for every situation $f(\Sigma)$, $P(f(\Sigma))(\Gamma) \leq P(f(\Sigma))(\varphi)$.

When $\Gamma = \emptyset$, instead of $\Gamma \models_{\mathbb{P}} \varphi$ we simply write $\models_{\mathbb{P}} \varphi$. The expression "$\Gamma \not\models_{\mathbb{P}} \varphi$" denotes that φ is not probabilistic logical consequence of Γ. The formulae of Γ are called *premises* and φ is termed *conclusion*.

Despite possessing certain specific characteristics, $S_{\mathbb{P}}$ is a semantics for CPL. The formulae considered valid by $S_{\mathbb{P}}$ are exactly the same as those considered valid by S_{V}. Furthermore, $\Gamma \models_{\mathbb{P}} \varphi$ if, and only if, $\Gamma \models_{\mathrm{V}} \varphi$. The similarities and differences between $S_{\mathbb{P}}$ and S_{V} are described by Alves [2]. In the following discussion, we address the notion of informational logical consequence.

4 Informational Logical Consequence

In this section, we propose a quantitative-informational definition of logical consequence and present some of its properties. We first introduce notions such as the quantity of information present in a formula of L and in a set of formulae, based on the notion of quantity of information developed by Shannon and Weaver [17].

Definition 4.1 (Quantity of Information of a Formula According to a Situation) The *quantity of information* or *informational value of a formula* φ *of* L *according to a situation* $f(\Sigma)$, denoted by "$I(f(\Sigma))(\varphi)$," is the numerical value defined by

$$I(f(\Sigma))(\varphi) =_{\mathrm{df}} -\log_2 P(f(\Sigma))(\varphi)\,.$$

When $P(f(\Sigma))(\varphi) = 0$, we define that $\log_2 0 = 0$, in other words, $I(f(\Sigma))(\varphi) = 0$. When there is no risk of ambiguity or imprecision, we shall use "$I(\varphi)$" instead of "$I(f(\Sigma))\varphi$." It can be shown that $I(f(\Sigma))$: Form$(L) \to \mathbb{Q}_+$.

Example 4.2 (Quantity of Information in Formulae Based on the Two Situations in Example 3.3)

φ	$f(\Sigma_1)(\varphi)$	$I(f(\Sigma_1))\varphi$	$f(\Sigma_2)(\varphi)$	$I(f(\Sigma_2))\varphi$
P_1	$\{2, 4, 6\}$	1	{H}	1
P_2	$\{1, 3, 5\}$	1	{T}	1
P_3	$\{1, 2, 3, 5\}$	0.58	$\emptyset$	0
P_4	$\{1\}$	2.58	$\emptyset$	0
$\neg P_1$	$\{1, 3, 5\}$	1	{T}	1
$P_1 \wedge P_3$	$\{2\}$	2.45	$\emptyset$	0
$P_1 \vee P_3$	U	0	{H}	1
$P_1 \to P_2$	$\{1, 3, 5\}$	1	{T}	1
$P_1 \to P_3$	$\{1, 2, 3, 5\}$	0.58	{T}	1
$P_2 \to P_3$	U	0	{H}	1

Definition 4.3 (Quantity of Information in a Set of Formulae) Let $\Gamma = \{\varphi_1, \ldots, \varphi_n\} \subseteq \text{Form}(L)$. The *informational value of* Γ (*quantity of information in* Γ), *according to* $f(\Sigma)$, denoted by "$I^*(f(\Sigma))(\Gamma)$," is defined by

$$I^*(f(\Sigma))(\Gamma) =_{\text{df}} -\log_2 P^*(f(\Sigma))(\Gamma)\,.$$

When there is no risk of ambiguity or imprecision, we shall use $I(\Gamma)$ instead of $I^*(f(\Sigma))(\Gamma)$. Although the domain of this function may be different from that of the information function concerning formulae, its image set is the same; in other words, $I^*(f(\Sigma))\colon \wp(\text{Form}(L)) \to \mathbb{Q}_+$.

Definition 4.4 (Informational Logical Consequence) A formula φ is *informational logical consequence* of a set Γ of sentences, which is denoted by "$\Gamma \mid\models \varphi$," if, for every $f(\Sigma)$, $I(f(\Sigma))(\Gamma) \geq I(f(\Sigma))(\varphi)$.

When $\Gamma = \emptyset$, instead of "$\Gamma \mid\models \varphi$," we simply write "$\mid\models \varphi$"; "$\Gamma \not\mid\models \varphi$" denotes that φ is not informational logical consequence of Γ. The formulae of Γ are called *premises* and φ is termed *conclusion*.

According to the above definition, a formula is informational logical consequence of a given set of formulae if, and only if, the quantity of information present in the conclusion is never greater than the quantity of information in the premises. In probabilistic logical consequence, the relation is inverse.

Theorem 4.5

(a) $\varphi \to \psi \mid\models \neg\varphi \vee \psi$.
(b) $\varphi \vee \psi \mid\models \neg(\neg\varphi \wedge \neg\psi)$.
(c) $\varphi \wedge \psi \mid\models \neg(\neg\varphi \vee \neg\psi)$.
(d) $\varphi \wedge \psi \mid\models \psi \wedge \varphi$.
(e) $\varphi \vee \psi \mid\models \psi \vee \varphi$.
(f) $\neg\neg\varphi \mid\models \varphi$.
(g) $\varphi \wedge \varphi \mid\models \varphi$.
(h) $\varphi \vee \varphi \mid\models \varphi$.
(i) $\top_{\mathbb{P}} \wedge \varphi \mid\models \varphi$.
(j) $\top_{\mathbb{P}} \vee \varphi \mid\models \top_{\mathbb{P}}$.
(k) $\bot_{\mathbb{P}} \wedge \varphi \mid\models \bot_{\mathbb{P}}$.
(l) $\bot_{\mathbb{P}} \vee \varphi \mid\models \varphi$.
(m) $\top_{\mathbb{P}} \mid\models \bot_{\mathbb{P}}$.
(n) $\varphi \mid\models \varphi$.

Since the quantity of information in the premise and conclusion of each one of the items of the above theorem is the same for each given situation, the reciprocal of each one of the items is also valid, such that one formula is informational logical consequence of the other.

The first three items of the above theorem show that the definitions of one connective, obtained from others, are maintained in informational logical consequence. The first item shows that the notion of implication presupposed here is that of material implication. The fourth and sixth items show, respectively, that the logic underlying informational logical consequence is neither temporal logic nor intuitionistic logic.

Theorem 4.6

(a) $|\models \varphi \Leftrightarrow I(\varphi) = 0$, for every f.
(b) $|\models \varphi \Leftrightarrow \Gamma \; |\models \varphi$, for every Γ.
(c) $I(\Gamma) = 0$, for every f, and $\Gamma \; |\models \varphi \Rightarrow |\models \varphi$.
(d) $\models_{\mathbb{P}} \varphi \Rightarrow |\models \varphi$.

Proposition 4.7

(a) $\varphi \wedge \psi \not{|\models} \varphi$.
(b) $\varphi \rightarrow \psi, \varphi \not{|\models} \psi$.
(c) $\varphi \rightarrow \psi, \neg\psi \not{|\models} \neg\varphi$.
(d) $\varphi \rightarrow \psi, \psi \rightarrow \gamma \not{|\models} \varphi \rightarrow \gamma$.
(e) $\varphi \not{|\models} \psi \rightarrow \varphi$.
(f) $\psi \not{|\models} \varphi \vee \psi$.
(g) $\gamma \vee \varphi, \neg\gamma \vee \psi \not{|\models} \varphi \vee \psi$.

The above proposition shows that the rules of inference of a large part of formal logical systems are not valid in informational terms. The second item is the rule of *modus ponens*, adopted in systems such as that of Mendelson [15]; the last two items are the expansion rule and the cut rule, adopted by Shoenfield [18]. The same can be said when the above items are treated as arguments. Arguments that are traditionally considered valid can present more information in the conclusion than in the set of premises, which means that, according to the perspective in question, they act to amplify information.

As shown by Alves [2], the invalidity of these rules of inference or arguments is generally due to the possibility of the set of premises being informationally empty in a given situation. In the case of the expansion rule, when, in a given situation Σ, $P(\psi) = 0$ and $I(\varphi) \neq 0$, then $I(\psi) < I(\varphi \vee \psi)$. Hence, $\psi \not{|\models} \varphi \vee \psi$. While the premise is informationally empty, the conclusion produces novelty and reduces uncertainty. For example, in the game of dice, the sentence "it fell on number seven," for which the probability of occurrence is zero, would not reduce the uncertainty about what occurred in the game. Meanwhile, the sentence "it fell on an even number or on number seven" possesses a quantity of information that is greater than zero, since it reduces uncertainty: the dice could have fallen on numbers two, four, or six, eliminating the possibility of having fallen on an odd number. Thus, in the conclusion, there is something that did not exist in the premise. There was an informational gain, given that the quantity of information in the premise was null.

In the case of *modus ponens*, the situation $f(\Sigma_1)$ of Example 4.2 above provides an example in which the informational value of the premises, interpreted as "if it falls on evens, then it falls on odds" and "it falls on evens," is less than the informational value of the conclusion, interpreted as "it falls on odds." Here, the informational value of the premises is null. Since the events "fall on odds" and "fall on evens" are mutually exclusive, the sentence "if it falls on evens, then it falls on odds" is equivalent to "it falls on odds or it falls on odds," which is equivalent to "it falls on odds." The probability of the set of premises is therefore defined from the union of "it falls on odds" and "it falls on evens," equivalent to "it falls on odds and does not fall on odds," for which the probability is zero. Meanwhile, the conclusion possesses a quantity of information that is greater than zero, since it possesses an informational value that is greater than the value of the set of premises.

The example discussed in the preceding paragraph, which is an individual case of Theorem 4.8b, outlined below, seems to fit an intuitive notion of informational logical consequence. Intuitively, "fall on odd and not fall on odd" provides no information concerning a circumstance. Hence, any probabilistically contingent conclusion can contain more information than is contained by the premises.

Several steps of the demonstrations of the following theorems have been omitted. These steps, indicated by "TPP," refer to theorems previously proved in Alves [2].

Theorem 4.8 Let $\Gamma = \{\varphi_1, \ldots, \varphi_n\}$. Then:

(a) $\Gamma \mid\models \varphi$ if and only if $I(\varphi) = 0$, for every f, or $(\varphi_1 \wedge \ldots \wedge \varphi_n) \equiv \varphi$;
(b) If $I(\varphi) > 0$, for a given f, then $\bot_{\mathbb{P}} \not\mid\models \varphi$;
(c) $\Gamma \mid\models \top_{\mathbb{P}}$;
(d) $\Gamma \mid\models \bot_{\mathbb{P}}$;
(e) If $I(\varphi) > 0$, for a given f, then $\top_{\mathbb{P}} \not\mid\models \varphi$.

Proof

(a) **(⇒) :** Let $\Gamma \mid\models \varphi$. Suppose that $I(\varphi) \neq 0$, for a given f, and $(\varphi_1 \wedge \ldots \wedge \varphi_n) \not\equiv \varphi$. It is then possible to show the existence of f' such that $I(f')(\Gamma) < I(f')(\varphi)$, contradicting the initial hypothesis: $I(f')(\Gamma) = 0$ and $I(f')(\varphi) = I(f)(\varphi)$. Hence, when $\Gamma \mid\models \varphi$ we have that if $I(\varphi) \neq 0$, for a given f, then $(\varphi_1 \wedge \cdots \wedge \varphi_n) \equiv \varphi$, in other words, $I(\varphi) = 0$, for every f, or $(\varphi_1 \wedge \cdots \wedge \varphi_n) \equiv \varphi$.
(⇐) : Case 1: Let $I(\varphi) = 0$, for every f. Then, by TPP, $I(\Gamma) \geq I(\varphi)$, for every f and every Γ. So, by Definition 4.4, $\Gamma \mid\models \varphi$.
Case 2: By Definition 3.6, $(\varphi_1 \wedge \cdots \wedge \varphi_n) \equiv \varphi$ if and only if $P(\varphi_1 \wedge \cdots \wedge \varphi_n) = P(\varphi)$, for every f. By Definition 3.7, $P(\varphi_1 \wedge \cdots \wedge \varphi_n) = P(\varphi)$, for every f if and only if $P(\Gamma) = P(\varphi)$, for every f. Then, by TPP, $I(\Gamma) = I(\varphi)$, for every f. So, by Definition 4.4, $\Gamma \mid\models \varphi$.
(b) Let $P(\varphi) = 1/2$. Then $I(\bot_{\mathbb{P}}) = 0 < I(\varphi) = 1$.
(c) Since $I(\top_{\mathbb{P}}) = 0$, then $I(\Gamma) \geq I(\top_{\mathbb{P}})$.
(d) Since $I(\bot_{\mathbb{P}}) = 0$, then $I(\Gamma) \geq I(\bot_{\mathbb{P}})$.
(e) Let $P(\varphi) = 1/2$. Then $I(\top_{\mathbb{P}}) = 0 < I(\varphi) = 1$. □

The first item above describes the arguments that are valid according to the informational perspective of logical consequence.

The second item explains that a probabilistically contradictory formula cannot lead informationally to any informative formula. In fact, only formulae that are probabilistically valid or invalid, in other words not informative, are informational logical consequence of contradictory formulae.

Presupposing the distinction between classical and nonclassical systems, as suggested by Da Costa [4, 5], or by Haack [11], we can conclude from the second item of the above theorem that *classical logic is not the logic underlying informational logical consequence*. This is because, in classical formal logical systems, a contradiction generates any formula. Considering that complementary logics, such as modal logic, retain the same principles of classical logic, we can also conclude that *no complementary logic underlies informational logical consequence*. The remaining candidates are heterodox logics, such as intuitionistic and paraconsistent systems, as described by D'Ottaviano [6].

In intuitionistic logics, negation possesses certain particular characteristics. Such characteristics do not permit, for example, recourse to proofs employing reduction to the absurd, given that formulae such as $\varphi \leftrightarrow \neg\neg\varphi$ are not valid in these systems. Meanwhile, it can be shown that $\neg\neg\varphi \mid\models \varphi$ and $\varphi \mid\models \neg\neg\varphi$ and, from Theorem 4.9b below, we have $\mid\models \varphi \leftrightarrow \neg\neg\varphi$. Therefore, *intuitionistic logic cannot provide a basis for informational logical consequence*.

Theorem 4.8, especially the second item, indicates that the logic underlying informational logical consequence is, at least, *paraconsistent sensu lato*. This is because this notion of consequence does not permit, for example, the *Ex Falso Quodlibet*, also known as the *Explosion Principle*, such that from a contradiction does not follow any formula.

The third item of the above theorem is shared by both probabilistic and veritative-functional logical consequence, but for different reasons. Informationally, a probabilistically valid formula is logical consequence of any formula, because its value is minimal; hence, it cannot provide more information than any other formula. In probabilistic and veritative-functional terms, it is logical consequence of any formula because its value is maximal; in other words, it possesses the probabilistic value "1" or the true value "V."

The fourth item, like the second, is not valid in the other two notions of logical consequence considered here. In these notions, an inconsistent formula is logical consequence solely of a contradictory set of formulae. In the informational version, it is logical consequence of any set of formulae. In informational terms, there is no distinction between valid and inconsistent formulae.

The fifth result expresses a similarity between informational logical consequence and the other notions of consequence: contingent formulae are not logical consequence of valid formulae.

Theorem 4.9

(a) If $\varphi \mid\models \psi$, and $\psi \mid\models \gamma$, then $\varphi \mid\models \gamma$.
(b) If $\varphi \equiv \psi$, then $\mid\models \varphi \leftrightarrow \psi$.
(c) Is not the case that: $\Gamma, \varphi \mid\models \psi$ if and only if $\Gamma \mid\models \varphi \to \psi$.
(d) Is not the case that: $\mid\models \varphi$ *e* $\mid\models \psi$ if and only if $\varphi \equiv \psi$.
(e) Is not the case that: $\Gamma \mid\models \varphi$ if and only if $f(\Gamma) \subseteq f(\varphi)$, for every f.
(f) Is not the case that: $\Gamma \mid\models \varphi$ if and only if $\Gamma \models_V \varphi$.

Proof

(a) Let $\varphi \mid\models \psi$, and $\psi \mid\models \gamma$. By Definition 4.4, $I(\varphi) \geq I(\psi) \geq I(\gamma)$, for every f. Then, by TPP, $I(\varphi) \geq I(\gamma)$, for every f. So, by Definition 4.4, $\varphi \mid\models \gamma$.
(b) By TPP, if $\varphi \equiv \psi$, then $\varphi \leftrightarrow \psi$ is $\top_{\mathbb{P}}$. And, by TPP, $\varphi \leftrightarrow \psi$ is $\top_{\mathbb{P}}$, then $\models_{\mathbb{P}} \varphi \leftrightarrow \psi$. So, by Theorem 4.6d, $\mid\models \varphi \leftrightarrow \psi$.
(c) ($\not\Rightarrow$): Let $\Gamma = \emptyset$. $\bot_{\mathbb{P}} \psi$, but $\mid\models \bot_{\mathbb{P}} \rightarrow \psi$.
($\not\Leftarrow$): $\not\mid\models \psi \rightarrow \bot_{\mathbb{P}}$, but $\psi \mid\models \bot_{\mathbb{P}}$.
(d) ($\not\Rightarrow$): $\mid\models \bot_{\mathbb{P}}$ e $\mid\models \top_{\mathbb{P}}$, but $\bot_{\mathbb{P}} \not\equiv \top_{\mathbb{P}}$.
($\not\Leftarrow$): $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi$, but $\models_{\mathbb{P}} \varphi \rightarrow \psi$ and $\models_{\mathbb{P}} \neg\varphi \vee \psi$ do not occur.
(e) ($\not\Rightarrow$): $\top_{\mathbb{P}} \mid\models \bot_{\mathbb{P}}$, but $f(\top_{\mathbb{P}}) = U \not\subseteq f(\bot_{\mathbb{P}}) = \emptyset$.
($\not\Leftarrow$): Let $f(\varphi) \neq \emptyset$. Then $f(\bot_{\mathbb{P}}) \subseteq f(\varphi)$, but $\bot_{\mathbb{P}} \varphi$.
(f) ($\not\Rightarrow$): $\varphi \vee \neg\varphi \mid\models \varphi \wedge \neg\varphi$, but $\varphi \vee \neg\varphi_{\mathrm{V}}\varphi \wedge \neg\varphi$.
($\not\Leftarrow$): $\psi \models_{\mathrm{V}} \varphi \vee \psi$, but $\psi \not\mid\models \varphi \vee \psi$. □

The above theorem expresses some of the characteristic properties of informational logical consequence, when compared to the probabilistic and veritative-functional versions. In contrast to the latter two versions, the reciprocal of Theorem 4.9b, especially, cannot be shown for the informational version; in other words, is not the case that if $\mid\models \varphi \leftrightarrow \psi$, then $\varphi \equiv \psi$, given that, for example, $\mid\models \bot_{\mathbb{P}} \leftrightarrow \top_{\mathbb{P}}$, but $\bot_{\mathbb{P}} \not\equiv \top_{\mathbb{P}}$.

Theorem 4.9c shows the invalidity of the corresponding semantics of the deduction theorem. The fourth and fifth items illustrate some of the innate characteristics of informational logical consequence, as discussed in the final considerations (below). Theorem 4.9f expresses a distinction between the relations of informational and veritative-functional logical consequence.

We show below that informational logical consequence is not a Tarskian logical consequence, according to the definition set out in the introduction of this chapter.

Theorem 4.10

(a) Is not the case that: if $\varphi \in \Gamma$, then $\Gamma \mid\models \varphi$.
(b) Is not the case that: if $\Gamma \subseteq \Delta$ and $\Gamma \mid\models \varphi$, then $\Delta \mid\models \varphi$.
(c) If $\Delta \mid\models \psi$, for each $\psi \in \Gamma$, and $\Gamma \mid\models \varphi$, then $\Delta \mid\models \varphi$.

Proof Let f be any situation.

(a) Let $P(\Gamma) = 0$ and $P(\varphi) = 1/2$. Then, by TPP, $I(\Gamma) < I(\varphi)$.
(b) Let $\Gamma = \{\varphi\}$, $\Delta = \{\varphi, \neg\varphi\}$ and $I(\varphi) > 0$. Then, by TPP, $\Gamma \mid\models \varphi$ and $\Delta \not\mid\models \varphi$.
(c) Case 1: Let $I(\Delta) = 0$. Then, by Hypothesis, $I(\psi) = 0$, for each $\psi \in \Gamma$. By TPP, $I(\Gamma) = 0$. Thus, by Hypothesis, $I(\varphi) = 0$. So, $I(\Delta) = I(\varphi)$.
Case 2: Let $I(\Delta) \neq 0$. By TPP, and by Hypothesis, $P(\Delta) \leq P(\psi)$, for each $\psi \in \Gamma$. In this case, it can be shown that $f'(\Delta) \subseteq f'(\Gamma) \subseteq f'(\varphi)$, for every f'. Then, by TPP and Definition 3.8, $P(\Delta) \leq P(\Gamma) \leq P(\varphi)$. So, by TPP, $I(\Delta) \geq I(\varphi)$. □

Thus, informational logical consequence is neither reflexive nor monotonic, although it may satisfy the condition of transitivity.

According to our objectives in this paper, we finish our presentation of informational logical consequence. Others results and properties can be found in Alves (2012b).

5 Final Considerations

The shift of perspective in the analysis of the notion of logical consequence, from the true value to the quantity of information of formulae, enables characteristics to emerge that are uniquely attributable to this approach. Alves [2] presents some of the main similarities and differences between this concept and the usual perspective whereby logical consequence is defined in terms of maintenance of the truth of the premises for the conclusion of an argument. In what follows, we highlight the main results obtained from the elements presented in this chapter.

FC1 Informationally empty formulae are informational logical consequence of the empty set (Theorem 4.6a). This means that probabilistically valid and contradictory formulae are self-sustained, which illustrates a first difference between, on one hand, the veritative-functional and probabilistic versions and, on the other hand, the informational version of logical consequence. In these two versions, a contradiction is generally not a logical consequence of a given set of premises.

FC2 The formulae that are informational logical consequence of a given set of formulae whose quantity of information is always null are informational logical consequence of the empty set (Theorem 4.6c). This result also illustrates an inherent characteristic of the informational perspective of logical consequence. From this, it follows that if a formula is logical consequence of a contradictory set of formulae, it is informational logical consequence of the empty set. This does not generally hold in the case of the veritative-functional and probabilistic perspectives of logical consequence.

FC3 Some of the rules of inference of classical formal logical systems, and some of the arguments traditionally considered valid, do not possess general validity in the informational perspective of logical consequence (Proposition 4.7). The cases in which the set of premises possesses null information provide examples showing that the conclusion can be more informative than the set of premises in these rules or arguments. Meanwhile, it can be shown that when the quantity of information in the set of premises is greater than zero, the quantity of information in the conclusion is always smaller than the quantity of information in the set of premises.

This result indicates that according to the veritative-functional perspective, the conclusion of a valid argument can possess more information than its set of premises. This seems to support the conception that in a valid argument, the information in the conclusion is already implicitly or explicitly given in the premises. We leave it open for future work to investigate the nature of the information underlying the veritative-functional perspective, as well as the association between the true and probability values and the sentences involved in a logical consequence relation. We believe that this analysis should resolve, at least partially, the strangeness indicated in this paragraph.

In the informational and veritative-functional perspectives, amplifying inductive arguments, where the conclusion is more informative than the set of premises, are invalid. Meanwhile, such arguments may be considered interesting in some areas of knowledge, given that they amplify information.

The theorem in question also shows that the informational perspective of logical consequence is not equivalent to the veritative-functional and probabilistic perspectives. There are formulae that are the veritative-functional and probabilistic logical consequence of a given set of formulae, but are not informational logical consequence of it. On the other hand, as already shown from Theorem 4.6d or 4.8d, some formulae can be informational logical consequence of a given set of formulae, but cannot be logical consequence from the probabilistic or veritative-functional points of view.

FC4 A formula is only informational logical consequence of a given set of formulae if that set of premises is probabilistically equivalent to the conclusion, or if the conclusion is informationally null (Theorem 4.8). This result shows that a probabilistically contradictory formula can be informational consequence of a given set of informative formulae. As an example: $\varphi \mid\models \neg\varphi \wedge \varphi$.

If, on one hand, informational logical consequence is distinct from traditional classical deductive logical consequence, on the other hand, it should not be considered to be an inductive inference. This is because a characteristic of induction is that it permits the conclusion to possess more information than its set of premises.

FC5 The logic underlying informational logical consequence is, at the least, paraconsistent *sensu lato* (Theorem 4.8b). One of the motives for arriving at this conclusion is that in paraconsistent logics, the principle of explosion is not valid: from a contradiction does not follow any formula. A proposal for future work is to analyze the elementary characteristics of a paraconsistent system, and show that they are satisfied by informational logical consequence.

The logic underlying informational logical consequence is not classical logic, since in classical formal logical systems a contradiction generates any formula. Consequently, the complementary logics are also unable to provide a basis for informational logical consequence, given that they satisfy the principles of classical logic. Heterodox logics, such as the intuitionistic and temporal systems, are also unable to fulfill this role, as shown in the commentaries on Theorems 4.8b and 4.5d.

From Theorem 4.8b, we find that, in contrast to the veritative-functional and probabilistic perspectives, in formal logical systems that adopt informational logical consequence, the inconsistency of a given theory does not imply its triviality, as shown by Alves [2]. From the point of view of formal logical systems such as the classical ones, this signifies that any formula would be considered a theorem of the given theory. Consequently, from the theorem of completeness, for these systems we have $\models_{\mathrm{V}} \psi$ and $\models_{\mathbb{P}} \psi$, for any ψ. Meanwhile, since in informational logical consequence, in general, $\perp_{\mathbb{P}} \mid\not\models \psi$, it is not possible to conclude that the inconsistency of a theory implies its triviality.

FC6 The corresponding semantics of the theorem of deduction is not valid in the informational perspective of logical consequence (Theorem 4.9c). The last three items of

Theorem 4.9 illustrate other significant results specific to the informational perspective. In the veritative-functional and probabilistic perspectives, if two formulae are logical consequence of an empty set of inferences, then they are equivalent, as they are both valid. In the informational version, as expressed in Theorem 4.9d, this does not possess general validity. A probabilistically valid formula and another that is probabilistically inconsistent provide an example to demonstrate the invalidity of this result.

Theorem 4.9d illustrates another specificity of informational logical consequence. This result can be easily shown for the probabilistic version, as reported by Alves [2]. In the case of the veritative-functional version, it is necessary to adapt the right hand side of the result: the true value of the set of premises is smaller than the true value of the conclusion, in all evaluations. This signifies that it is never possible that all the premises can be true while the conclusion is false in the same single evaluation. Finally, the last item of the theorem states that informational and veritative-functional logical consequence cannot be satisfied by the same sets of sentences.

FC7 Informational logical consequence is not a Tarskian logical consequence (Theorem 4.10). A large part of the characteristics specific to the informational perspective of logical consequence, such as that which refers to nonmonotonicity, derives from cases in which the quantity of information is null. The specific difference of this perspective is found in cases that concern extreme probability values, whether in individual situations or in all situations.

While this may be the basic difference, it should not be considered a small difference. It produces results that might be considered discrepant, when compared to the traditional perspectives of logical consequence. Among these, we recall that some rules of inference do not constitute valid arguments, and that not every formula is informational logical consequence of a contradictory set of formulae. Furthermore, it shows that logical consequence, when analyzed from the informational viewpoint in question, ceases to be Tarskian, and that the logic underlying this perspective is not classical, but at the very least paraconsistent.

FC8 The set of premises of an informational inference is always finite, and the sample space is constituted of a finite set of events. These two characteristics represent serious restrictions in our proposal. The first restriction indicates the impossibility of dealing with arguments using a potentially infinite set of premises, such as the ω-arguments, described by Tarski [20]. This author claims to have constructed a theory in which sentences of the type "n possesses the property P," for natural n, are theorems of the theory, and the sentence "every natural number possesses the property P" cannot be proved in the theory. Hence, this sentence is not their logical consequence, which in this case seems absurd.

The second restriction limits the possible models for a language of a formal system. A proposal for future work would be to consider informational logical consequence based on a definition that involves an infinite probability space. It would then become viable to analyze informational logical consequence for languages of first-order theories, which has not been addressed in the present work.

References

1. Alves, M.A.: Informação e conteúdo informacional: notas para um estudo da ação. In: Gonzalez, M.E.Q., BROENS, M.C. (Orgs.) Informação, Conhecimento e Ação Ética, pp. 98–112. Cultura Acadêmica, São Paulo (2012)
2. Alves, M.A.: Lógica e Informação: Uma Análise da Consequência Lógica a Partir de uma Perspectiva Quantitativa da Informação. Doctorate thesis, Universidade Estadual de Campinas, Campinas/SP (2012)
3. Bresciani Filho, E., D'Ottaviano, I.M.L.: Conceitos básicos de sistêmica. In: D'Ottaviano, I.M.L, Gonzales, M.E.Q. (Orgs.) Auto-organização: Estudos Interdisciplinares, pp. 283–306. Coleção CLE, v. 30. Universidade Estadual de Campinas/CLE, Campinas (1990)
4. Da Costa, N.C.A.: Lógica Indutiva e Probabilidade, 2nd edn. Ed. HUCITEC, São Paulo (1993)
5. Da Costa, N.C.A.: Logiques Classiques et Non Classiques. Essai sur les Fundements de la Logique. Masson, Paris (1997)
6. D'Ottaviano, I.M.L.: A lógica clássica e o surgimento das lógicas não-clássicas. In: Évora, F.R.R. (Org.) Século XIX: O Nascimento da Ciência Contemporânea, pp. 65–93. Coleção CLE, v. 11. Universidade Estadual de Campinas/CLE, Campinas (1992)
7. Dretske, F.: Knowledge and the Flow of Information. Basil Blackwell, Oxford (1981)
8. Enderton, H.B.: Elements of Set Theory. Academic Press, San Diego (1977)
9. Gonzalez, M.E.Q.: Informação e cognição: uma proposta de (dis)solução do problema mente-corpo. In: Encontro Brasileiro/Internacional de Ciências Cognitivas, 2, 1996, Campos dos Goytacazes. Anais... pp. 53–60. Universidade Estadual do Norte Fluminense, Campos de Goytacazes (1996)
10. Feitosa, H.A, D'Ottaviano, I.M.L.: Um olhar algébrico sobre as traduções intuicionistas. In: Sautter, F.T., Feitosa, H.A. (Orgs.) Lógica: Teoria, Aplicações e Reflexões, pp. 59–90. Coleção CLE, v. 39. Universidade Estadual de Campinas/CLE, Campinas (2004)
11. Haack, S.: Philosophy of Logics. Cambridge University Press, Cambridge (1978)
12. Hartley, R.: Transmission of information. Bell Syst. Tech. J. **7**, 535–563 (1927)
13. Hershberger, W.: Principles of Communication Systems. Prentice-Hall, New York (1955)
14. Mates, B.: Elementary Logic. Oxford University Press, New York (1972)
15. Mendelson, E.: Introduction to Mathematical Logic. D. Van Nostrand, Princeton, NJ (1964)
16. Shannon, C.: A mathematical theory of information. Bell Syst. Tech. J. **27**, 379–423 (1948)
17. Shannon, C., Weaver, W.: The Mathematical Theory of Information. University of Illinois Press, Urbana (1949)
18. Shoenfield, J.: Mathematical Logic. Addison Wesley, Reading, MA (1967)
19. Tarski, A.: On some fundamental concepts of metamathematics. In: Tarski, A. (ed.) Logic, Semantics, Metamathematics: papers from 1923 to 1938, Translated by J.H. Woodger, pp. 30–37. Clarendon Press, Oxford (1956)
20. Tarski, A.: Fundamental concepts of methodology of deductive sciences. In: Tarski, A. (ed.) Logic, Semantics, Metamathematics: papers from 1923 to 1938, Translated by J.H. Woodger, pp. 60–109. Clarendon Press, Oxford (1956)
21. Tarski, A.: Concept of truth in formalized languages. In: Tarski, A. (ed.) Logic, Semantics, Metamathematics: papers from 1923 to 1938, Translated by J.H. Woodger, pp. 152–278. Clarendon Press, Oxford (1956)
22. Tarski, A.: On the concept of logical consequence. In: Tarski, A. (ed.) Logic, Semantics, Metamathematics: papers from 1923 to 1938, Translated by J.H. Woodger, pp. 409–420. Clarendon Press, Oxford (1956)

M. A. Alves
UNESP, Philosophy Department, State University of São Paulo, Marília/SP, Brazil
e-mail: marcosalves@cle.unicamp.br

I. M. Loffredo D'Ottaviano (✉)
Unicamp, Philosophy Department, Centre for Logic, Epistemology and the History of Science – CLE, University of Campinas, Campinas, Brazil
e-mail: itala@cle.unicamp.br

Finite-Variable Logics Do not Have Weak Beth Definability Property

Hajnal Andréka and István Németi

Abstract We prove that n-variable logics do not have the weak Beth definability property, for all $n \geq 3$. This was known for $n = 3$ (Ildikó Sain and András Simon), and for $n \geq 5$ (Ian Hodkinson). Neither of the previous proofs works for $n = 4$. In this paper, we settle the case of $n = 4$, and we give a uniform, simpler proof for all $n \geq 3$. The case for $n = 2$ is left open.

Keywords Definability theory · Weak Beth definability property · Finite-variable fragment of first-order logic · Mathematical logic

Mathematics Subject Classification (2010) Primary 03B10, 03C40 · Secondary 03B20, 03C07, 03C40

1 Introduction

Definability theory is one of the most exciting and important parts of logic. It concerns concept formation and structuring our knowledge by investigating the category of theories. Implicit definitions are important in understanding concept formation and explicit definitions are vital ingredients of interpretations between theories. This has applications in the methodology of sciences [4, 6, 15].

The Beth definability theorem for first-order logic (FOL) states that each implicit definition is equivalent to an explicit one, these are modulo theories. Investigating whether this theorem holds for fragments of first-order logic gives information about the complexity of the explicit definition equivalent to the implicit one. The Beth definability property is equivalent to surjectivity of epimorphisms in the associated class of algebra (this is a theorem of Németi [17]; see also [7, 13, 18]).

The failure of the Beth definability property for finite variable fragments was first proved in 1983 [3] (for all $n \geq 2$) by showing that epimorphisms are not surjective in finite-dimensional cylindric algebras, see [2]. That proof, translated to logic, relies inherently on the fact that the implicit definition it uses is not satisfiable in each model of the theory. The question came up whether the so-called weak Beth definability property holds for finite-variable fragments. The weak Beth definability property differs from the original Beth definability property in that we require not only the uniqueness, but also

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_4

the existence of the implicitly defined relation. In some sense, the weak Beth definability property is more intuitive and is considered to be more important than the (strong) Beth definability property, see e.g., [5].

In this paper, we prove that n-variable logics do not have the weak Beth definability property either, for all $n \geq 3$. This means that there are a first-order logic theory, and an implicit definition that has exactly one solution in each model of the theory, such that both the theory and the implicit definition are written up using n variables only, yet any explicit definition equivalent to this implicit one has to use more than n variables. For more on finite-variable logics and Beth definability properties see [12] and the remarks at the end of this chapter.

2 The Main Theorem

The n-variable fragment $\mathcal{L}_n$ of an FOL language $\mathcal{L}$, where n is any finite number, is the set of all formulas in $\mathcal{L}$ which use n variables only (free or bound). To make this more concrete, we may assume that $\mathcal{L}$ uses the variables $v_0, v_1, \ldots$, while $\mathcal{L}_n$ uses only the variables $v_0, v_1, \ldots, v_{n-1}$. In finite variable fragments we do not allow function or constant symbols, but we allow equality. Here is a definition of the formulas of $\mathcal{L}_n$:

- $R(v_{i_1}, \ldots, v_{i_k})$ is a formula of $\mathcal{L}_n$ if R is a k-place relation symbol and $i_1, \ldots, i_k < n$.
- $v_i = v_j$ is a formula of $\mathcal{L}_n$ if $i, j < n$.
- $\neg\varphi,\ \varphi \wedge \psi,\ \exists v_i\, \varphi$ are formulas of $\mathcal{L}_n$ whenever φ, ψ are formulas of $\mathcal{L}_n$ and $i < n$.

The above are all the formulas of $\mathcal{L}_n$. We use other logical connectives, e.g., $\forall v_i, \vee, \rightarrow$ as derived ones. Models, satisfiability of formulas under evaluations of the variables, and validity in $\mathcal{L}_n$ are the same as in FOL. The following theorem says that $\mathcal{L}_n$ does not have even the weak Beth definability property whenever $n \geq 3$.

Theorem 2.1 (No Weak Beth Property for $\mathcal{L}_n$.) Let $n \geq 3$. There are a theory Th in the language of an n-place relation symbol R and a binary relation symbol S, and a theory $\Sigma(D)$ in the language of Th enriched with a unary relation symbol D such that

- in each model of Th there is a unique relation D for which $\Sigma(D)$ holds (we call such $\Sigma(D)$ a strong implicit definition of D in Th);
- there is no explicit definition for D in Th, i.e., for each n-variable formula φ in the language of Th we have

$$\mathsf{Th} \cup \Sigma(D) \not\models \forall v_0[D(v_0) \leftrightarrow \varphi].$$

Proof We write out the proof in detail for $n = 3$. Generalizing this proof to all $n \geq 3$ will be easy. We will often write x, y, z for v_0, v_1, v_2 and we will write simply R for $R(x, y, z)$. We will use $U_0(x), U_1(y), U_2(z)$ as abbreviations of the formulas on the right-hand sides of the respective $\leftrightarrow$ below:

$$U_0(x) :\leftrightarrow \exists yzR\,, \qquad U_1(y) :\leftrightarrow \exists xzR\,, \qquad U_2(z) :\leftrightarrow \exists xyR\,.$$

These formulas express the domain of R, i.e., the first projection of R, and the second and third projections of R. We will include formulas in Th that express that U_0, U_1, U_2 are sets of cardinalities $3, 2, 2$ respectively, and they form a partition of the universe. We will formulate these properties with three variables after describing the main part of the construction. Let us introduce the abbreviations T and $\mathsf{big}(R)$ as

$$T :\leftrightarrow U_0(x) \wedge U_1(y) \wedge U_2(z), \quad \text{and}$$
$$\mathsf{big}(R) :\leftrightarrow \bigwedge\{\exists v_i\, R \leftrightarrow \exists v_i (T \wedge \neg R) : i = 0, 1, 2\}.$$

In the above, T is the "rectangular hull" of R, and $\mathsf{big}(R)$ expresses that R cuts this hull into two parts, each of which is sensitive in the sense that as soon as we quantify over them, the information on how R cuts T into two parts disappears. (Note that $\mathsf{big}(R)$ implies that $\exists v_i\, R \leftrightarrow \exists v_i T \leftrightarrow \exists v_i (T \wedge \neg R)$.) Assume that $|U_0| = 3, |U_1| = 2, |U_2| = 2$ and $\mathsf{partition}(U_0, U_1, U_2)$ are formulas in $\mathcal{L}_3$ that express the associated meanings. Then we define

$$\mathsf{Th} := \{|U_0| = 3, |U_1| = 2, |U_2| = 2, \quad \mathsf{partition}(U_0, U_1, U_2), \quad \mathsf{big}(R)\}.$$

We will show that Th has exactly one model, up to isomorphism. However, before doing that, let us turn to expressing the promised properties about the U_i's using three variables.

We will use *Tarski's way of substituting* one variable for the other. That is, we introduce the abbreviations

$$U_1(x) :\leftrightarrow \exists y(x = y \wedge U_1(y)), \qquad U_2(x) :\leftrightarrow \exists z(x = z \wedge U_2(z)).$$

We now can express that U_0, U_1, U_2 form a partition of the universe:

$$\forall x(U_0(x) \vee U_1(x) \vee U_2(x)), \qquad \forall x(U_i(x) \to \neg U_j(x)) \quad \text{for} \quad i \neq j, \ i, j < 3.$$

To express the sizes of the sets U_i we will use the abbreviations

$$U_1(z) :\leftrightarrow \exists y(z = y \wedge U_1(y)), \qquad U_2(y) :\leftrightarrow \exists z(y = z \wedge U_2(z)).$$

Now, for $i = 1, 2$ we define the formulas

$$|U_i| \leq 2 :\leftrightarrow \neg\exists xyz(x \neq y \wedge x \neq z \wedge y \neq z \wedge U_i(x) \wedge U_i(y) \wedge U_i(z)),$$
$$|U_i| \geq 2 :\leftrightarrow \exists xy(x \neq y \wedge U_i(x) \wedge U_i(y)),$$
$$|U_i| = 2 :\leftrightarrow |U_i| \geq 2 \wedge |U_i| \leq 2.$$

It remains to express that U_0 has exactly three elements. In $\mathcal{L}_n$ with $n \geq 4$ we can express $|U_0| = 3$ similarly to the above, but in $\mathcal{L}_3$ we have to use another tool. To express in $\mathcal{L}_3$ that U_0 has exactly three elements, we will use the binary relation S. (This is the sole use of S in Th, for $n \geq 4$ we can omit S from the language.) We shall express that S is a cycle of order 3 on U_0. The following formulas express that S is a function on U_0 without a fixed point:

$$\forall x \exists y\, S(x,y), \quad S(x,y) \wedge S(x,z) \to y = z, \quad S(x,y) \to (U_0(x) \wedge U_0(y) \wedge x \neq y).$$

The following formula expresses that U_0 consists of exactly one 3-cycle of S:

$$S(x,y) \leftrightarrow \exists z(S(y,z) \wedge S(z,x)), \qquad S(x,y) \vee S(y,x) \vee x = y.$$

In the above, we used Tarski-style substitution of variables without mentioning (e.g., $U_0(y)$) and we omitted universal quantifiers in front of formulas (e.g., we wrote $S(x,y) \wedge S(x,z) \to y = z$ in place of $\forall xy(S(x,y) \wedge S(x,z) \to y = z)$). This expresses that U_0 has exactly three elements.

We turn to showing that Th has exactly one model up to isomorphism. Let $\mathfrak{M} = \langle M, R, S\rangle \models \mathsf{Th}$. Let U_i, T be defined as above. Then M is the disjoint union of the U_i's, and the sizes of the U_i's for $i = 0,1,2$ are $3,2,2$, respectively. (So M has seven elements.) Let $U_1 = \{b_0, b_1\}$, let c, d be the two elements of U_2, and let

$$X := \{u \in U_0 : \langle u, b_0, c\rangle \in R\}.$$

By $\mathfrak{M} \models \mathsf{big}(R)$ and $|U_2| = 2$ we have that $\langle u, b_0, d\rangle \notin R$ if $u \in X$ and $\langle u, b_0, d\rangle \in R$ if $u \in U_0 - X$. Hence

$$U_0 - X = \{u \in U_0 : \langle u, b_0, d\rangle \in R\}.$$

Also, by $\mathfrak{M} \models \mathsf{big}(R)$, X has one or two elements (it cannot be that X has zero or three elements). If $|X| = 1$ then let us use the notation $c_0 = c, c_1 = d$, and if $|X| = 2$ then let $c_0 = d, c_1 = c$. Let us name the elements of U_0 as a_0, a_1, a_2 such that $X = \{a_0\}$ if $|X| = 1$, $X = \{a_1, a_2\}$ if $|X| = 2$ and $S = \{\langle a_i, a_j\rangle : j = i + 1 \pmod 3 \text{ and } i, j \leq 3\}$. This can be done by $\mathfrak{M} \models \mathsf{Th}$. The setting so far determines R by $\mathfrak{M} \models \mathsf{big}(R)$, as follows. For all $i \leq 2,\ j,k \leq 1$ we have $\langle a_i, b_j, c_k\rangle \in R$ if and only if $\langle a_i, b_{j+1 \pmod 2}, c_k\rangle \in T - R$ if and only if $\langle a_i, b_j, c_{k+1 \pmod 2}\rangle \in T - R$. This is so by $\mathfrak{M} \models \mathsf{big}(R)$ and by $|U_i| = 2$ for $i = 1,2$. From this we have that

$$\begin{aligned} R = \{\langle u, b_i, c_j\rangle : u = a_0 \quad & \text{and} \quad i + j = 0 \pmod 2\} \cup \\ \{\langle u, b_i, c_j\rangle : u = a_1 \vee u = a_2 \quad & \text{and} \quad i + j = 1 \pmod 2\}. \end{aligned}$$

We have seen that all models of Th are isomorphic to each other. The above also show that there is no automorphism of $\mathfrak{M}$ that would move $\{a_0\}$.

We are ready to formulate our implicit definition $\Sigma(D)$. We design $\Sigma(D)$ so that, by using the above notation, it specifies $\{a_0\}$. We will write D in place of $D(x)$.

$$\begin{aligned} \Sigma(D) := \{ \quad & T \wedge \neg D \wedge R && \to \forall x(T \wedge \neg D \to R)), \\ & T \wedge \neg D \wedge \neg R && \to \forall x(T \wedge \neg D \to \neg R)), \\ & D \to U_0(x), \quad |D| = 1 && \}. \end{aligned}$$

Then in each model of Th there is exactly one unary relation D for which $\Sigma(D)$ holds, namely D has to be the unary relation $\{a_0\} \subseteq U_0$. Thus $\Sigma(D)$ is a strong implicit definition of D in Th.

It remains to show that Σ cannot be made explicit in $\mathcal{L}_3$, i.e., there is no three-variable formula φ in the language of Th for which $\mathsf{Th} \cup \Sigma(D) \models D \leftrightarrow \varphi$. Our plan is to list all the $\mathcal{L}_3$-definable relations in the above model and observe that $\{a_0\}$, the relation that Σ defines, is not among them. For any $\varphi \in \mathcal{L}_3$ define

$$\mathsf{mn}(\varphi) := \{\langle a, b, c\rangle : \mathfrak{M} \models \varphi[a, b, c]\}.$$

In the above, $\mathfrak{M} \models \varphi[a, b, c]$ denotes that the formula φ is true in $\mathfrak{M}$ when the variables v_0, v_1, v_2 are evaluated to a, b, c, respectively, and mn abbreviates "*meaning*". Let

$$A := \{\mathsf{mn}(\varphi) : \varphi \in \mathcal{L}_3\}.$$

Clearly, A is closed under the set Boolean operations because

$$\begin{aligned} \mathsf{mn}(\varphi \wedge \psi) &= \mathsf{mn}(\varphi) \cap \mathsf{mn}(\psi), \\ \mathsf{mn}(\neg\varphi) &= M^3 - \mathsf{mn}(\varphi), \end{aligned}$$

and so A is closed under intersection and complementation with respect to M^3, the set of all M-termed three-sequences. Since M is finite, this implies that A is atomic and the elements of A are exactly the unions of some atoms.

We will list all the atoms of A. It is easy to see that the elements $U_i \times U_j \times U_k$ for $i, j, k \leq 2$ are all in A and they form a partition of M^3. To list the atoms of A, we will list the atoms below each $U_i \times U_j \times U_k$ by specifying a partition of each. For $i, j, k \leq 2$ let us abbreviate the sequence $\langle i, j, k\rangle$ by ijk.

$U_0 \times U_1 \times U_2$ is T, and the partition of T will be $\{R, T - R\}$. For ijk a permutation of 012, the partition of $U_i \times U_j \times U_k$, the permuted version of T, will be the correspondingly permuted versions of R and $T - R$. Assume i, j, k are all distinct, i.e., they form a permutation of $0, 1, 2$. We define

$$\begin{aligned} X(ijk, r) &:= \{\langle u_i, u_j, u_k\rangle : \langle u_0, u_1, u_2\rangle \in R\}, \\ X(ijk, -r) &:= \{\langle u_i, u_j, u_k\rangle \in U_i \times U_j \times U_k : \langle u_0, u_1, u_2\rangle \notin R\}. \end{aligned}$$

We note that

$$X(012, r) = R, \quad \text{and} \quad X(012, -r) = T - R.$$

Note that

$$\mathsf{mn}(R(v_i, v_j, v_k)) = X(ijk, r),$$

and the same for $-r$ in place of r, so $X(ijk, r)$, $X(ijk, -r)$ are elements of A.

Assume now that ijk is not repetition free, i.e., $|\{i, j, k\}| < 3$. In these cases, the blocks of the partition of $U_i \times U_j \times U_k$ will be put together from partitions of $U_m \times U_n$ $(m, n < 3)$. Recall that $S = \{\langle a_0, a_1\rangle, \langle a_1, a_2\rangle, \langle a_2, a_0\rangle\}$. We define

$$\begin{aligned} \overline{S} &:= \{\langle a, b\rangle : \langle b, a\rangle \in S\}, \\ \mathsf{id}_i &:= \{\langle a, a\rangle : a \in U_i\}, \\ \mathsf{di}_i &:= \{\langle a, b\rangle : a \neq b, \ \ a, b \in U_i\}. \end{aligned}$$

Above, id_i, di_i abbreviate "*identity* on U_i", and "*diversity* on U_i", respectively, and $\overline{S}$ is the inverse of S. Since S is a cycle on the three-element set U_0, its inverse $\overline{S}$ is its complement in the diversity element of U_0, so $\{S, \overline{S}, \mathsf{id}_0\}$ is a partition of $U_0 \times U_0$. Also, $\{\mathsf{di}_i, \mathsf{id}_i\}$ is a partition of $U_i \times U_i$ for $i = 1, 2$. We are ready to define the "binary partitions" as follows

$$\mathsf{Rel}_{00} := \{S, \overline{S}, \mathsf{id}_0\}, \quad \mathsf{Rel}_{11} := \{\mathsf{di}_1, \mathsf{id}_1\}, \quad \mathsf{Rel}_{22} := \{\mathsf{di}_2, \mathsf{id}_2\},$$
$$\mathsf{Rel}_{ij} := \{U_i \times U_j\} \quad \text{for} \quad i \neq j \ .$$

Note that for all $e \in \mathsf{Rel}_{ij}$, $e' \in \mathsf{Rel}_{jk}$ we have $e \circ e' \in \mathsf{Rel}_{ik}$, where $\circ$ denotes the operation of composing binary relations. In general, when $|\{i, j, k\}| < 3$ and $e = \langle e_0, e_1\rangle \in \mathsf{Rel}_{ij} \times \mathsf{Rel}_{jk}$ we define

$$X(ijk, e) := \{\langle a, b, c\rangle \in U_i \times U_j \times U_k : \langle a, b\rangle \in e_0,\ \langle b, c\rangle \in e_1\}.$$

Notice that we already defined $X(ijk, e)$ for the case when i, j, k are distinct and $e \in \{r, -r\}$. Let $\mathsf{choice}(e, ijk)$ denote $e \in \{r, -r\}$ when ijk is repetition free, and $e = \langle e_0, e_1\rangle$, $e_0 \in \mathsf{Rel}_{ij}$, $e_1 \in \mathsf{Rel}_{jk}$ otherwise. Define

$$B := \{X(ijk, e) : i, j, k \leq 2, \mathsf{choice}(e, ijk)\},$$
$$C := \{\bigcup Y : Y \subseteq B\}.$$

The following notation will be convenient when $\mathsf{choice}(e, ijk)$ and ijk are not repetition free.

$$e_{01} := e_0, \quad e_{12} := e_1, \quad e_{02} := e_0 \circ e_1,$$
$$e_{ij} := \overline{S} \quad \text{when} \quad i > j \ \text{and} \ e_{ji} = S\ ,$$
$$e_{ij} := e_{ji} \quad \text{when} \quad i > j \ \text{and} \ e_{ji} \neq S\ .$$

The intuitive meaning of e_{ij} is that $\langle a_i, a_j\rangle \in e_{ij}$ whenever $\langle a_0, a_1\rangle \in e_0$ and $\langle a_1, a_2\rangle \in e_1$.

We want to prove that $A = C$. We show $A \subseteq C$ by showing $\mathsf{mn}(\varphi) \in C$ for all $\varphi \in \mathcal{L}_3$, by induction on φ. The atomic formulas are:

$$\mathsf{mn}(R(v_i, v_j, v_k)) = X(ijk, r) \quad \text{when} \quad |\{i, j, k\}| = 3\ ,$$
$$\mathsf{mn}(R(v_i, v_j, v_k)) = \emptyset \quad \text{otherwise},$$
$$\mathsf{mn}(S(v_i, v_j)) = \bigcup\{X(n_1n_2n_3, e) : n_i = n_j = 0, e_{n_in_j} = S\},$$
$$\mathsf{mn}(v_i = v_j) = \bigcup\{X(n_1n_2n_3, e) : n_i = n_j, e_{n_in_j} \in \{\mathsf{id}_0, \mathsf{id}_1, \mathsf{id}_2\}\}.$$

Clearly, $M^3 \in C$, and C is closed under complementation with respect to M^3 and intersection, because B is finite and its elements form a partition of M^3. Thus,

$$\mathsf{mn}(\neg\varphi) \in C, \quad \mathsf{mn}(\varphi \wedge \psi) \in C \quad \text{whenever} \quad \mathsf{mn}(\varphi), \mathsf{mn}(\psi) \in C.$$

To deal with the existential quantifiers, let us define for arbitrary $H \subseteq M^3$

$$\mathsf{C}_0 H := \{\langle a, b, c\rangle \in M^3 : \langle a', b, c\rangle \in H \quad \text{for some} \quad a'\},$$
$$\mathsf{C}_1 H := \{\langle a, b, c\rangle \in M^3 : \langle a, b', c\rangle \in H \quad \text{for some} \quad b'\},$$
$$\mathsf{C}_2 H := \{\langle a, b, c\rangle \in M^3 : \langle a, b, c'\rangle \in H \quad \text{for some} \quad c'\}.$$

Then we have, by the definition of the meaning of the existential quantifiers, that for all $i \leq 2$

$$\mathsf{mn}(\exists v_i \varphi) = \mathsf{C}_i \mathsf{mn}(\varphi).$$

Thus, to show that

$$\mathsf{mn}(\exists v_i \varphi) \in C \quad \text{whenever} \quad \mathsf{mn}(\varphi) \in C$$

it is enough to show that C is closed under C_i, i.e., $\mathsf{C}_i X \in C$ whenever $X \in C$ (and $i \leq 2$). Since C_i is additive, i.e., $\mathsf{C}_i(X \cup Y) = \mathsf{C}_i(X) \cup \mathsf{C}_i(Y)$, it is enough to show that

$$\mathsf{C}_m X(ijk, e) \in C \quad \text{for all} \quad i, j, k, m \leq 2, \quad \text{and a good choice } e \text{ for } ijk\,.$$

Assume that i, j, k are distinct and $e \in \{r, -r\}$. Then by $\mathfrak{M} \models \mathsf{big}(R)$

$$\begin{aligned}\mathsf{C}_0 X(ijk, e) &= M \times U_j \times U_k\,,\\ \mathsf{C}_1 X(ijk, e) &= U_i \times M \times U_k\,,\\ \mathsf{C}_2 X(ijk, e) &= U_i \times U_j \times M\,.\end{aligned}$$

It is easy to check that $U_i \times U_j \times U_k \in C$ for all i, j, k, and hence $V_0 \times V_1 \times V_2 \in C$ whenever V_i are unions of U_0, U_1, U_2. When i, j, k are not all distinct

$$\begin{aligned}\mathsf{C}_0 X(ijk, e) = M \times e_{12} = \{\langle a, b, c\rangle : \langle b, c\rangle \in e_{12}\} = \\ \bigcup\{X(mjk, e') : m \leq 2, e'_{12} = e_{12}\},\end{aligned}$$

$$\begin{aligned}\mathsf{C}_1 X(ijk, e) = \{\langle a, b, c\rangle : \langle a, c\rangle \in e_{02}\} = \\ \bigcup\{X(imk, e') : m \leq 2, e'_{02} = e_{02}\},\end{aligned}$$

$$\mathsf{C}_2 X(ijk, e) = \bigcup\{X(ijm, e') : m \leq 2, e'_{01} = e_{01}\}.$$

We have seen that $A \subseteq C$.

To show that $C \subseteq A$ we have to check that each $X(ijk, e)$ is the meaning of a formula $\varphi \in \mathcal{L}_3$ in $\mathfrak{M}$. We already did this for $X(ijk, r)$, i, j, k distinct. For $ijk = 000$ and $e = \langle S, S\rangle$

$$X(000, \langle S, S\rangle) = \mathsf{mn}(U_0(x) \wedge U_0(y) \wedge U_0(z) \wedge S(x, y) \wedge S(y, z)),$$

where $U_0(x) = \exists yzR$, $U_0(y) = \exists x(x = y \wedge U_0(x))$, $U_0(z) = \exists x(x = z \wedge U_0(x))$ are the abbreviations introduced before. The other cases are similar, we leave it to the reader to check them.

Finally, to show that $\mathsf{mn}(D(x)) = \{\langle a_0, b, c\rangle : b, c \in M\} \notin A$, observe that the domain of each element in B either contains U_0 or else is disjoint from it, and therefore the same holds for their unions. Clearly, this is not true for $\mathsf{mn}(D(x))$. This shows that $\mathsf{mn}(D) \notin A$, i.e., D cannot be explicitly defined in $\mathfrak{M}$. Since $\mathfrak{M}$ is a model of Th, this means that $\Sigma(D)$ is not equivalent to any explicit definition that contains only three variables.

To generalize the construction and the proof from $n = 3$ to $n \geq 4$ is straightforward. In the general case M has $2n + 1$ elements, it is the disjoint union of sets $U_0, U_1, \ldots U_{n-1}$ of sizes $3, 2, \ldots, 2$, respectively, and $R = \{s \in U_0 \times \cdots \times U_{n-1} : (s_0 = a_0 \wedge \Sigma\{a_i : 1 \leq i < n\}$ is even$) \vee (s_0 \in \{a_1, a_2\} \wedge \Sigma\{a_i : 1 \leq i < n\}$ is odd$)\}$. □

There is an FOL-formula $\varphi(v_0)$ for Th and $\Sigma(D)$ as in Theorem 2.1, which explicitly defines $D(v_0)$, since the Beth definability theorem holds for FOL. The above theorem then implies that this explicit definition has to use more than n variables. Thus, both the theory and the implicit definition use only n variables, but any equivalent explicit definition has to use more than n variables. In our example, $D(v_0)$ can be defined by using $n + 1$ variables. By using a construction from [10], Ian Hodkinson [12] proved that for any number k there are also a theory and a (weak) implicit definition using only n variables such that any explicit definition to which this implicit definition is equivalent has to use more than $n + k$ variables.

Theorem 2.1 implies (the known fact) that Craig's interpolation theorem does not hold for n-variable logic, either, for $n \geq 3$. This is so because in the standard proof of the Beth's definability theorem in, e.g., [8, Thm. 2.2.22], the explicit definition is constructed from an interpoland. Complexity investigations for Craig's theorem were done earlier, see, e.g., Daniel Mundici [16].

The proof given here proves more than what Theorem 2.1 states. In the proof, Th and $\Sigma(D)$ are written in the so-called *restricted* n-variable logic, and $\Sigma(D)$ is not equivalent to any n-variable formula using even infinitary conjunctions and disjunctions in a finite model of Th. A formula is called restricted if substitution of variables is not allowed in it, i.e., it uses relational atomic formulas of form $R(v_0, \ldots, v_k)$ only (and it does not contain subformulas of form $R(v_{i0}, \ldots, v_{ik})$ where $\langle i0, \ldots, ik\rangle \neq \langle 0, \ldots, k\rangle$), see [11, Part II, Sect. 4.3]. Thus the weak Beth definability property fails for a wide variety of logics, from the restricted n-variable fragment with finite models only, to $L^n_{\infty,\omega}$.

The variant of $\mathcal{L}_n$ in which we allow only models of size $\leq n + 1$ has the strong Beth definability property, for all n, this is proved in [2]. Another variant of $\mathcal{L}_n$ that has the strong Beth definability property is when we allow models of all sizes but in a model truth is defined by using only a set of selected (so-called admissible) evaluations of the variables (a generalized model then is a pair consisting of a model in the usual sense and this set of admissible evaluations). The so-called guarded fragments of n-variable logics also have the strong Beth definability property. For more on this see [1, 9, 14].

We note that $\mathcal{L}_2$ does not have the strong Beth definability property (this is proved in [2]), and we do not know whether it has the weak one. There are indications that it might have it. If so, $\mathcal{L}_2$ would be a natural example of a logic distinguishing the two Beth definability properties. At present, we only have artificial examples for this, see Chapter XVIII by J. Makowsky in [5, p. 689, item 4.2.2(v)].

Acknowledgment This work was completed with the support of Hungarian National Grant for Basic Research No 81188.

References

1. Andréka, H., van Benthem, J.F.A.K., Németi, I.: Modal languages and bounded fragments of predicate logic. J. Philos. Log. **27**, 217–274 (1998)
2. Andréka, H., Comer, S.D., Madarász, J.X., Németi, I., Sayed-Ahmed, T.: Epimorphisms in cylindric algebras and definability in finite variable logic. Algebra Univers. **61**, 3–4, 261–282 (2009)
3. Andréka, H., Comer, S.D., Németi, I.: Epimorphisms in cylindric algebras. Preprint, 1983. Abstracted in Abstracts of AMS, 83T-08-186. April 1983, vol. 4, no. 3, p. 293
4. Andréka, H., Németi, I.: Comparing theories: the dynamics of changing vocabularies. In: Baltag, A., Smets, S. (eds.) Johan F. A. K. van Benthem on Logic and Information Dynamics. Outstanding Contributions to Logic, vol. 5, pp. 143–172. Springer (2014)
5. Barwise, J., Feferman, S.: Model-Theoretic Logics. Springer, New York, Berlin, Heidelberg, Tokyo (1985)
6. van Benthem, J.F.A.K.: The logical study of science. Synthese **51**, 431–472 (1982)
7. Blok, W.J., Hoogland, E.: The Beth property in algebraic logic. Stud. Log. **83**(1–3), 49–90 (2006)
8. Chang, C.C., Keisler, H.J.: Model Theory. North-Holland, Amsterdam and London (1973)
9. Grädel, E.: On the restraining power of guards. J. Symb. Log. **64**(4), 1719–1742 (1999)
10. Gurevich, Y., Shelah, S.: On finite rigid structures. J. Symb. Log. **62**(2), 549–562 (1996)
11. Henkin, L., Monk, J.D., Tarski, A.: Cylindric Algebras. Parts I–II. North-Holland, Amsterdam (1985)
12. Hodkinson, I.: Finite variable logics. Updated version of paper appeared in Bull. Europ. Assoc. Theor. Comp. Sci. **51**, 111–140 (1993), with addendum in vol. 52, pp. 37. http://www.doc.ic.ac.uk/~imh/papers/fvl_revised.pdf
13. Hoogland, E.: Algebraic characterisation of two Beth definability properties. Master Thesis, University of Amsterdam (1996)
14. Hoogland, E., Marx, M., Otto, M.: Beth definability for the guarded fragment. In: Logic for Programming and Automated Reasoning. Lecture Notes in Computer Science, vol. 1705, pp. 273–285. Springer Verlag, Berlin Heidelberg (1999)
15. Madarász, J.X.: Logic and relativity (in the light of definability theory), pp. xviii, 367. PhD Dissertation, ELTE Budapest (2002)
16. Mundici, D.: Complexity of Craig's interpolation. Fundam. Inform. **3**(2), 261–278 (1982)
17. Németi, I.: Surjectivity of Epimorphisms is Equivalent to Beth Definability in Algebraic Logic. Manuscript (1982)
18. Sain, I.: Definability issues in universal logic. In: Andréka, H., Ferenczi, M., Németi, I. (eds.) Cylindric-like Algebras and Algebraic Logic. Bolyai Society Mathematical Studies, vol. 22, pp. 393–419. Springer, Berlin, Heidelberg, New York (2012)
19. Sain, I., Simon, A.: Beth Properties of Finite Variable Fragments of First Order Logic, Preprint. Mathematical Institute of the Hungarian Academy of Sciences, Budapest (1992)

H. Andréka (✉) · I. Németi
MTA Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary
e-mail: andreka.hajnal@renyi.mta.hu

Peirce's Role in the History of Logic: *Lingua Universalis and Calculus Ratiocinator*

Irving H. Anellis

Abstract This chapter discusses central issues of Peirce's conception of logic, comparing his work with the works of other logicians of the time, in particular Boole, Venn, Schröder, and Frege. It presents a detailed analysis of Peirce's approach to notation, including pasigraphy and Peirce's claw.

Keywords Peirce · Boole · Frege · Semiotics · Lingua universalis · Calculus ratiocinator

Mathematics Subject Classification (2010) Primary 01A55 · Secondary 03A05

1 General Account

Recently, Vladimir Ivanovich Shalak [120] undertook to describe two different conceptions of logic: the *traditional*, which derives from Aristotle and whose approach is ontological; and the *semiotic*, which understands logic to be both formal and linguistic, and as such is a normative science. While the concept of logic as semiotic, and as a normative science, is familiar to historians of logic, and in general to Peirce scholars, Shalak, regrettably, does not examine the origin of this distinction; either does he give any indication of the significance of the distinction. Thus, for example, the very well known differences between the traditional and the Boolean squares of opposition, in which the validity of inferences depends upon the question of the existential import of propositions, are not considered.[1]

A major historiographic development was undertaken by Jean van Heijenoort (1912–1986) in 1967, when he published his anthology *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931* [128], which undertakes to trace, through the most salient and representative publications, the origins and formation of mathematical logic, or, "modern" logic, which he traced to 1879 and the publication in that year of the *Begriffsschrift*

For a detailed account of the full panoply of characteristics of modern logic claimed to have been due to Frege, but which occur in the classical Boole–Schröder calculus, many of them due to Peirce, see [4].

[1] We know, for example, that there are inferences that fail in traditional logic in the presence of the empty set that are valid in the class calculus and modern propositional logic. For a discussion of the issue with respect to the traditional versus Boolean squares of opposition, see, e.g., Wu [144, 145].

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_5

of Gottlob Frege (1848–1925), and which culminated with the incompleteness theorems of Kurt Gödel (1906–1978) in 1931. In this work, van Heijenoort remarked [128, vi] that “Considered by itself, the period would, no doubt, leave its mark on the history of logic, but it would not count as a great epoch,” but a great era in the history of logic did open in 1879, when Gottlob Frege’s *Begriffsschrift* was published [128, vi]. Not only that, together with the *Arithmetices principia* [61] of Giuseppe Peano (1858–1932), it also initiated a “new era.” In a posthumously published manuscript of 1974 van Heijenoort [134] made an even stronger remark, “Modern logic began in 1879, the year in which Gottlob Frege (1848–1925) published his *Begriffsschrift*.” This seems to be a very far cry, for example, from Peirce’s [87] judgment that on the history of modern or “exact” logic, that logic was begun by De Morgan. His article, *Logic as Calculus and Logic as Language*, [129]was published at the same time when *From Frege to Gödel* was published. van Heijenoort sought to explain and analyze the differences between the algebraic tradition and the logistic tradition, or, to use the terminology subsequently proposed by Hans Sluga [122]: the “Booleans” and the “Fregeans.” Specifically, whereas both traditions were inspired by Gottfried Wilhelm Leibniz (1646–1716) (see especially Leibniz’s “Fundamental calculi ratiocinatoris” of *circa* Summer of 1688; see Leibniz [52, pp. 917–922]; see also Leibniz [53]), the algebraic tradition took logic to be a calculus, inspired by Leibniz’s *calculus ratiocinator*, in which one dealt with a specific, limited universe of discourse, appropriately defined for the occasion, and the logicist tradition took logic to be a language, inspired by Leibniz’s *lingua universalis* or *lingua characterica*, in which the universe of discourse was *the universe*.[2] This universe was fixed; but it encompasses every universe of discourse, in effect all that there is, allowing it to formulate propositions about everything

[2] Esquisable [20] understands Leibniz’s *characteristica*, however, to be a formal combinatorial system, in which symbols (such as numbers as the quintessential symbol) represent symbolic knowledge (*cognitio symbolica*) that is subject to manipulation in accordance with combinatorial rules. Though symbolic knowledge is entirely blind (*cognitio caeca*; *cognitio suppositiva*), it differs from intuitive knowledge (*cognitio intuitiva*) by its complexity and symbolic mediation and formal representation. Leibniz’s earliest treatment of symbolic knowledge is found in his “Meditationes de cogitationes, veritate et ideis” of November 1684; see Leibniz [52, pp. 585–592]. Legris [50] examines Frege’s use of this Leibnizian concept to evaluate Frege’s Begriffsschrift notation. Legris [51] also uses the analysis of the *cognitio symbolica* of Esquisable [20] to examine the extent to which Boole, Schröder, and Frege developed mathematical logic as the attainment of Leibniz’s program, and concludes that both Schröder’s algebra of relatives and Frege’s *Begriffsschrift* share the goal of simultaneously constructing both a language and a calculus, and that, as a consequence, van Heijenoort [129] and Hintikka [40] made too sharp a distinction between logic as calculus and logic as language. Legris [50] examines Frege’s use of this Leibnizian concept to evaluate Frege’s Begriffsschrift notation. See also Patzig [60] for an account of Frege’s and Leibniz’s respective conceptions of the *lingua characteristica* (or *lingua charactera*) and their relationship. Patzig [60, p. 103] notes that Frege wrote of the idea of a *lingua characteristica* along with *calculus ratiocinator*, using the term “lingua characteristica” for the first time only in 1882 in “über den Zweck der Begriffsschrift” (in print in [24]) and then again in “über die Begriffsschrift des Herrn Peano und meine einige” [27], in the *Begiffsschrift* [21] terming it a “formal language” – “Formelsprache.” In the foreword to the *Begriffsschrift*, Frege wrote only of a *lingua characteristica* or “allgemeine Charakteristik” and a *calculus ratiocinator*, but not of a *charcteristica universalis*. Peckhaus [75] likewise argues that both logic as calculus and logic as language are to be found in the work of Schröder. Other aspects of demonstrating the extent to which Schröder’s work satisfied the criteria defined for a modern mathematical logic include Peckhaus [71–75] and Thiel [124, 125], and Peckhaus [76] is an extensive sustained survey of Schröder’s work.

that can be said, without changing one's semantic reference. For Frege, for example [26], the comparison of the Boolean conception of logic given by Ernst Schröder (1841–1902) in his review [115] of the *Begriffsschrift* is erroneous, based upon a misunderstanding of what the *Begriffsschrift* has taken as its goal and has accomplished. That is, Frege [26] declared that while George Boole (1815–1864) had produced a *mere* calculus, he had in his own *Begriffsschrift* produced *both* a calculus *and* a language, but a language first and foremost.[3] Similarly, he declared [27] that, whereas Peano had moved in the direction of producing a *caracteristica*, his *Arithmetices principia* nevertheless still remained essentially a calculus, not a language, and thus a pasigraphy. Schröder [118, 119], meanwhile, set out to compare Peano's pasigraphy, as displayed in the *Formulario* [64–67], with that of Peirce (albeit using his variation of Peirce's notation), and found it to be superior to that of Peano.

It is not an easy task to explicate the conception of logic of Charles Sanders Peirce (1839–1914), since his use of the term evolved over time, and many of his pronunciations seem to be contradictory. His use of the term "logic" seems, at best, equivocal, since in a broad sense, it encompasses the entire theory of signs, or semiotics; at other times, taken more narrowly, he called that branch of semiotics "critic," which is best understood as the formal theory of deduction, akin to what Aristotle called "analytic" and medieval logicians called "dialectic." If pressed, I would argue that, given Peirce's use of the broad sense of logic as semiotic and the narrow sense of logic as "critic," it served him both as a language and as a calculus. Peirce himself sought to explain his equivocal use of logic, for example, in a letter to Victoria Welby (1837–1912) of March 14, 1909, by noting [98, 421] that he "at first defined logic as the general science of the relation of *symbols* to their *objects*. And I still think that this defines the *Critic of Argument* which is the central part of logic."

In what follows, I want to emphasize in particular the semiotic aspect of Peirce's conception of logic. And, taking critic as a branch of semiotic (along with speculative grammar and methodeutic),[4] it follows that, for Peirce, logic serves both as a calculus and as a language. It is the speculative grammar that establishes the vital connection between logic as a calculus and logic as a language.

Frege evidently borrowed the term "Begriffsschrift" from the characterization [127, 4] by Friedrich Adolf Trendelenburg (1802–1872) of Leibniz's general characteristics. The term had already been used by Wilhelm von Humboldt (1767–1835) in the treatise "Ueber die Buchstabenschrift und ihren Zusammenhang mit dem Sprachbau" of 1824 (see [43]) on the letter script and its influence on the construction of language, and by Franz B. (also Frantisek Bolemír) Kvet (1825–1864) in *Leibnitz'ens Logik*. Trendelenburg's work was known to, and cited, by both Frege in the *Begriffsschrift* [21, V] and by Schröder in both *Der Operationskreis des Logikkalkuls* [114, VI] and in the *Vorlesungen über die Algebra der Logik* [116, I, 38].

[3] Korte [46, 183], however, rejects the nearly universal claims, not only by Frege, but by van Heijenoort and Sluga, among others, that the *Begriffsschrift* is, indeed, a language, but *not* a calculus. Korte's claim is argued upon the basis of Frege's logicism.

[4] Peirce also divides the branches of semiotics into "pure grammar," "logic proper," and "pure rhetoric," the latter his alternative name for methodeutic (*q.v.* [97, 99]).

Peirce [92, 2.227;[5]], [97, 98] wrote: "Logic, in its general sense, is, as I believe I have shown, only another name for *semiotic* (σημειωτηκή), the quasi-necessary, or formal, doctrine of signs."

We can use this as a starting point to understand his conception of logic and to understand his contributions to and role in the development of logic.

In his algebra of relatives, Peirce introduced the concept of truth-functional analysis, defining, in particular, his relation "$\prec$" such that the expression "$a \prec b$" is defined as being true whenever b is true, and false only in case b is false while a is true, that is, as material implication. With this, Peirce explicitly introduced formal logic into the concept of syntactic truth, or validity, which we understand as truth invariant with respect to the extra-logical references of the terms of the proposition "$a \prec b$." Considering the matter in more detail, we note that in the manuscript "On the Algebraic Principles of Formal Logic" written in the autumn of 1879 – the very year in which Gottlob Frege's *Begriffsschrift* appeared, Peirce [103, 23] explicitly identified his "claw" as the "copula of inclusion" and defined material implication or logical inference, illation. From there, he immediately connected his definition with truth-functional logic, by asserting [103, 23] that:

> This definition is sufficient for the purposes of formal logic, although it does not distinguish between the relation of inclusion and its converse. Were it desirable thus to distinguish, it would be sufficient to add that the real truth or falsity of $A \prec B$, supposes the existence of A.

The following year, Peirce continued along this route: in "The Algebra of Logic" of 1880 [81, 21], [103, 170], where $A \prec B$ is explicitly defined as "A implies B."

In the manuscript fragment "Algebra of Logic (Second Paper)" written in the summer of 1884, Peirce [103, 111–115] reiterated his definition of 1880, and explained in greater detail there [1989, 112] that, "In order to say 'If it is a it is b,' let us write $a \prec b$. The formulae relating to the symbol '$\prec$' constitute what I have called the algebra of the copula … The proposition $a \prec b$ is to be understood as true if either a is false or b is true, and is only false if a is true while b is false."

It was at this stage that Peirce undertook the truth-functional analysis of propositions and proofs, and also introduced specific truth-functional considerations, saying that, **v** the symbol for "true" (*verum*) and **f** the symbol for false (*falsum*), the propositions $\mathbf{f} \prec a$ and $a \prec \mathbf{v}$ are true, and either one or the other of $\mathbf{v} \prec a$ or $a \prec \mathbf{f}$ is true, depending upon the truth or falsity of a, and going on to further analyze the truth-functional properties of the "claw."

In its formal sense, what I have called the formal or syntactic aspect of Peirce's semiotics is *logical critic*, the analysis and articulation of formal deductive reasoning, what in the traditional Aristotelian terms is the analytic of arguments. It is the "science of the sheer form of thought in general" [99, 164].

Thus, it is now made entirely explicit that the truth or falsity of a proposition is dependent exclusively and wholly upon the structure of the proposition, that is, upon the definition of the connective or relation and the truth of the terms (or relata) of the propo-

[5] The standard mode of reference in Peirce scholarship for citing material from Peirce's *Collected Papers* is by volume number and paragraph; thus, e.g., 4.239 cites volume IV, paragraph 239 of the *Collected Papers*.

sition. Peirce explained that, in asserting that logic is semiotic in the sense of being "quasi-necessary or formal," he understands that it is concerned with signs and their character, which we consider through abstraction. This is precisely how Hilbert's remark was meant to be understood, that "Man muss jederzeit an Stelle von Punkten, Geraden und Ebenen Tische, Stühle oder Bierseidel sagen konnen" (as quoted in Blumenthal [8, 403]; see also Freudenthal [29, 391]). It forms the basis of the formalist philosophy of mathematics that the syntactic structure of propositions, determined by the definitions, axioms, and logical inference rules of a mathematical system, is the basis for the validity of the formulas (or propositions) derived within that system. That is, "logic proper is the formal science of the conditions of the truth of representations ..." [97, 99], [92, 2.229]. The logician is interested in the "formal responsibility" for the truth of propositions; hence, "The question for him (the logician) is: What is the nature of the sort of sign of which a principle variety is called a proposition, which is the matter upon which the act of judging is exercised?" [97, 103], [104, 292]. The difference between a formula and a proposition for Peirce is that the former makes no commitment to extra-logical assumptions about the truth which it expresses, in Peirce's words, "carries no positive truths," but "must hold in *any* universe" [104, 382][6].

For Peirce, the three subdivisions of semiotic, or the theory of signs included syntactics, semantics, and pragmatics. The latter is concerned with the relation of signs to interpreters. Modern mathematical logic, to the extent that it is formal in the Hilbertian sense, dismisses pragmatics, relegating it to the realm of psychologistic philosophies of logic or the concern of psychology, rather than of logic properly so-called. Semantics, however, is a critical component of modern logic.

The semantic component of Peirce's algebra of relatives rests upon the concept of the *universe of discourse* first formulated by Augustus De Morgan (1806–1871), who was the first to introduce the concept of a universe of discourse, or, as he termed it, a "universe of a proposition, or of a name" [17, 380], [18, 2] that, unlike the fixed universe of all things that was employed by Aristotle and the medieval logicians, and remained typical of the traditional logic, "may be limited in any manner expressed or understood." The concept was subsequently borrowed by George Boole, who in *The Mathematical Analysis of Logic* [9] used algebraic symbols to represent classes and members of classes, without, however, using De Morgan's terminology [9, 5]. He called the entire universe of discourse 1 and the empty or null class 0, and under the appellation *the Universe*, he understood it "as comprehending every conceivable class of objects whether actually existing or not ..." [9, 15].

Peirce (in MS 493, an undated, unpaginated manuscript notebook) defined the *universe of discourse* as "aggregate of the individual objects which 'exist', that is are independently side by side in the collection of experiences to which the deliverer and interpreter of a set of symbols have agreed to refer and to consider." The extensional conception of a universe of discourse, comprised of individuals and classes, was adopted by Peirce partially from De Morgan, but also partially from Mitchell, who added the concept of *dimensionality* to De Morgan's universe [92, 2.536]. Underlying the semantic interpretation of a universe of discourse for Peirce was the ontological commitment to individuals and the classes to

[6] My emphasis.

which they belong. The inhabitants of the universe of discourse may be physical, determined by experience, through the senses; or they may be imaginary, as populated by the contents of a work of art. As Peirce [92, 2.536]) wrote:

> In every proposition the circumstances of its enunciation show that it refers to some collection of individuals or of possibilities, which cannot be adequately described, but can only be indicated as something familiar to both speaker and auditor. At one time it may be the physical universe of sense, at another it may be the imaginary 'world' of some play or novel, at another a range of possibilities.

This suggests that, for Peirce, like Hilbert, the universe of discourse is chosen at will, and depends upon the particular circumstances of the discourse. In other words, the propositions of logic with which one deals may be propositions concerning either tables, chairs, and beer mugs, or points, lines, and planes. What is subject to change is the semantic interpretation of a logical system; what remain unchanged are the formal logical properties of the logical relations between the terms (relata) of propositions (relations) as defined by the syntactic structure of the propositions, the axioms of the systems, and the inference rules holding for the connectives (relations) of the system. Using a textbook example, whereas it is both syntactically valid and semantically true that, "If all Greeks are men, and all men are mortal, then all Greeks are mortal," it is syntactically valid and semantically false (or meaningless) that, "If all boojams are snarks and all snarks are burdips, then all boojams are burdips." As further expressed by Peirce [96, 6.351], therefore,

> "... I wish my description of what is true or false, to apply to what is not only true or false generally, but also to what is true or false under conditions already assumed. Whatever may be the limitations previously imposed, that to which the truth or falsity is limited may be called the universe of discourse. For example, at the mention of a certain name, every person initiated into the Eleusinian mysteries invariably experiences a feeling of awe. This is true. It is therefore true that every person initiated into the Eleusinian mysteries always experiences a sentiment of awe; not universally, but only under the limitations already understood before this is said."

The formalism that rests upon the syntactical or structural relation between terms of propositions and between propositions in determining the validity or truth of arguments or proofs, in Peirce's terminology, the formal conditions of truth, coupled with the extra-syntactic dependence upon the semantic interpretation of propositions, rooted in the universe of discourse, in the case of Boole and Peirce these being defined by classes, is one of the principal conditions or characteristics of modern formal logic. Historiography initially attributed the unification of the semantic and syntactic streams of logic to the work in the first instance of Leopold Löwenheim (1878–1957) [54] and Thoralf Albert Skolem (1887–1963) [121], and in the second instance to Jacques Herbrand (1908–1931) [33], who applied the definitions of the universal and existential quantifiers, taken from Peirce as presented in the *Vorlesungen über die Algebra der Logik* [116] of Ernst Schröder (1841–1902), treating logic as a calculus, in terms of logical sums and products, and applied it to the concept of logic as language, dealing with a universal universe of discourse, Frege's *Universum*, and applying David Hilbert's (1862–1943) concept of proof as formal (referring to [35–38] and [39] – to which I would have added [34]) to Bertrand Russell's (1872–1970) conception, as found in the *Principia Mathematica* [139, 140], of logic as language in which the universal universe of discourse is the sum total of all classes of classes. The contrast between logic, on the one hand, as a *mere* calculus, exemplified by the algebraic logicians, Boole, De Morgan, Peirce, and Schröder, and, on the other hand, as a language, exemplified by Frege and Russell, was first enunciated by van Heijenoort [130]. It is clear, however, that Peirce had already taken into consideration both aspects of logic, but did not develop it methodically or systematically. Indeed, in some respects he is close to Hilbert; specifically, to the extent that for Hilbert, as for Peirce, the major task for logic is to serve as a metamathematics. Expanding upon Peirce's contrast between logic and mathematics with his father Benjamin Peirce's (1809–1880) conception of mathematics as "the science which draws necessary conclusions" [78, 97]; C. Peirce [93, 3.558], [94, 4.229], and Charles Peirce's [94, 4.239], that logic is "the *science* of drawing conclusions," we may take as elucidating the distinction Peirce's [94, 4.373], assertion that the task of the logician "is simply and solely the investigation of the theory of logic, and not at all the construction of a calculus to aid the drawing of inferences." Van Heijenoort in "Set-theoretic Semantics" [130] elaborated this distinction as the one between syntax and semantics, and held that, the algebraic logicians, in producing logic as a calculus, provided the concept of a semantic as an interpretation for their calculi (adding to van Heijenoort's claim that the basis for this aspect of modern logic is owing to De Morgan and his conception [17, 380], [18, 2] of universes of discourse), whereas from Frege and later Russell and to some extent Peano comes the concept of logic as a formal system, providing a syntax.[7] It is the combination of these two aspects of language in the work largely of Löwenheim and Skolem, who arose from the semantic stream, with the syntactic stream, that established modern mathematical logic as typified by the first-order predicate, or first-order functional logic.

The argument that it was Frege's [21] *Begriffsschrift* that inaugurated the era of modern mathematical logic as a *lingua universalis* or *lingua characterica*, rather than a mere *calculus ratiocinator*, and that the logic of Boole, De Morgan, Peirce, and Schröder satisfied the condition of being a calculus, but not a language, was formulated by Frege himself, who designated his *Begriffsschrift* a *Formelsprache* [21, X–XI], [27, 371]. In defense of this claim, Frege and his modern adherents point to Frege's *Universum* as a fixed and universal universe of discourse, outside of which there is nothing. This universality enables one to formulate propositions within the logical system about anything whatever. the fact that Frege chose to replace the subject–predicate syntax that held since Aristotle with the function–argument syntax is a subsidiary, if still essential, issue in the claim that the *Begriffsschrift* is both a calculus and a language. Van Heijenoort [133] and others argue that the universality of Frege's *Begriffsschrift* and that of Whitehead and Russell's *Principia Mathematica* rule out the possibility of asking about the properties (such as completeness and consistency) of their logical systems, because there is nothing extra systematic. The corollary claim is that the reintroduction, by Löwenheim, Skolem, and Herbrand, of the calculus aspect of the algebraic logic of Boole, De Morgan, Peirce, and Schröder, and, thus, the simultaneous reintroduction of their model-theoretic approach through universes

[7] See, e.g., Thiel [126] for a discussion of the relation of syntax and semantics according to Frege; see Anellis [4] for an analysis of the relation between these dichotomies and where Peirce's work in particular stood.
Anellis [4, 261–262] argues that, even if Peirce had no explicit and formal definition of *formal system*, it is present and at least implicit in his work, for example in his [82] "*On the Logic of Number*" providing an informal axiomatization of number theory.

of discourse, enabled logicians thereafter, and in the wake in 1931 of Kurt Gödel's [30] incompleteness theorems for *Principia*-like systems, to begin to ask about the properties of logical systems, and thereby establishing first-order predicate logic as the exemplar of modern mathematical logic (*q.v.* [6, 11, 129]). For Leibniz in his "Fundamenta calculi ratiocinatoris," the *calculus ratiocinator* and the *lingua characterica* were indeed intended to be two aspects of a single program, of establishing a logical *mathesis universalis*.

Peirce's conception of logic as semiotic, however, encompassing semantic and syntactic aspects, and in which one is free to operate either with a specific universe of discourse or with the universe that includes both the actual and the possible, presents his algebra of relatives both as a calculus and as a language, and one in which, as a calculus, the focus is upon the formal truth (or, strictly speaking, the validity) of formulas, and one in which, as a language, the focus is upon the positive truth of propositions. This is the case underlying Schröder's argument of 1898 [118, 119] that Peirce's system, but not that of Peano, provides a pasigraphy, or formal language; Schroder denies Peano's claim in *Notations de logique mathématique* [64, 52] to having attained Leibniz's aim of providing a pasigraphic language – that "Le probléme proposé par Leibniz *est* (donc) résolu," and compares Peirce's logic of relatives with Peano's logical system, to the detriment of Peano; but even so admitted that there was still much work to be done, and recognizes Peirce's logic of relatives as a *calculus ratiocinator*, which is capable of serving as the formal basis of a *lingua characteristica* or *scriptura universalis*. But Schröder doubts that such a *lingua* could be constructed, even in principle. Meanwhile, he dismissed Frege's *Begriffsschrift* as having been surpassed while failing to take account of the work already done. But Schröder did not take into account the semantic aspects of Peirce's conception of logic as semiotic.

The connection between the semantic and the syntactic is established in the case of both Peirce and Frege within the context of the universe, or universe of discourse. For both Peirce and Frege, the Universe, whatever its ontological (or existential) cardinality is, is ultimately resolved into two objects. For Peirce in "Truth and Falsity and Error," following Boole in *The Mathematical Analysis of Logic* [9], the universe of discourse resolves into two classes, X and not-X; "Truth is a character which attaches to an abstract proposition..." [95, 5.567]; moreover, "Truth and falsity are characters confined to propositions.... To say that a proposition is true is to say that every interpretation of it is true," and to be false entails that there is at least one interpretation for which the character asserted by the proposition fails [95, 5.567]; this is the basis upon which a proposition must either be true or false. For Frege, the *Universum*, which is fixed and includes every [logical] object (*Gegenstand*) reduces to two objects: *The True* (*das Wahre*) or *The False* (*das Falsche*); and every proposition is the *name* either of *The True* or *The False*, or, more formally, the meaning (*Bedeutung*) of propositions are the names of a truth values (*Namen von Wahrheitswerthen*) [25, 34].

We turn now to a detailed historical account and analysis of Peirce's development of his pasigraphy or semiotics and of the algebraic logic which he and his coworkers developed as the calculus for his pasigraphic system.

2 Peirce's Semiotics and Pasigraphy[8]

Charles Peirce has been considered by many to be the founder of modern semiotics, and his theory of signs has not only attracted the attention of historians and philosophers of logic and language, but has also been the basis or starting point for much work in semiotics. Peirce was also deeply involved in the technical development of logic in the 19th century, and was an important and direct influence upon the work in the algebra of logic of Ernst Schröder. Indeed, Schröder's *magnum opus*, his *Vorlesungen über die Algebra der Logik* [116] can readily be characterized as the codification and systematization of Peirce's work in logic, which in turn can be understood as the unification of the logic of relatives of De Morgan with the algebraic logic of Boole. Moreover, although Peirce was preceded by a half-dozen years by Gottlob Frege in the introduction of quantifiers into logic, and, hence, contributing to the development of first- and higher order predicate logic, Peirce's quantification theory had an immediate and profound impact upon the work of other logicians, whereas Frege's quantification theory largely languished in the library, awaiting the attention paid to it by Bertrand Russell in 1903 in the *Principles of Mathematics* [112].

In a letter of 1871 to Peirce concerning notations employed for symbolizing algebraic operations, James Edward Oliver (1829–1895) commends Peirce for the invention of new symbols, and in particular expresses agreement with Peirce's contention that the introduction of new concepts in mathematics which are outside the parameters of the established use of symbols that represent settled concepts, warrant, and indeed require the creation of new symbols (see Peirce [100, 492–497]). Oliver does not say, in his letter, which of Peirce's writings he has in mind, but it is clear from the context that he is referring to Peirce's paper of 1870, "*Description of a Notation for the Logic of Relatives...*" [79], in which, among other things, Peirce introduced his "claw" ($-\!\!<$), which does duty, in this particular context as a copula, and by virtue of which Peirce was able to express Aristotle's Barbara *syllogism*, "All S are M, all M are P; therefore, all S are P" as $((S -\!\!< M)\&(M -\!\!< P)) -\!\!< (S -\!\!< P)$. This is the work in which Peirce introduced an algebraic format into De Morgan's logic of relations (for which, see Merrill [58]). He employed algebraic symbols extending to relations those used by Boole for class terms. This new sign was understood by Peirce to be the logical analogue of "$\leq$," the mathematician's sign for the relation of less than or equal to. In order to suggest, however, that the relation he had in mind is less complex than less than or equal to he chose to invent a symbol which was typographically less complex than "$\leq$." Peirce's choice of the "claw" over the algebraist's "$\leq$" would seem to suggest that he considered de Saussure to be mistaken in claiming that arbitrary signs are better than what de Saussure called "symbols," that is, signs for which there are connections between the sign and its meaning. That this is what Peirce believed is confirmed by his later creation of a new notation for the 16 binary connectives which was carefully designed precisely to conform to their logical meanings (and to avoid arbitrary elements).

[8] This section is an expanded version of the original English text, translated into French by Jean-Marie Chevalier, under the title "La Logique et la théorie de la notation (sémiotique) de Peirce" [5].

Initially, Peirce conceived of his "claw" as a symbolic representation of inclusion rather than as implication. More specifically, Peirce most probably initially had in mind for this connective general inclusion, that is, any relation that is transitive, reflexive, and asymmetrical, rather than material implication specifically (see Houser [41, 431, 437 n. 6]). However, its interpretation, as either inclusion or as implication, could vary according to the specific context in which it occurred. Its use was also further adapted as set membership. One of the arguments that Bertrand Russell later set forth against the Boole–Peirce–Schröder notation as against the notation of Peano, as a proof of its inferiority as compared with the latter, was precisely that the former failed to distinguish between these interpretations. Peano employed the "horseshoe," a backward "C" (which evolved into the more familiar "⊃"), to symbolize implication for "conseguenza," which was subsequently adopted by Russell, and the epsilon, written as "∈," to symbolize set membership [61]. Russell told Welby in 1904 [113] that "a symbolism based on Peano's is practically more convenient" than the Boole–Peirce–Schröder notation. On the question of the logic of relations, Russell [111] raises specific criticisms of Peirce's work, granted, as we remarked, that in 1870 Peirce introduced the same symbol ($-\!\!<$) for class inclusion and for implication. This was seen by Russell [1901] as a serious weakness; and he takes pains there to distinguish his own work from that of Peirce and Schröder. (His complaint is presumably based upon his own distinction [112, 187] between "Universal Mathematics," meaning universal algebra in what he understood to be Whitehead's sense, and the "Logical Calculus," the former "more formal" than the latter. In particular, for Universal Mathematics, the signs of operations are variables, whereas for the Logical Calculus, as for every other branch of mathematics, the signs of operations have a constant meaning.) But Russell misses the point, inasmuch as he interprets Peirce's notation to be a conflation not of class inclusion with implication, but of class inclusion with set membership. In Peirce's opinion his algebraic logic of 1880 was highly abstract and could be interpreted as a class calculus, predicate calculus, or propositional calculus. (This erroneous interpretation is reiterated by Kennedy [45, 367–368].) In a letter to Russell of January 27, 1901, Louis Couturat (1868–1914) expresses his agreement with the need to distinguish implication from set membership [16]. In fact, however, Peirce made no distinction in his work prior to "On the Algebra of Logic: A Contribution to the Philosophy of Notation" [84] between sets and classes, and so the charge that he conflates the notation for class inclusion with the notation for set membership is moot for Peirce, if there is an issue here at all. Schröder too, following Peirce, used the same symbol "€," for class inclusion and implication, for which he, in turn, was criticized by Frege [26].[9]

In the manuscript "On the Algebraic Principles of Formal Logic" written in the autumn of 1879 (in Peirce [103, 21–37]), Peirce explicitly identified the "claw" as the "copula of inclusion" and defined it [103, 23] and defined material implication or logical inference as

$$1^{\text{st}},\quad A -\!\!< A,\ \text{whatever } A \text{ may be.}$$

$$2^{\text{nd}},\quad \text{If } A -\!\!< B,\ \text{and } B -\!\!< C,\ \text{then } A -\!\!< C.$$

[9] For details of Russell's and Peirce's accounts and criticisms of one another's contributions to logic, see Anellis [1995]; for discussion of the views of Peirce and his adherents towards Russell's work in logic, see [3].

From there he immediately connected his definition with truth-functional logic, by asserting that:

> "This definition is sufficient for the purposes of formal logic, although it does not distinguish between the relation of inclusion and its converse. Were it desirable thus to distinguish, it would be sufficient to add that the real truth or falsity of $A \mathrel{-\!\!<} B$, supposes the existence of A."

Schröder replaced Peirce's claw with the symbol "€," which he named *Subsumption*, and, like Peirce's symbol, was capable, depending upon the specific context in which it was employed, of doing duty as either implication, set membership, or inclusion (the copula of predication).[10]

In his own "Elementary Notes. I. General and Logico-Mathematical Notation" published in 1888 in the *Annals of Mathematics*, which is a disorganized catalog of mathematical concepts and associated notational devices that had been introduced into mathematics research works and textbooks in modern times, much of the discussion focuses on diverse applications of the copula, only one of which is arithmetic identity. Among the variety of uses, we are most interested in the logical copula, about which Oliver [59, 192] wrote:

> We need a logical copula to assert the dependence of one statement upon another without first asserting the truth of either. This, ∴ and ∵ cannot do; but it is done by Prof. C.S. Peirce's $\mathrel{-\!\!<}$ and $\mathrel{>\!\!-}$; thus $P \mathrel{-\!\!<} Q$, or $Q \mathrel{>\!\!-} P$, means "If P be true, so is Q;" and $(A = B) \mathrel{-\!\!<} (C = D)$, or $(C = D) \mathrel{>\!\!-} (A = B)$ means "If $A = B$, then $C = D$." Such a statement neither implies nor denies its own converse; but we might write $P \mathrel{>\!\!-\!\!<} Q$, for "If P be true, so is Q, and conversely, if Q, then P;" i.e., "Statements P and Q are equivalent;" and this, or some other sign of equivalence, would often be wanted. Peirce's sign of equivalence is not $\mathrel{>\!\!-\!\!<}$, but $=$; which is the best in his studies of Pure Logic, but which, of course, is preoccupied in Mathematics. Hence, it may be as well not to employ the forms $\mathrel{-\!\!<}$, $\mathrel{>\!\!-}$, if, as I think is the case, another form (the square half bracket) is already somewhat used for a quite analogous purpose, and is rather more convenient. When we write
>
> $$C = D, \qquad [(8)$$
>
> the marginal reference, [(8), is commonly a reason for the statement $C = D$; and if the formula (8), referred to, be $A = B$, we have thus virtually written $C = D \;[\; A = B$, in the sense $(C = D) \mathrel{>\!\!-} (A = B)$, above; that is, we have the copulas ["*Is implied by* or necessary to;"] "*Implies* or is sufficient to:"] ["*Is equivalent*, or both necessary and sufficient to." These copulas [,],][, or their equivalents $\mathrel{>\!\!-}$, $\mathrel{-\!\!<}$, $\mathrel{>\!\!-\!\!<}$, can, at need, be variously modified by indices. Thus, we could form copulas for "Does not imply," "Implies but is not implied by," etc., analogous to the familiar "Is not equal to." When a logical copula connects only equations not continued and whose second members are zero, we might, for shortness, write only their first members; as $A + B \,]\, C, D$ for $A + B = 0 \,]\, C = D = 0$; i.e., for $A + B = 0 \,]\, C = 0, D = 0$.

Peirce's notation was developed as a linear expression of formulas suitable for algebraic relations. The notation which Peirce devised was based in large measure upon that presented by Boole and originating as the symbolic representation for an algebra of logic, but which was further expanded by Peirce to accommodate De Morgan's development of relations between terms. For Peirce, the old Aristotelian structure of subject-predicate syntax, which had been arithmeticized by Boole, so that the traditional proposition "All X are Y"

[10] In a review published in *Mind* in January 1892 [48] of the first volume of Schröder's *Vorlesungen über die Algebra der Logik*, Christine Ladd-Franklin remarked on Peirce's view that "for the purposes of Logic, there is no difference between the transitive relation for terms and the transitive relation for propositions" [48, 128] and discussed some respects in which Schröder's views differed from Peirce's.

was rewritten as $XY = 1$, "No X are Y" was rewritten as $XY = 0$, etc., could be still further expanded, so that propositions such as "A is the father of B" could be rewritten as relational expressions, and that relations involving more than two relata could also readily be accommodated, so that we now have, for example, X (Y to express "All X are Y" and X , Y to express "No X are Y." Writing $X \prec Y$ to express "All X are Y" and $X \prec \overline{Y}$ to express "No X are Y," it was a simple matter for Peirce to formulate longer expressions which involved more than two terms. Refining his notation further by the use of indices, Peirce was able to represent relations, so that, for example, if f is the symbol for the relation "father" and a and b are the notational representation of the relata, $f_{a,b}$ expressed the relation "a is the father of b," which was easily distinguished from the relation $f_{b,a}$ to represent the relation "b is the father of a." It is a simple notational matter, then, to express a triple relation, for example f , $f_{a,b,c}$ to express the relation between grandfather, father, and son. Moreover, the use of quantifiers different from the relational term and the relata-terms could be introduced, and the addition of indices applied to quantifiers as well as to relations, enabled Peirce to easily algebraicize Aristotle's categorical propositions. Denoting the existential and universal quantifiers by "$\sum_i$" and "$\prod_i$," respectively, as logical sums and products, e.g., $\sum_i x_i = x_i + x_j + x_k + \ldots$ and individual variables, $i, j, \ldots,$ are assigned both to quantifiers and predicates, Peirce was able not only to express relations, but to enumerate the terms (or conditions) that were required to render a quantified proposition true or false.

It is Peirce's notation that Schröder, with minor variation, adopted in his *Vorlesungen*. And this is the notation that logicians employed until it was in the main supplanted, starting, under the influence of Russell's *Principles*, of Peano's notation.[11] (The Peirce–Schröder symbols for the quantifiers were the primary exception, these being adopted from Hilbert and integrated into the Peanesque notation with which we are most familiar today.) In a letter to Peirce of March 2, 1897, Schröder wrote (as quoted in Houser [42, 223–224]):

> As one instance I should like to draw your attention to the pasigraphic movement *in Italy*. Have you ever noticed the 5 vols. of Peano's Rivista di Matematica together with his "Formulario" and so many papers of Burali-Forti, Peano, Pieri, de Amicis, Vivanti, Vailati, etc. therein, as well as in the reports of the Academia di Torino and in other Italian periodicals (also in the German Mathematische Annalen)?
>
> By the bye you had better not [emphasize] the comparatively trifling divergencies of our systems of notation in view of the contrast with the latter of the one, unanimously employed by those most active Italian investigators, which is, at least with regard to relative notions, so very inferior to ours. I have so to say to stand out nearly alone against them all; whereby "the Good" again and again proves to be an enemy of "the Better" – as is averred by the [German] proverb: Das Gute ist des Bessern Feind.

The origin of the term "pasigraphy" can possibly be traced to Joseph de Maimieu (1753–1820) and his books *Pasigraphie … ou … premiers élémens du nouvel art-science d'écrire d'imprimer en un langue de maniére a etre lu et entendu dans toute autre langue sans traduction* [55], and *Pasigraphie … oder … Anfangsgründe der neuen Kunst-Wissenschaft in einer Sprache alles so zu schreiben und zu drucken, dass es in jeder*

[11] In 1891, in his "Principii di logica matematica," [62, n. 5] remarked on the connection of his "sign of deduction" and notations used by others, including that of Peirce.

andern ohne übersetzung gelesen und verstehen werden kann erfunden und verfasset [56] and Johann Severin Vater (1771–1826) and his book *Pasigraphie und Antipasigraphie; oder Ueber die neueste Erfinderung einer allgemeinen Schriftsprache für alle Völker, und von Wolkens, Leibnitzens, Wilkens und Kalmars pasigraphischen Ideen*. It is possible that Peano, whose interest went beyond mathematics and logic to include artiticial languages, borrowed the term for the linguistic aspects of his axiomatic system from these early uses. Benjamin Hawkins [32, 120] suggests that Schröder was the likely source for the labeling by Peirce of Peano's logical system as a *pasigraphy*.

The value of a uniformization of notation was recognized by late 19th- and early 20th-century logicians, who not only remarked the wide variety of definitions and symbolizations for the most basic elements of logic, but complained about the confusing profusion of notational systems. (See Fig. 1 for the compilation by John Venn (1834–1923) [136] of notations for the most common propositions of logic from the late-17th century to 1879, from the logical systems devised from Leibniz through Frege.) Venn divided the 33 forms into seven different general types. The authors whose notations are considered range from Leibniz to Boole and Hamilton, and from Charles Peirce and his students to Frege. That same year, he was one of those who published a review of Frege's *Begriffsschrift* [135] in which the main point was the cumbrousness of Frege's notation. In 1888, Sophie Willock Bryant (1850–1922), in her article "On the Nature and Functions of a Complete Symbolic Language" [12] – not unnaturally, then – complained of the existence of too many competing logical notations and systems, and she advocated a return to Boole's original system. The last two decades of the 19th century saw a vigorous, if not rancorous, debate about the various notations being devised by logicians; the more crucial aspects of this debate were carried out within the context of the broader philosophical issues of the purpose of the notation as reflective of the nature, purpose and proper role of logic itself. The leading participants in this aspect of the history of the philosophy of logic at that time were Schröder, Frege, and Peano. Schröder's review [115] of Frege's *Begriffsschrift* was the most serious, sustained, systematic, and comprehensive treatment of Frege's tract; nevertheless, Schröder, likewise, found the *Begriffsschrift*-notation to be clumsy and argued that, more importantly, it both ignored the work of the algebraic logicians and failed to add anything new or important. Echoes of Schröder's criticisms of Frege's *Begriffsschrift* can be seen in the thesis [141] of Norbert Wiener (1894–1964) for Harvard University of 1913, in which a comparison is rendered by Wiener between the classical Boole–Schröder algebra of logic as represented in Schröder's *Vorlesungen über die Algebra der Logik* on the one hand and the logic of *Principia Mathematica* on the other, in which Wiener argued against Russell's claims that the Boole–Schröder algebra cannot express what Russell's method expresses.[12] Peirce for his part always nevertheless held that logicians should experiment with alternative notations and that notational systems should evolve as the conceptual systems they represented evolved, since new conceptions call for new symbols. Additionally, Peirce believed that different kinds of logical analysis might best be handled with different notational systems.

Among the recipients of a large number of complaints was the two-dimensional representation of as simple a proposition as A implies B in Frege's *Begriffsschrift*, written

[12] See also [31] for a survey and analysis of Wiener's thesis.

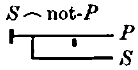
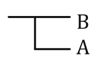
I. Existential			
(1) Negative	1.	$SP=0$	Boole.
	2.	$S(0=P)$	Macfarlane.
(2) Affirmative	3.	$(\bar{S}+\bar{P})_1$	Mitchell.
II. Identity	4.	$S=v(1-P)$	Boole.
	5.	$S=\frac{0}{0}(1-P)$	Boole.
	6.	$S=v(X-P)$	Wundt.
	7.	$S=Sp$	Jevons.
	8.	$S=1-P-y$	Delboeuf, Murphy.
	9.	$S=\frac{P}{\infty}$	Holland.
III. Subsumption	10.	$S<X-P$	Drobisch.
	11.	$S<-P$	Segner.
	12.	$S \angle \bar{P}$	R. Grassmann.
	13.	$S \prec \bar{P}$	Peirce.
	14.	$S \not\leqq P_1$	Schröder.
IV. Mutual Exclusion	15.	$S \overline{\vee} P$	Ladd-Franklin.
V. Implication	16.	$+S-P$	Darjes.
	17.	$S : P'$	McColl.
	18.	$S \frown$ not-P	Pokorny.
	19.		Frege.
VI. Predicational	20.	$Sx=-P$	Maimon.
	21.	$S=-P$	Victorin.
	22.	$S.P$	De Morgan.
	23.	$S)(P$	Wundt.
	24.	$S-nP$	Jäger, Lichtenfels.
	25.	$S>P$	Ploucquet.
	26.	$nS-P$	Ploucquet.
	27.	$tS \parallel tP$	Bentham.
	28.	$S :$ ━━━┿━━━ $: P$	Hamilton.
VII. Notional	29.	$L-S \infty P$	Leibnitz.
	30.	$\frac{S}{m}=\frac{P}{n}$	Lambert.
	31.	$S>\frac{P}{n}$	Lambert.
	32.	$S=-P+M$	Castillon.
	33.	$S \not\leqq P_1$	Husserl.

Fig. 1 John Venn's Classification of Notations for the most common propositions of logic [136, 38] and [137, 481]

as



which Venn [135] termed "cumbrous and inconvenient" as compared with a linear notation such as employed by the majority of logicians, mainly working in the Boolean style, who took their cue from the "symbolical algebra" which Boole, De Morgan, George Peacock (1791–1858), Charles Babbage (1792–1871), Benjamin Peirce (1809–1880), Charles Peirce, Arthur Cayley (1821–1895), and James Joseph Sylvester (1814–1897), and their colleagues were developing through the 19th century. Schröder thought it a "monstrous waste of space" which "indulges in the Japanese custom of writing vertically." (See Fig. 2 for Frege's Theorem 71 in the *Begriffsschrift* notation, which is rendered

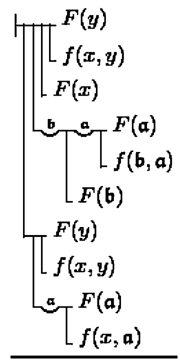


Fig. 2 Frege's Theorem 71, *Begriffsschrift* [21, 59]

as $[\forall a\ (f(x,a) \supset F(a)) \supset (f(x,y) \supset F(x,y))] \supset [\forall b[F(b) \supset \forall a(f(b,a) \supset F(a))] \supset [F(x) \supset (f(x,y) \supset F(y))]$ in the Peano–Russell notation.) One of the few early reviewers who defended Frege's notation against the linear notation of the algebraic logicians was the philosopher Kurd Lasswitz (1848–1910), who [49] attacked the Boole–Schröder notation as "one-sided" and thought Frege's notation in some respects the more preferable. Terrill Ward Bynum noted [14, 19] that Schröder read Lasswitz's review prior to writing his own, and concluded that Schröder consequently came to consider Frege's notation as a "threatening alternative to his own".

In examining the pasigraphy of Peano and his school and comparing it with that of Peirce for the algebra of relatives and as adopted for his own use in the *Algebra der Logik*, Schröder [118, 148], [119, 46] wrote that:

> 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 the 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 (with perfect 'Konsequenz', as we Germans say, or mathematical strictness, 'Strenge').

Schröder explained in his article on pasigraphy that he was led to abandon consideration of Peano's notation in favor of his own modification of Peirces upon coming to the conclusion that Peano's notation was incapable of expressing relations (see Peckhaus [71] for a discussion of Schröder's comparison of Peirce's and Peano's notation).[13]

Schröder [118, 147], [119, 46] then described the purpose of the pasigraphy, or "pasigraphic language," as "to serve and forward on account of its logical structure the purposes of Science; first of all of that science, which the ancient Greek called '*the* science (katexochen)', Mathesis, and next: of Logic and an exact Philosophy. . . ." Schröder's assessment argues for the superiority and greater flexibility of the Peircean pasigraphy

[13] Schröder announced the shift in his attitude towards formal logic as a result of having read Peirce's [83] "The Logic of Relatives." Schröder's also explained his shift in letters to Felix Klein (1849–1925) (see Peckhaus [71, 198–202] for transcriptions of the letters).

in comparison with Peano's, although he evidences considerable respect for the efforts of Peano and his school in that direction. The bulk of his treatment considers in detail the similarities and differences between the Peircean and Peanesque, and in particular examines various examples in their respective treatments of various concepts of logic and set theory. (The specific cases which Schröder offers for comparison would be of interest primarily to the specialist in algebra, algebraic logic, or set theory, and we shall consequently forego these illustrations.) On the other hand, Schröder has little of comfort to say about Frege's *Begriffsschrift*, and dismisses it with but a few words: Frege, Schröder writes [118], [119, 60–61], is "heedless of anything accomplished in the same direction by others, took immense pains to perform what had already been much better done and was therefore superseded from the outset, thus delivering a still-born child. . . ." Schröder had also already undertaken a brief consideration of signs in 1890 [116]. Schröder expressed the nature of his own plan in the *Algebra der Logik* as devising the algebraic logic as: "Der Gedanke einer philosophisch wissenschaftlichen *Universalsprache*" [116, I, 93] in which the "Verwirklichung des gedachten Ideals einer wissenschaftlichen Klassifikation und systematischen Beziehung alles Benennbaren" [116, I, 93] could be fulfilled. The fulfillment of this plan would yield "die vollendete Kenntnis der die Begriffselemente der verknüpfen bestimmte Grund*operationen* und die Bekanntschaft mit deren Gesetzen." Thus, Schröder was conceiving his system, as presented in the *Algebra der Logik*, in terms of an ideal language, which would also serve as the foundation for the algebra of logic and the algebra of relatives. It is on this basis that Peckhaus came to understand that Schröder was moving to adopt a logicist position in philosophy of mathematics [71, 72].

Schröder concluded that the Boolean calculus of equivalent statements offers little more than a system of denotation, while the Peanesque pasigraphy, in its failure to employ Peirce's algebra of relatives, offers an obstacle to the profitable application of Peirce's algebra to the Peano's system of denotation [118, 161], [119, 61]. He compares Peano's system, as presented in the *Formulario*, to a sailboat, in contrast with his own Algebra der Logik, which in comparison with Peano's sailboat, is a steamship [118, 161], [119, 61].

Frege responded to the criticisms of his *Begriffsschrift*, and in particular to Schröder's criticisms, by arguing that the purpose of his work was misunderstood. He argued, in particular, that whereas the algebraic logicians had presented either a *mere* calculus, or at best a calculus which was also, but only secondarily, a language, his *Begriffsschrift* was at once *both* a calculus and a language, and *primarily and foremostly*, a language. We find these in "Booles rechnende Logik und die Begriffsschrift" [22], "Booles logische Formelsprache und die Begriffsschrift" [23], "über den Zweck der Begriffsschrift" [Frege 1883], and in "über die Begriffsschrift des Herrn Peano und meine einige" [27], Frege undertakes to compare his own *Begriffsschrift* with Peano's pasigraphy and to argue that Peano's system of notation is on its way towards becoming a fully fledged *lingua characterica*, but remains has not yet attained that status; that, moreover, while it is a pasigraphy, and more or less a *lingua*, it remains a calculus.[14]

[14] Van Heijenoort [129, 325, n. 3] understood Frege [27, 371] to have asserted that "Boole's logic is a *calculus ratiocinator*, but no *lingua charaterica*; Peano's mathematical logic is in the main a *lingua characterica* and subsidiarily, also a *calculus ratiocinator*, while my *Begriffsschrift* intends to be both with equal stress."

How did Peirce understand the concept of *pasigraphy*? And how did he understand the relation of pasigraphy to logic, to language, and to a calculus? Once these questions are answered, we can the more readily consider the merits of Peirce's notation for logic with its major competitors.

We may begin by noting that the concept of *logic* for Peirce was equivocal. In the narrow sense, it is a *normative* science, establishing the rules for correctly drawing, or *deducing*, conclusions from given propositions. It is on this basis that Peirce was able, as we have seen, to translate the Aristotelian syllogism as an implication. Thus, for Peirce [93, 4.239] "To *draw necessary* conclusions is one thing, to draw conclusions is another, and the *science* of drawing conclusions is another; and that science is Logic." Logic in this usage is a deductive methodology,[15] and in that case a system of logical symbols is the means by which we can "analyze a reasoning into its last elementary steps" [94, 4.239]. In an unpublished manuscript of 1873, intended as part of a larger work on logic, Peirce went so far as to defined logic as a study of signs; he wrote [80], [102, 82–84]:

> A sign is something which stands for another thing to a mind. To its existence as such three things are requisite.
>
> In the first place, it must have characters which shall enable us to distinguish it from other objects.
>
> In the second place, it must be affected in some way by the object which it signifies, or at least something about it must vary as a consequence of a real causation with some variation of its object. One of the simplest examples of this is a weathercock, which is directly moved by the force of the wind. A photograph is caused by a radiant light from the object it represents. In the case of a picture executed by hand the causation is less direct, but none the less exists. The relation of a historical statement with its object is that of being caused by it. If a promise is made, this is a sign of the thing promised only so far as it will itself cause the existence of its being, unless we are to regard it as a prophecy which is caused by that state of mind which will cause the thing prophesied to be carried out. Thus the causation may either be from the object to the sign, or from the sign to the object, or from some third thing to both; but some causation there must be.
>
> The third condition of the existence of a sign is that it shall address itself to the mind. It is not enough that it should be in relation to its object but it is necessary that it shall have such a relation to its object as will bring the mind into a certain relation with that object namely, that of knowing it. In other words, it must not only be in relation with its object, but must be regarded by the mind as having that relation. It may address the mind directly, or through a translation into other signs. In some way it must be capable of interpretation. We have seen that thoughts themselves have intellectual significance only so far as they prove themselves to other thoughts. So that thoughts are themselves signs which stand for other objects of thought. And since, on the other hand, there is no sign which the mind may not make use of in reasoning, it follows that the science of thought in its intellectual significance is one and the same thing with the science of the laws of signs. Now there are many general truths with regard to signs which hold good for all signs whatever, of necessity; being involved in the essential nature of signs. The origin of these principles is undoubtedly the nature of the mind. But they are involved in so much of what is true of the mind as is implied in our capability of reasoning at all and which may therefore be said to be implicitly taken for granted by all men, that is, to be deducible from what everybody agrees to and must agree to before we can begin any discussion whatever in a rational way, and which is thus taken out of the special domain of psychology and made the common property of science. These principles might be evolved from a study of the mind and of thought, but they can also be reached by the simple consideration of any signs we please. Now the latter mode of studying them is much the easiest, because the examination

[15] In a much broader sense, Peirce would also include induction and deduction as means of inference along with deduction.

> of external signs is one of the most simple researches which we can undertake, and least susceptible to error, while the study of the mind is one of the most difficult and doubtful. We shall therefore proceed in the remainder of this part of the work to compare signs, and generalize our results, being guided in doing so by a certain feeling of the necessity that this or that must be true, such as is felt in mathematics the origin of which necessity clearly is, in this case at least, that the principles are involved in the postulate, that the mind is so constituted as to investigate.
>
> The business of Algebra in its most general signification is to exhibit the manner of tracing the consequences of supposing that certain signs are subject to certain laws. And it is therefore to be regarded as a part of Logic. Algebraic symbols have been made use of by all logicians from the time of Aristotle, and probably earlier. Of late, certain logicians of some popular repute, but who represent less than any other school the logic of modern science, have objected that Algebra is exclusively the science of quantity, and is therefore entirely inapplicable to Logic. This argument is not so weak that I am astonished at these writers making use of it, but it is open to three objections: In the first place, Algebra is not a science of quantity exclusively, as every mathematician knows; in the second place these writers themselves hold that logic is a science of quantity; and in the third place, they, themselves, make a very copious use of algebraic symbols in Logic.

For Peirce, *signs* may be either *icons*, an *indices*, or *symbols*, according to whether they are, respectively, in some relevant manner like the objects they represent; or physically or causally related to the their objects; or represent their objects by agreement or convention. Peirce introduced this conception in 1880 in "*On the Algebra of Logic*," and subsequently further developed this characterization. He held that all three of these types of sign are necessary for logic. Schröder [117, 3431] explained that "the sign is a handle by which we take hold of the things of thought." In "*On the Algebra of Logic: A Contribution to the Philosophy of Notation*" [84], Peirce explained his theory of signs. General signs, soon to be called *symbols*, convey conventional meanings; demonstrative signs, called *indexes*, to denote individuals in appropriate universes of discourse; and signs of resemblance, called *icons*, to exhibit relations and structures. It was in this paper that he gave the first clear introduction of truth-values and a decision procedure as a general procedure for determining whether a proposition could be proven to be a theorem.

Once more, Peirce's choice of the "claw" over the algebraist's "$\leq$" would seem to suggest that he would have considered de Saussure was mistaken in claiming that arbitrary signs (Peirce's *symbols*) are better than symbols, as signs – Peirce's *icons* and *indices* –, for which there are nonarbitrary connections between the sign and its meaning.

It is, said Peirce, the icon that represents its object by sharing some quality or structural feature. By examining an icon of an object we are *at the same time* examining a feature of the object itself. Carefully constructed iconic notations provide us with a means for learning about the objects and structures they represent through the manipulation and examination of the signs themselves.

In 1906, Peirce wrote a series of articles for a journal of philosophy in which he defended his idiosyncratic graphical notation for logic. He imagined himself responding to a great general who questioned why one should bother to perfect a notation to represent some course of thought that must already be present to us. Peirce argued that the function of a good diagram in logic is similar to that of a good map in a military campaign. Just as the general might stick pins into a map to experiment with different military strategies so the logician might manipulate his diagrams to learn about the relations or structures they represent. Peirce – whose degree at Harvard was in chemistry – expressed it thusly [88, 493], [94, 4.530]:

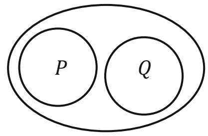
> one can make exact experiments upon uniform diagrams; and when one does so, one must keep a bright lookout for unintended and unexpected changes thereby brought about in the relations of different significant parts of the diagram to one another. Such operations upon diagrams, whether external or imaginary, take the place of the experiments upon real things that one performs in chemical and physical research. Chemists have ere now, I need not say, described experimentation as the putting of questions to Nature. Just so, experiments upon diagrams are questions put to the Nature of the relations concerned.

Peirce then argues that just as the chemist is really investigating the molecular structure of his samples, so the logician is investigating the form of relation instantiated in his diagram. It could be suggested that in a certain respect the chemist appears to have an advantage over the logician; namely, in experimenting with actual samples of the chemical compound under investigation the chemist has a sort of guarantee that the genuine article is being studied. If the molecular structure of interest is H_2O we can hardly go wrong if we experiment with a sample of pure water. Furthermore, because the molecular structure of interest is so precisely represented by our sample, by experimenting with our beaker of water we may learn unexpected and perhaps even unimagined things about H_2O. While it is true that chemists sometimes have this advantage over logicians (and mathematicians), that does not have to be the case. Peirce's idea was that logicians can have this same "advantage" by designing notations to precisely embody the structures they represent. If that were accomplished the logician and mathematician could learn a great deal by the mere rule-governed manipulation of signs.

The solution for Peirce was to devise graphical methods for experimentally manipulating the constituents of propositions to investigate their logical relations. Therefore, beginning around 1890, Peirce began to concentrate on graphical logic. In "*The Critic of Arguments*" [85], Peirce extended the concept of function to many-place function and insists that even algebra is a kind of diagram, but that it is not especially powerful as graphical systems go. After 1892, in a number of papers mostly involving graphical systems or semiotics, Peirce turned his attention to quantification theory. Much of the work which he carried out in logic in the final decade of his life was devoted to developing graphical systems of logic. His entitative graphs came first, and treated propositional logic. The surface of the graph was a sheet which represented a truth-theoretic plane, and the letters representing the terms of the calculus were connected by lines representing the relations between these terms. A "cut" in the sheet, depicted by a circle around a letter representing a term in the universe of discourse, indicated a whole in the sheet, and thus represented the negation, or falsity, of the encircled term. In his existential graphs, the next phase of his work, Peirce used a similar graphical technique to deal with quantified propositions.[16] In the entitative graphs, P together with Q, i.e., their mere concatenation, means "P or Q" while in the existential graphs it would mean "P and Q."

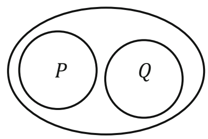
Placing a cut around two concatenated terms P or Q which have each already been negated yields, by De Morgan's Laws, the proposition *not (not-P and not-Q)*, i.e., *P or Q*:

[16] Similar, because it would be incorrect to understood existential graphs are basically just the entitative graphs with quantification.



Entitative graph: "P and Q," i.e., "It is not the case that not-P or not-Q"

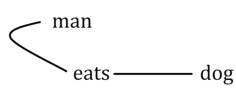
In 1897, as [123] reminds us, Peirce switched to the dual form, existential graphs, which were based on conjunction, negation, and existential quantification. For the existential graphs,



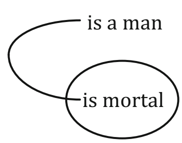
Existantial graph: "P or Q," i.e., "It is not the case that not-P and not-Q"

now represents De Morgan's Law as *not (not-P and not-Q)*, i.e., P or Q.

Peirce continued to develop versions of existential graphs for modal logic and higher-order logic until his death in 1914. Peirce's α-graphs were existential graphs for sentential calculus and Boolean algebra. The empty sheet of assertion again represented the universe of discourse, and the cut indicated the negation of a term. The concatenation of two terms represented their conjunction. Lines of identity [straight or curved] are introduced for β-graphs, which are existential graphs for first-order logic. In the existential graph, writing "man" on the sheet is read "there exists a man." Thus, the graph



read intuitively represents a proposition which should probably be read "Either something is a man or it eats a dog;"; see read existentially, it represents the proposition "Some man eats some dog." The existential graph



is to be read as "There is some man who is not mortal" (see the manuscript "Existential Graphs" [94, 4.407].)

Peirce's γ-graphs, which added elements of color and stacks of sheets of assertion, were still in the experimental stage at the time of Peirce's death, and were developed for modal logic and for higher order quantification theory.

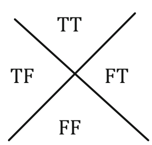
EG, Peirce's existential graphs, are his highly iconic, graphical system of logic. He described them, in the manuscript "Phaneroscopy" of 1906 as a moving picture of thought, "rendering literally visible before one's very eyes the operation of thinking *in actu*," and as a "generalized diagram of the Mind" [88], [94, 4.582].

Also among the graphical treatment for representation of the propositions of logic devised by Peirce were truth tables. Thus, one of Peirce's manuscript of "*The Simplest Mathematics*" written in January 1902 ("Chapter III. The Simplest Mathematics (Logic III)," MS 431; see [94, 4:260–262]) contained a table presenting the 16 possible sets of truth values for a two-term proposition:

Charles Peirce's table for the 16 binary connectives (as presented in Clark [1997, 309])

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
F	F	F	F	T	T	T	T	F	F	F	F	T	T	T	T
F	F	F	T	F	T	F	F	T	T	F	T	F	T	T	T
F	F	T	F	F	F	T	F	T	F	T	T	T	F	T	T
F	T	F	F	F	F	F	T	F	T	T	T	T	T	F	T

As part of this work, Peirce developed a special iconic notation for the 16 binary connectives that enabled him to give a quasi-mechanical procedure for identifying thousands of tautologies from the substitution sets of expressions with up to three term (or proposition) variables and five connective variables. In his X-frames notation, the open and closed quadrants are indicate truth or falsity respectively, so that for example, ⊠, the completely closed frame, represents row 1 of the table for the 16 binary connectives, and $\times$, the completely open frame, represents row 16. The X-frame notation is based on the representation of truth values for two terms as follows:[17]



A well-designed notation enables us to investigate our subject by manipulating the notation. But we must keep our purposes in mind. If we want to investigate inferential relationships between propositions, or discover what conclusions can be drawn from given premises, the Existential Graphs should be our choice. But for investigating the internal structure of propositions, for exploring variations of form, or for studying the group properties of the propositional connectives, EG will rank far behind Peirce's advanced algebraic logic. For Peirce, then, the benefits of designing notational signs to iconically represent their objects are by no means limited to logical connectives. Peirce liked to distinguish mathematics from logic by pointing out that mathematics is concerned with deducing consequences from given assumptions, while logic is concerned with analyzing procedures and methods of reasoning. As a result, mathematicians want efficient and

[17] For details on computational aspects Peirce's X-frame notation and applications, see Clark [15] and Zellweger [146].

powerful notational systems, while logicians tend to prefer weak systems which, though inefficient, are more analytical and reveal more about the inferential process. It is reasonable to suppose that different systems of notation are called for and that a good notation for deduction will model large inferential steps while a good notation for fine analysis will model the smallest steps of inference and the most elementary relations between terms.

In the broader sense, Peirce associated logic with the theory of signs, and thus he devoted considerable attention to the distinctions between icon, index, and symbol. "Logic," he wrote [97, 98], "in its general sense, is ... only another name for *semiotic* (σημειωτηκή), the quasi necessary or formal, doctrine of signs." Adam Vile [138, 1] consequently argues that Peirce's work in logic, "from a semiotic underpinning," was an attempt at developing logic which "aimed at clarifying reasoning and formulating easily understandable general principles of reasoning" [138, 1]. In doing so, he took the rationale for this study of logic to be concerned with what the medieval had called speculative grammar (*grammatica speculativa*), and it is the relation between symbols and the relations between objects or facts that they represent that are examined. The speculative grammar, or pasigraphy, is the syntactic basis for the research into the non-logical universe that provides the model-theoretic semantic – or as Jaakko Hintikka [40] would have it in reference to Peirce, the game-theoretic – interpretation for the research into reasoning. For Peirce, the distinct tools of methodologies of research, and thus the three branches of logic in this broad senese, are deduction, induction, and abduction. Logic in this broad sense, then, is the theory of all reasoning, and the speculative grammar, or theory of signs, inseparably linked to logic, is the syntactic aspect of logic. This conception was inherited by Peirce from the work of Thomas Maulfelt (Thomas of Erfurt; 14th cent.) and his school, in particular, the *Tractatus de modis significandi seu Grammatica speculativa* (see, e.g., [57]), which, in Peirce's day, was still attributed to Johannes Duns Scotus (1266?–1308). As recently expressed by Alessandro Isnenghi [44, 150], for Thomas and his school:

> Logic depends upon and presupposes Grammar, as in Peirce's theory. Grammar does not concern direct references to the thing: on the contrary it is the dimension of truth, belonging to logic, that presupposes the notion of congruity of grammar. ... [T]he existence or non-existence of an object... is something that concerns only the work of the logician.

As Peirce said for himself, for example, in 1903 in "A Syllabus of Certain Topics of Logic,"

> Logic may be regarded as the science of the general laws of signs. It has three branches: (1) *Speculative Grammar*, or the general theory of the nature and meaning of signs, whether they be icons, índices, or symbols; (2) *Critic*, which classifies arguments and determines the validity and degree of force of each kind; (3) *Methodeutic*, which studies the methods that ought to be pursued in the investigation, in the exposition, and in the application of truth. Each division depends on that which precedes it [96, 2:260].

Scotus himself distinguished theoretical logic from practical logic and proper from improper inferences, a distinction which we may suppose played a role in Peirce's differentiation between logic as speculative grammar or semiotics, and logic as a deductive science.

This approaches the notion, formulated by Descartes, Leibniz, and their intellectual heirs, of a *mathesis universalis*. And whereas Descartes had sought to make *more geometrico* the *mathesis*, on the ground that Euclid's *Elements* provided the most perfect example of rigorous deductive reasoning, Leibniz undertook the Project, taken up by Boole, Peirce, and, in different directions, Frege and Russell, to devise a *mathesis* based upon a reformulation of traditional logic as a deductive science. The basis of such a mathesis was to be, on the one hand, a pasigraphy, or formal universal language, a *lingua characteristica*. The deductive aspect was to be provided for the language by a calculus, a *calculus ratiocinator*. For Leibniz, Boole, Peirce, Frege, Russell, and their colleagues, logic was either a calculus, or a language, or both. The question, when considering the primacy of one aspect of logic over another was determined by the selection of the calculus. The *logica utens* was the choice of the "Booleans," or algebraic logicians and the American postulate theorists who followed them: Boole, De Morgan, Peirce, Schröder, Huntington et al. The *logica utens* is a calculus designed for application to a specific universe of discourse, Schröder's *Denkbereich*. The axioms and the primitives, or undefined terms of the syntax, are selected on the basis of the purpose of the specific field for which the system was devised. Thus, for example, in formulating one's axioms, if the purpose was to formulate and develop a geometric theory rather than, say, an area of algebra. The axioms and primitives would be designed that were appropriate to geometry. And if a narrower theory was at issue, one would select axioms and primitives appropriate to, say, projective, rather than to metric, geometry, or to Euclidean, rather than to non-Euclidean geometry, as the case might be. The *logica utens* was not a language, but a calculus, since the primitives to which the axioms were to be applied could be left uninterpreted until after the axiomatic system was developed. This is what David Hilbert (1862–1943) was attempting to articulate when he was reported as saying that it was a matter of indifference whether the primitives of an axiomatic system were composed of *points*, *lines* and *planes*, or *tables*, *chairs*, and *beer* mugs (as quoted in Blumenthal [8, 403] and Freudenthal [29, 391]).

A key idea of the "Leibniz programme" of developing a *mathesis universalis* which was at once both a *calculus ratiocinator* and a *lingua characteristica* was that the logical symbols can, and should, be treated formally, without recourse to special interpretations. According to Boole,

> [W]e may in fact lay aside the logical interpretation of the symbols in the general equation; convert them into quantitative symbols, susceptible only of the values 0 and 1; perform upon them as such all the requisite processes of solution; and finally restore to them their logical interpretation. (italics in the original) [10, 70]

Thus, from the beginning Boole understood that the resulting algebra could be studied either in the abstract as a special branch of mathematics, or under various interpretations as different kinds of logics. He also makes it clear from the outset that his explicit goal is to devise logic "in the form of a calculus" [10, 11]. The general laws of thought for this calculus are "the signs appropriate for the science of logic in particular" and of the investigation which he undertakes for his logical calculus are concerned with "the laws to which that class of signs are subject" [10, 24]. In devising and working through the deductive consequences of his algebra, Boole left his calculus uninterpreted, so that it could perform precisely as a pure calculus, rather than as a language whose formulas or expressions concerned propositions, or classes, or sets.

For Peirce, the distinction between *logica docens* and *logica utens* was consistently formulated in terms of the *logica utens* as a "logical theory" or "logical doctrine" as a means for determining between good and bad reasoning (see, e.g., "The Proper Treatment of Hypotheses: a Preliminary Chapter, toward an Examination of Hume's Argument against Miracles, in its Logic and in its History" (MS 692, 1901) [101, 2:891–892]; from the "Minute Logic," "General and Historical Survey of Logic. Why Study Logic? Logica Utens," *ca*. 1902 [Peirce 1933, 2.186]; "Logical Tracts. No. 2. On Existential Graphs, Euler's Diagrams, and Logical Algebra," *ca*. 1903 [94, 4.476]; Harvard Lectures on Pragmatism, 1903 [95, 5.108]), and the *logica docens* in terms of specific cases. In the entry on "Logic" for Baldwin's *Dictionary* [106, II, 21], Peirce, in collaboration with his former student Christine Ladd-Franklin (1847–1930), wrote:

> In all reasoning, therefore, there is a more or less conscious reference to a general method, implying some commencement of such a classification of arguments as the logician attempts. Such a classification of arguments, antecedent to any systematic study of the subject, is called the reasoner's *logica utens,* in contradistinction to the result of the scientific study, which is called *logica docens.* [See REASONING]
>
> That part of logic, that is, of *logica docens,* which, setting out with such assumptions as that every assertion is either true or false, and not both, and that some propositions may be recognized to be true, studies the constituent parts of arguments and produces a classification of arguments such as is above described, is often considered to embrace the whole of logic; but a more correct designation is Critic (Gr. κριτική. According to Diogenes Laertius, Aristotle divided logic into three parts, of which one was πρὸς κρίσιν). . . .

In the next paragraph, Peirce and Ladd-Franklin [106, II, 21], establish the connection between logic as critic and the *grammatica speculativa*:

> It is generally admitted that there is a doctrine which properly antecedes what we have called critic. It considers, for example, in what sense and how there can be any true proposition and false proposition, and what are the general conditions to which thought or signs of any kind must conform in order to assert anything. Kant, who first raised these questions to prominence, called this doctrine transcendentale Elementarlehre, and made it a large part of his *Critic of the Pure Reason.* But the *Grammatica Speculativa* of Scotus is an earlier and interesting attempt. The common German word is Erkenntnistheorie, sometimes translated EPISTEMOLOGY (q.v.).

Ahti-Veikko Pietarinen [107] has characterized the distinction for Peirce as one between the *logica utens* as a logic of action or use and the *logica docens* as a general theory of correct reasoning. In the terms formulated by van Heijenoort (see, e.g., [129]), a *logica utens* operates with a specific, narrowly defined and fixed universe of discourse, and consequently serves as a *logic as calculus*, and thus as a *calculus ratiocinator*, whereas a *logica docens*, or, more precisely, a *logica magna*, operates with a universal domain, or universal universe of discourse, characterized by Frege as the *Universum*, which is in fact universal and fixed. There are several interlocking layers to van Heijenoort's thesis that, as a result of its universality, it is not possible to raise or deal with metalogical, i.e., metasystematic, properties of the logical system of *Principia Mathematica*. These aspects were dealt with in a series of papers by van Heijenoort over the course of more than a decade. The writings in question, among the most relevant, include "Logic as calculus and Logic as Language" [129], "Historical Development of Modern Logic" (1974) [134]; "Set-theoretic Semantics," [130], "Absolutism and Relativism in Logic" (1979) [131], and "Systéme et métasystéme chez Russell" [133].

At the same time, however, we are obliged to recognize that Peirce's own understanding of *logicautens* and *logica docens* is not precisely the same as we have represented them here and as understood by van Heijenoort. For Peirce, a *logica utens* is, or corresponds to a logical theory, or logic as critic, and *logica docens* is the result of the scientific study, and more akin to an uncritically held but deeply effective logical theory, and hence normative, which because it governs our actions almost instinctively amounts almost to a moral theory.

We should distinguish more carefully *logica docens* from *logica utens* as conceived by van Heijenoort as it relates to Peirce.[18] A *logica utens* is specific calculus designed to serve a specific purpose or narrow field of operation, and is typically associated to one *universe of discourse* (a term coined by De Morgan) which applies to a specific, well-defined domain. The classical Boole–Schröder algebra of logic is understood by van Heijenoort as a *logica utens* in this sense. Although admittedly the universe of discourse can have more than one semantic interpretation, that interpretation is decided ahead of time, to apply specifically to sets, or to classes, or to propositions, but never does duty for more than one of these at a time. In a more practical sense, we might consider the axiomatic systems developed by the postulate theorists, who set forth a specific system of axioms for specific fields of mathematics, and for whom, accordingly, the universe of discourse is circumspect in accordance with the area of mathematics for which an axiomatic system was formulated. For example, we see one for group theory, and another for geometry; even more narrowly, we find one for metric geometry, another for descriptive geometry; etc. The universe of discourse for the appropriate postulate system (or *logica utens*) for geometry would consist of points, lines, and planes; another universe of discourse, might, to borrow Hilbert's famous example be populated by tables, chairs, and beer mugs. We may, correspondingly, understand the *logica docens* as an all-purpose logical calculus which does not, therefore, operate with one and only one or narrowly constrained specific universe of discourse or small group of distinct universes of discourse.

Associated with the *logica utens*/*logica docens* distinction is the logic as calculus/logic as language distinction. *Logic as calculus* is understood as a combinatorial tool for the formal manipulation of elements of a universe of discourse. Typically, but not necessarily, this universe of discourse is well-defined. We should, perhaps, better the logic as calculus on a purely syntactic level. The "Booleans" (and, although van Heijenoort did not specifically mention them, the Postulate theorists), reserved a formal deductive system for combinatorial-computational manipulation of the syntactic elements of their system. The semantic interpretation of the syntactic elements to be manipulated was external to the formal system itself. The semantic interpretation was given by the chosen universe of discourse. Again, the axioms selected for such a system were typically chosen to suit the needs of the particular field of mathematics being investigated, as were the primitives that provided the substance of the elements of the universe of discourse, whether sets, classes, or propositions, or points, lines, and planes, or tables, chairs and beer mugs.

[18] In its original conception, as explicated by the medieval philosophers, the *logica utens* was a practical logic for reasoning in specific cases, and the *logica docens* a teaching logic, or theory of logic, concerning the general principles of reasoning. These characterizations have been traced back at least to the *Logica Albertici Perutilis Logica* (*ca*. 1360); see [1] of Albertus de Saxonia (1316–1390) and then his school in the 15th century, although the actual distinction can be traced back to the *Summulae de dialectica* of Johannes Buridanus (*ca*. 1295 or 1300–1358 or 1360) (see Buridanus [13]). See, e.g., Bíard [7] for Buridan's distinction, and Ebbesen on Albert.

On the other hand, the *logica docens* is intended as an all-purpose formal logical system which is applicable regardless of the universe of discourse which provides the contents for its manipulation, regardless of the primitive terms upon which it operates, or their semantic interpretation, if any. It is in these terms that van Heijenoort also therefore distinguishes between *relativism* and *absolutism* in logic; a *logica docens* is appropriate relative to its specific universe of discourse; a *logica utens* is absolute in being appropriate to any and every universe of discourse. More broadly, there are many *logica utens*es, but only one, universally applicable, *logica docens*.

Pietarinen, for one, would agree with van Heijenoort with regard at least to Peirce, that his work belongs to logic as calculus; as Pietarinen ([108, 19] in the preprint version of [109]) expresses it, "in relation to the familiar division between language as a universal medium of expression and language as a re-interpretable calculus," Peirce and his significant followers took language to serve the latter role." Elsewhere, Pietarinen makes the case even more strongly and explicitly, asserting [110, 63] that: "Peirce's disaffection with unreasonably strong realist assumptions is shown by the fact than he did not advocate any definite, universal logic that would deserve the epithet of being the logic of our elementary thought. Logical systems are many, with variable interpretations to be used for the various purposes of scientific inquiry." Pietarinen makes it clear that Peirce was "not a believer in unrestricted classical logic" and he bases this claim precisely upon the distinction between *logica utens* and *logica docens*.

For Russell, as for Frege, says van Heijenoort, it is the character of this inclusiveness that makes their logical systems suitable not merely as a *calculus ratiocinator*, but as a *lingua characteristica* or *charcteristica universalis*. Thus, Frege's *Begriffsschrift* and Whitehead and Russell's *Principia Mathematica* are *both* calculus and language *at once*. Moreover, Frege would argue, that his *Begriffsschrift*, unlike the calculi of the Booleans, not simply *both* calculus and language, but a language first and foremost. As we know, Schröder and Peano would argue over whether the classical Boole–Schröder or the logic of the *Formulaire* was the better pasigraphy, or *lingua universalis*,[19] and Frege and Schröder, along the same lines, whether the *Begriffsschrift* or the classical Boole–Schröder was a *lingua*, properly so-called, and, if so, which was the better. Van Heijenoort would argue for the correctness of Frege's appraisal.

The other aspect of this universality is that, as a language, it is not restricted to a specific universe of discourse, but that it operates on the universal domain, what Frege called the *Universum*. Thus, the universe of discourse for Frege and Russell is the universal domain, or the universe. It is in virtue of the *Begriffsschrift*'s and the *Principia* system's universe of discourse being *the universe*, that enables these logical systems to say (to put it in colloquial terms) everything about everything in the universe. One might go even further, and with van Heijenoort understand that, ultimately, Frege was able to claim that there are only two objects in the *Universum*: *the True* and *the False*, and that every proposition in his system assigns the *Bedeutung* of a proposition to one or the other.

If there is a difference between Frege and Peirce, it is that, taking logic as a normative science and connecting it explicitly with ethics, Peirce ultimately adheres to the duality of the antipodes of the True and the False into which Frege divides the universe, with

[19] Peano's assessment of Schröder's system is found in Peano [63, 68].

propositions as serving as the *names* of *das Wahre* and *das Falsche*, Peirce [86, 188–189, 200] taking arguments, rather than propositions, as being either good or bad. Peirce says [86, 188]: "the *logica utens* is a particular species of morality. Logical goodness and badness, which we shall find is simply the distinction of *Truth* and *Falsity* in general, amounts, in the last analysis, to nothing but a particular application of the real general distinction of Moral Goodness and Badness, or Righteousness and Wickedness." He is also clear that it is the argument or proof, or as he says, inferences, that are either good or bad; and that [103, 200], since "no argument can exist without being referred to some special class of arguments," "[t]he act of inference consists in the thought that the inferred conclusion *would be true*," making it clear that it is the proof, or argument which, being good or bad, must yield a conclusion which is either true or false, and that, since each argument is a case of a "special class of arguments, these arguments is a representation or name of *Truth* or of *Falsity*."

What makes the logic of the *Begriffsschrift* (and of the *Principia*) a language preeminently, as well as a calculus, rather than a "mere" calculus, was that it is a *logica docens*, and it is absolute. The absoluteness guarantees that the language of the *Begriffsschrift* is a language, and in fact a universal language, and fulfills the Leibniz program of establishing it as a *mathesis universa*, which is both a language and a calculus.

But, because of this universality, there is, van Heijenoort argues, nothing "outside" of the *Universum*. (This should perhaps set us in mind of Ludwig Wittgenstein (1889–1951), and in particular of his proposition 5.5571 of the *Tractatus Tractatus logico-philosophicus* [142], that "The limits of my language are the limits of my world" – "Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt.") If van Heijenoort had cared to do so, he would presumably have quoted Proposition 7 from the *Tractatus*, that, by virtue of the universality of the *logica docens* and its universal universe of discourse, anything that can be said must be said within and in terms of the *logica docens* (whether Frege's variant or Whitehead–Russell's), and any attempt to say anything *about* the system is "wovon man nicht sprechen kann." In van Heijenoort's terminology, then, given the universality of the universal universe of discourse, one cannot get outside of the system, and the system/metasystem distinction becomes meaningless, because there is, consequently, nothing outside of the system. It is in this respect, then, that van Heijenoort argued that Frege and Russell were unable to pose, let alone answer, metalogical questions about their logic. Or, as Wittgenstein stated it in his *Philosophische Grammatik* [143, 296]: "*Es gibt keine Metamathematik*," explaining that "Der Kalkül kann uns nicht prinzipielle Aufschlüssen über die Mathematik geben," and adding that "Es kann darum auch keine 'führenden Probleme' der mathematischen Logik geben, denn das wären solche. . . ."

Turning then specifically to Peirce, we can readily associate his concept of a *logica docens* as a general theory of semiotics with van Heijenoort's conception of Frege's *Begriffsschrift* and Whitehead–Russell's *Principia* as instances of a *logica magna* with logic as language; and likewise, we can associate Peirce's concept of *logica utens* as with van Heijenoort's concept of algebraic logic and the logic of relatives as instances of a *logica utens* with logic as a calculus. It is on this basis that van Heijenoort argued that, for the "Booleans" or algebraic logicians, Peirce included, the algebraic logic of the Booleans was *merely* a calculus, and not a language. By the same token the duality between the notions of logic as a calculus and logic as a language is tantamount to Peirce's narrow conception of logic as critic on the one hand and to his broad conception of logic as a general theory of signs or semiotics. It is on this basis that Volker Peckhaus has concluded [71, 174–175] that in fact Peirce's algebra and logic of relatives "wurde zum pasigraphischen Schlüssel zur Schaffung einer schon in den frühen zeichentheoretischen Schriften programmatisch geforderten wissenschaftlichen Universalsprache und zu einem Instrument für den Aufbau der 'absoluten Algebra', einer allgemeinen Theorie der Verknüpfung," that is, served as both a *characteristica universalis* and as a *calculus ratiocinator*, the former serving as the theoretical foundation for the latter. Quoting Peirce from a manuscript of 1906 in which he offered a summary of his thinking on logic as a calculus and logic as a language, Hawkins leads us to conclude that Peirce would not be content to consider satisfactory a logic which was merely a calculus, but not also a language, or pasigraphy also; Peirce, comparing his dual conception of logic both as a calculus and as a language with the conceptions which he understood to be those of Peano, on the one hand, of Russell on the other, writes ([90, MS 499:1–5]; as quoted by [32, 120]):

> The majority of those writers who place a high value upon symbolic logic treat it as if its value consisted in its mathematical power as a calculus. In my (1901) article on the subject in Baldwin's Dictionary I have given my reasons for thinking... if it had to be so appraised, it could not be rated as much higher than puerile. Peano's system is no calculus; it is nothing but a pasigraphy; and while it is undoubtedly useful, if the user of it exercises a discreet [*sic*] freedom in introducing additional signs, few systems have been so wildly overrated as I intend to show when the second volume of Rusell and Whitehead's Principles of Mathematics appears.[20] ... As to the three modifications of Boole's algebra which are much in use, I invented these myself, – though I was anticipated [by De Morgan] as regards to one of them, – and my dated memoranda show ... my aim was ... to make the algebras as analytic of reasonings as possible and thus to make them capable of every kind of deductive reasoning....It ought, therefore, to have been obvious in advance that an algebra such as I am aiming to construct could not have any particular merit [in reducing the number of processes, and in specializing the symbols] as a calculus.

Taking Peirce's words here at face value, we are led to conclude that, unlike those Booleans who were satisfied to devise calculi which were not also languages, Peirce, toward the conclusion of his life, if not much earlier, required the development of a logic which was both a calculus (or critic, "which are much in use, I invented these myself") and a language (or semiotic), and indeed in which the semiotic aspect was predominant and foundational, while considering the idea of logic as a language of paramount utility and importance.

It is with these considerations that we can comprehend that the duality between the notions of logic as a calculus and logic as a language is tantamount to Peirce's narrow conception of logic as critic, on the one hand, and to his broad conception of logic as a general theory of signs or semiotics, on the other. It is on this basis that Volker Peckhaus has concluded [71, 174–175] that, in fact, Peirce's algebra and logic of relatives

> "wurde zum pasigraphischen Schlüssel zur Schaffung einer schon in den frühen zeichentheoretischen Schriften programmatisch geforderten wissenschaftlichen Universalsprache und

[20] Peirce tended to conflate Russell and Whitehead even with respect to Russell's *Principles of Mathematics*, even prior to the appearance of the co-authored *Principia Mathematica* [139], presumably because of their earlier joint work "On Cardinal Numbers" [139] in the *American Journal of Mathematics*, to which Russell contributed the section on, a work with which Peirce was already familiar.

> zu einem Instrument für den Aufbau der 'absoluten Algebra', einer allgemeinen Theorie der Verknüpfung"

that is, served both as a *characteristica universalis* and as a *calculus ratiocinator*, the former serving as the theoretical foundation for the latter.

Acknowledgement The author is grateful to Nathan Houser for suggestions that considerably improved the text and helped clarify a number of points.

References

1. Albertus de Saxonia: Logica Albertucii: Perutilis logica excellentissimi sacre theologie professoris magistri Alberti de Saxonia ordinis Eremitarum divi Augustini: per reverendum sacre pagine doctorem magistrum Petrum Aurelium Sanutum Venetum eiusdem ordini professum: quaque diligentissime castigata: nuperrimeque impressa. Venetiis: Ere ac Sollertia Heredum Octavianus Scotus (1522)
2. Anellis, I.H.: Peirce rustled, Russell pierced: How Charles Pierce and Bertrand Russell viewed each other's work in logic, and an assessment of Russell's accuracy and role in the historiography of logic. Mod. Log. **5**, 270–328 (1995). http://www.cspeirce.com/menu/library/aboutcsp_anellis/csp&br.htm
3. Anellis, I.H.: Some views of Russell and Russell's logic by his contemporaries. Rev. Mod. Log. **10**(1/2), 67–97; electronic version: Some views of Russell and Russell's logic by his contemporaries, with Particular Reference to Peirce http://www.cspeirce.com/menu/library/aboutcsp/anellis/views.pdf (2004–05)
4. Anellis, I.H.: How Peircean was the 'Fregean' revolution in logic? Log. Issled. **18**, 239–272; preprint: http://vfc.org.ru/eng/events/conferences/smirnov2011/members/; http://vfc.org.ru/rus/events/conferences/smirnov2011/members/; final version: http://www.cspeirce.com/menu/library/aboutcsp/anellis/csp-frege-revolu.pdf; http://arxiv.org/abs/1201.0353 (2012)
5. Anellis, I.H.: La Logique et la théorie de la notation (sémiotique) de Peirce. In: Béziau, J.-Y. (ed.) Pointure du symbole, forthcoming. Editions Petra, Paris (2013)
6. Badesa, C.: The Birth of Model Theory: Löwenheim's Theorem in the Frame of the Theory of Relatives (Maudsley, M. trans.). Princeton University Press, Princeton, Oxford (2004)
7. Bíard, J.: Matiére et forme dans la théorie buridanienne des consequences. Arch. d'hist. doctrin. litt. Moyen Age **LXI**, 151–168 (1989)
8. Blumenthal, O.: Lebensgeschichte. In: David Hilbert, Gesammelte Abhandlungen, vol. III, pp. 388–429. Springer, Berlin (1932)
9. Boole, G.: The Mathematical Analysis of Logic. Macmillan, Barclay, & Macmillan, Cambridge; George Bell, London (1847)
10. Boole, G.: An Investigation of the Laws of Thought, on Which are Founded the Mathematical Theories of Logic and Probabilities. Walton & Maberly, London (1854), reprinted: Dover, New York (1951)
11. Brady, G.: From Peirce to Skolem: A Neglected Chapter in the History of Logic. North-Holland/Elsevier Science, Amsterdam/New York (2000)
12. Bryant, S.W.: On the nature and functions of a complete symbolic language, Mind **13**, 188–208 (1888)
13. Buridanus, J.: Summulae de dialectica (Klima, G. translator and editor). Yale University Press, New Haven (2001)
14. Bynum, T.W.: On the Life and Work of Gottlob Frege. In: Frege [1972], pp. 1–54 (1972)
15. Clark, G.: New light on Peirce's iconic notation for the sixteen binary connectives. In: Roberts, D.D., Van Evra, J. (eds.) Studies in the Logic of Charles Sanders Peirce Nathan Houser, pp. 304–333. Indiana University Press, Bloomington, Indianapolis (1987)
16. Couturat, L.: Letter to Bertrand Russell, 27 January 1901. Russell Archives: C35 (1901)

17. De Morgan, A.: On the syllogism: I. On the structure of the syllogism. Trans. Camb. Philos. Soc. **8**, 379–408 (1846), reprinted in De Morgan [1966], pp. 1–21
18. De Morgan, A. in: Heath, P. (ed.) On the Syllogism and Other Logical Writings. Yale University Press, New Haven (1966)
19. Ebbesen, S.: Is logic theoretical or practical knowledge? In: Bíard, J. (ed.) Itinéraires d'Albert de Saxe Paris Vienne au XIVe Siécle, pp. 267–276. Librairie Philosophique J. Vrin, Paris (1991)
20. Esquisable, O.M.: Representing and abstracting: An analysis of Leibniz's concept of symbolic knowledge. In: Cassanave, A.L. (ed.) Symbolic Knowledge from Leibniz to Husserl, pp. 1–49. College Publications, London (2012)
21. Frege, G.: Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Verlag von Louis Nebert, Halle (1879). English translation by Stefan Bauer-Mengelberg in van Heijenoort [1967a], pp. 1–82; English translation by Terrell Ward Bynum: [Frege 1972], pp. 101–203
22. Frege, G.: Booles rechnende Logik und die Begriffsschrift. In: Hermes, H., Kambartel, F., Kaulbach, F.C.S.J. (eds.) Gottlob Frege, Nachgelassene Schriften, pp. 9–52 (1880/81). Felix Meiner Verlag, Hamburg (1969), 2nd enlarged edn. (1983); English translation: Boole's logical calculus and the concept-script. In: Long, P., White, R. (trans.) Gottlob Frege, Posthumous Writings, pp. 9–46. University of Chicago Press, Chicago (1979)
23. Frege, G.: Booles logische Formelsprache und die Begriffsschrift. In: Hermes, H., Kambartel, F., Kaulbach, F.C.S.J. (eds.) Gottlob Frege, Nachgelassene Schriften, pp. 53–59 (1882). Felix Meiner Verlag, Hamburg (1969), 2nd enlarged edn. (1983); English translation: Boole's logical formula-language and my concept-script. In: Long, P., White, R. (trans.) Gottlob Frege, Posthumous Writings, pp. 47–52. University of Chicago Press, Chicago (1979)
24. Frege, G.: Über den Zweck der Begriffsschrift. Jena. Z. Naturwiss. **16**, Suppl.-Heft II, 1–10 (1883)
25. Frege, G.: Über Sinn und Bedeutung. Z. Philos. philos. Krit. (N.F.) **100**, 25–50 (1892)
26. Frege, G.: Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik. Arch. Syst. Philos. **1**, 433–456 (1895)
27. Frege, G.: Über die Begriffsschrift des Herrn Peano und meine einige. Verhandlungen der Königliche Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Math.-Phys. Kl. **48**, 362–368 (1896)
28. Frege, G.: In: Bynum, T.W. (ed.) Conceptual Notation and Related Articles. Clarendon Press, Oxford
29. Freudenthal, H.: David Hilbert. In: Gillispie, C.C. (ed.) Dictionary of Scientific Biography, vol. 6, pp. 388–395. Scribner, New York (1973)
30. Gödel, K.: Über die formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme, I. Monatsh. Math. Phys. **38**, 173–198 (1931); English translation in van Heijenoort [1967a], pp. 596–617 (1967)
31. Grattan-Guinness, I.: Wiener on the logics of Russell and Schröder: An account of his doctoral thesis, and of his discussion of it with Russell. Ann. Sci **32**, 103–132 (1975)
32. Hawkins, B.S., Jr.: Peirce and Russell: The history of a neglected 'Controversy'. In: Houser, N., Roberts, D.D., Van Evra, J. (eds.) Studies in the Logic of Charles Sanders Peirce, pp. 111–146. Indiana University Press, Bloomington, Indianapolis (1997)
33. Herbrand, J.: Recherches sur la théorie des démonstration. Ph.D. thesis, University of Paris (1930); reprinted in Prace Towarzystwa Naukowego Warszawskiego, Wydział III, no. 33; English translation of chapter 5 in van Heijenoort [1967a], pp. 525–581 (1967)
34. Hilbert, D.: Axiomatisches Denken. Math. Ann. **78**, 405–415 (1918)
35. Hilbert, D.: Neubegründung der Mathematik (Erste Mitteilung). Abh. math. Semin. hambg. Univ. **1**, 157–77 (1922)
36. Hilbert, D.: Die logischen Grundlagen der Mathematik. Math. Ann. **88**, 151–165 (1923)
37. Hilbert, D.: Über das Unendliche. Math. Ann. **95**, 161–190 (1926); English translation by Stefan Bauer-Mengelberg in van Heijenoort [1967a], 367–392 (1967)
38. Hilbert, D.: Die Grundlagen der Mathematik. Abh. math. Semin. hambg. Univ. **6**, 91–122 (1926); English translation by Stefan Bauer-Mengelberg and Dagfinn Føllesdal in van Heijenoort [1967a], 464–479 (1967)
39. Hilbert, D., Ackermann, W.: Grundzüge der theoretischen Logik. Springer, Berlin (1928)

40. Hintikka, J.: On the development of the model-theoretic viewpoint in logical theory. Synthése **77**, 1–36 (1988)
41. Houser, N.: Peirce's early work on the algebra of logic: Remarks on Zeman's account. Transact. Charles S. Peirce Soc. **23**, 425–440 (1987)
42. Houser, N. (ed.): The Schröder-Peirce Correspondence. Mod. Log. **1**, 206–236. http://projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.rml/1204834616 (1990/91)
43. von Humboldt, W.: Ueber die Buchstabenschrift und ihren Zusammenhang mit dem Sprachbau. Gelesen in der Akademie der Wissenschaften am 20. Mai 1824, in Druckerei der Königlichen Akademie der Wissenschaften (Hsg.), Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin. Aus dem Jahre 1824. Historisch-philologische Klasse, pp. 161–188. F. Dümmler, Berlin (1826)
44. Isnenghi, A.: C. S. Peirce's Semiotic and Modistae's Grammatica Speculativa. Cognitio-Estudos: Rev. Electrón. Filos. **5**, 148–151 (2008). http://www.pucsp.br/pos/filosofia/Pragmatismo/cognitioestu-dos/cog_estudos_v5n2/cog_est_52_Isnenghi.pdf
45. Kennedy, H.C.: What Russell learned from Peano. Notre Dame J. Form. Log. **14**, 367–372. (1973)
46. Korte, T.: Frege's Begriffsschrift as a lingua characteristica. Synthése **174**, 283–294 (2010)
47. Květ, F.B.: Leibnitz'ens Logik. Nach den Quellen dargestellt. F. Tempsky, Prague (1857)
48. Ladd-Franklin, C.: Review of Vorlesungen über die Algebra der Logik (Exakte Logik). By Dr. Ernst Schröder. Vol. I. Mind (n.s.) **1**, 126–132 (1892)
49. Lasswitz, K.: Rezension von G. Frege, Begriffsschrift. Jena. Literaturztg. **16**, 248–249 (1879); English translation by Terrell Ward Bynum: [Frege 1972], pp. 210–212 (1972)
50. Legris, J.: 'In Symbolen denken': Symbolic Knowledge in Frege's Conceptual Notation. In: Bernhard, P., Peckhaus, V. (eds.) Exactes Denken in Kontexte, Festschrift für Christian Thiel mit einem unveröffentlichen Brief Gottlob Freges, pp. 133–153. Paderborn, Mentis (2008)
51. Legris, J.: Between calculus and semantic analysis: Symbolic knowledge. In: Lassalle Cassanave, A. (ed.) Symbolic Knowledge from Leibniz to Husserl, pp. 79–113. College Publications, London (2012)
52. Leibniz, G.W.: Berlin-Brandenburgische Akademie der Wissenschaften und der Akademie der Wissenschaften in Göttingen (ed.) Sämtliche Schriften und Briefe, Riehe. 6. Philosophische Schriften, Bd. 4: 1677–Juni 1690. 4 Th. Akademie-Verlag, Berlin (1999)
53. Leibniz, G.W.: Die Grundlagen des logischen Kalküls, herausgegeben, übersetzt und mit einem Kommentar versehen von Franz Schupp, unter der Mitarbeit von Stephanie Weber: lateinisch–deutsch. Felix Meiner Verlag, Hamburg (2000)
54. Löwenheim, L.: Über Möglichkeiten im Relativkalkül. Math. Ann. **76**, 447–470 (1915); English translation by Stefan Bauer-MengelBerg in van Heijenoort [1967a], 228–251 (1967)
55. De Maimieu, J.: Pasigraphie... ou... premiers élémens du nouvel art-science d'écrire d'imprimer en un langue de maniére a etre lu et entendu dans toute autre langue sans traduction, 2nd edn.. Bureau de la Pasigraphie, Paris (1799)
56. De Maimieu, J.: Pasigraphie... oder... Anfangsgründe der neuen Kunst-Wissenschaft in einer Sprache alles so zu schreiben und zu drucken, dass es in jeder andern ohne übersetzung gelesen und verstehen werden kann erfunden und verfasset. Bureau de la Pasigraphie, Paris (1797)
57. Maulfelt, T.: In: Bursill-Hall, G.L. (ed. & translator) Thomas of Erfurt: Grammatica Speculativa. Longmans, Green, London (1972)
58. Merrill, D.D.: Augustus De Morgan and the Logic of Relations. Kluwer Academic, Dordrecht (1990)
59. Oliver, J.E.: Elementary notes. I. General and logico-mathematical notation. Ann. Math. **4**, 186–193 (1888)
60. Patzig, G.: Leibniz, Frege und die sogenannte 'lingua characteristica universalis'. Stud Leibnit. **3**, 103–112 (1969)
61. Peano, G.: Arithmetices principia, nova methodo exposita. Bocca, Torino (1889), abridged English translations in van Heijenoort [1967a], pp. 83–97 (1967); [Peano 1973], 101–134 (1973)
62. Peano, G.: Principii di logica matematica, Riv. Mat. **1**, 1–10 (1891); Reprinted: [Peano 1957–59], II, 92–101 (1957–59); English translation by Hubert C. Kennedy: [Peano 1973], 153–161 (1973)

63. Peano, G.: Review of [Schröder 1890–1905, I]. Riv. Mat. **1**, 164–170 (1899)
64. Peano, G.: Notations de logique mathématique (Introduction auFormulaire de mathématiques). Tipografia Guadagnini, Torino (1094)
65. Peano, G.: Formulaire de mathématiques, tm. 1. Bocca, Torino (1895)
66. Peano, G.: Formulaire de mathématiques, tm. 2. Sec. 1. Bocca, Torino (1897)
67. Peano, G.: Formulaire de mathématiques, tm. 2. Sec. 2. Bocca, Torino (1898)
68. Peano, G.: Review of [118]. Riv. Mat. **6**, 95–101 (1898)
69. Peano, G.: In: Cassina, U. (ed.) Opera scelte, 3 vols. Edizioni Cremonese, Rome (1957–59)
70. Peano, G.: In: Kennedy, H.C. (ed.) Selected Works of Giuseppe Peano. University of Toronto Press, Toronto, Buffalo, NY (1973)
71. Peckhaus, V.: Ernst Schröder und die 'pasigraphischen Systeme' von Peano und Peirce. Mod. Log. **1**, 174–205 (1990/91)
72. Peckhaus, V.: Ernst Schröder und der Logizismus. In: Stelzner, W. (ed.) Philosophie und Logik. Frege-Kolloquien Jena 1989/1991, pp. 108–119. Walter de Gruyter, Berlin (1993); reprinted: Beaney, M., Reck, E.H. (eds.) Frege: Critical Assessments of Leading Philosophers, pp. 176–190. Routledge, London, New York (2005)
73. Peckhaus, V.: Wozu Algebra der Logik? Ernst Schröders Suche nach einer universalen Theorie der Verknüpfungen. Mod. Log. **4**, 357–381 (1994)
74. Peckhaus, V.: The Axiomatic Method and Ernst Schröder's Algebraic Approach to Logic. Philos. Sci. **1**, 1–15 (1996)
75. Peckhaus, V.: Calculus ratiocinator versus Characteristica Universalis?: The two traditions in logic. Revis. Hist. Philos. Log. **25**, 3–14 (2004); reprinted: Beaney, M., Reck, H. (eds.) Gottlob Frege. Critical Assessments of Leading Philosophers, vol. 1: Frege's Philosophy in Context, pp. 176–190. Routledge, London, New York (2005)
76. 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. 567–609. Elsevier, North-Holland, Amsterdam (2004)
77. Peckhaus, V.: Frege, Friedrich Ludwig Gottlob, Complete Dictionary of Scientific Biography. 2008. Encyclopedia.com. (January 9, 2013). http://www.encyclopedia.com/doc/1G2-2830905684.html (2013)
78. Peirce, B.: Linear associative algebra (with notes and addenda by Charles S. Peirce). Am. J. Math. **4**, 97–229 (1881)
79. Peirce, C.S.: Description of a notation for the logic of relatives, resulting from an amplification of the conceptions of Boole's calculus of logic. Mem. Am. Acad. **9**, 317–378 (1870); reprinted: [Peirce 1933a], 3.45–149; [Peirce 1984], 359–429 (1984)
80. Peirce, C.S.: Of Logic as the Study of Signs, MS 221, Robin catalog # 380, March 14 (1873); published: [Peirce 1986], 82–84 (1986)
81. Peirce, C.S.: On the Algebra of Logic. Am. J. Math. **3**, 15–57 (1880); reprinted: [Peirce 1933b], 104–157 (1933); [Peirce 1989], 163–209 (1989)
82. Peirce, C.S.: On the logic of number. Am. J. Math. **4**, 85–95 (1881); reprinted: [Peirce 1989], 299–309 (1989)
83. Peirce, C.S.: The Logic of Relatives. In: Peirce, C.S. (ed.) Studies in Logic. Little, Brown & Co., Boston (1883); reprinted: Johns Hopkins University Press, Baltimore, pp. 187–203 (1983); reprinted: [Peirce 1932], pp. 195–210 (1932) and [Peirce 1989], pp. 453–466 (1989)
84. Peirce, C.S.: On the algebra of logic: A contribution to the philosophy of notation. Am. J. Math. **7**, 180–202 (1885); reprinted: [Peirce 1933a], pp. 27–98 (1933); [Peirce 1989], pp. 359–429 (1989)
85. Peirce, C.S.: The critic of arguments. Open Court **6**, 3391–3394 (1892); reprinted: [Peirce 1933a], pp. 3.404–424 (1933)
86. Peirce, C.S.: The Simplest Mathematics; January 1902 (Chapter III. The Simplest Mathematics (Logic III), MS 431; pp. 2–200 (p. 199 missing), including long alternative or rejected efforts (1902); see [Peirce 1933b], pp. 4:260–262 (1933)
87. Peirce, C.S.: C. S. Peirce's Lowell Lectures for 1903. Lecture 4, MS 467, 2 notebooks, pp. 1–96 (1903)

88. Peirce, C.S.: Prolegomena to an apology for pragmaticism. The Monist **16**, 492–546 (1906); reprinted: [Peirce 1933b], pp. 4.530–572 (1933)
89. Peirce, C.S.: Paper on the Existential Graphs for presentation to the National Academy of Science (April 16–18, 1906, Washington D.C.), MS 490 (1906); published in part in Peirce [1933b], pp. 4.573–584 (1933)
90. Peirce, C.S.: Phaneroscopy, MS 298 (1906)
91. Peirce, C.S.: On the System of Existential Graphs Considered as an Instrument for the Investigation of Logic A. MS., notebook (Harvard Cooperative), MS 499, n.d. (1906)
92. Peirce, C.S.: In: Hartshorne, C., Weiss, P. (eds.) Collected Papers of Charles Sanders Peirce, vol. II, Elements of Logic. Harvard University Press, Cambridge, MA (1932); 2nd edn. (1960)
93. Peirce, C.S.: In: Hartshorne, C., Weiss, P. (eds.) Collected Papers of Charles Sanders Peirce, vol. III, Exact Logic (Published Papers). Harvard University Press, Cambridge, MA (1933); 2nd edn. (1960)
94. Peirce, C.S.: In: Hartshorne, C., Weiss, P. (eds.) Collected Papers of Charles Sanders Peirce, vol. IV, The Simplest Mathematics. Harvard University Press, Cambridge, MA (1933); 2nd edn. (1960)
95. Peirce, C.S.: In: Hartshorne, C., Weiss, P. (eds.) Collected Papers of Charles Sanders Peirce, vol. V, Pragmatism and Pragmaticism. Harvard University Press, Cambridge, MA (1934); 2nd edn. (1960)
96. Peirce, C.S.: In: Hartshorne, C., Weiss, P. (eds.) Collected Papers of Charles Sanders Peirce, vol. VI. Harvard University Press, Cambridge, MA (1935); 2nd edn. (1960)
97. Peirce, C.S.: In: Buchler, J. (ed.) Philosophical Writings. Dover Publications, New York (1955)
98. Peirce, C.S.: In: Wiener, P.P. (ed.) Charles S. Peirce: Selected Writings (Values in a Universe of Chance). Dover Publications, New York (1958)
99. Peirce, C.S.: In: Fisch, M.H. (ed.) Writings of Charles S. Peirce: A Chronological Edition, vol. 1: 1857–1866. ndiana University Press, Bloomington (1982)
100. Peirce, C.S.: In: Moore, E.C. (ed.) Writings of Charles S. Peirce: A Chronological Edition, vol. 2: 1867–1871. Indiana University Press, Bloomington (1984)
101. Peirce, C.S.: In: Eisele, C. (ed.) Historical Perspectives on Peirce's Logic of Science. A History of Science. 2 vols. Mouton Publishers, Berlin (1985)
102. Peirce, C.S.: In: Kloesel, C.J.W. (ed.) Writings of Charles S. Peirce: A Chronological Edition, vol. 3: 1872–1878. Indiana University Press, Bloomington (1986)
103. Peirce, C.S.: In: Kloesel, C.J.W. (ed.) Writings of Charles S. Peirce: A Chronological Edition, vol. 4: 1879–1884. Indiana University Press, Bloomington (1989)
104. Peirce, C.S.: In: (Peirce Edition Project (eds.) The Essential Peirce. Selected Philosophical Writings, vol. 2 (1893–1913). Indiana University Press, Bloomington (1998)
105. Peirce, C.S.: The Principles of Logical Graphics, MS 493; small red leather notebook, n.p., n.d.
106. Peirce, C.S., Ladd-Franklin, C.: Logic. In: Baldwin, J.M. (ed.) Dictionary of Philosophy and Psychology: Including Many of the Principal Conceptions of Ethics, Logic, Aesthetics, Philosophy of Religion, Mental Pathology, Anthropology, Biology, Neurology, Physiology, Economics, Political and Social Philosophy, Philology, Physical Science, and Education and Giving a Terminology in English, French, German and Italian, vol. II, pp. 21–23. Macmillan, New York, London (1902)
107. Pietarinen, A.-V.: Cultivating habits of reason: Peirce and the logica untens versus logica docens distinction. Hist. Philos. Q. **22**, 357–372 (2005). http://www.helsinki.fi/ pietarin/publications/Habits%20of%20Reason-Pietarinen.pdf
108. Pietarinen, A.-V.: Significs and the Origins of Analytic Philosophy, preprint, 33 pp (2008)
109. Pietarinen, A.-V.: Significs and the origins of analytic philosophy. J. Hist. Ideas **70**, 467–490 (2009)
110. Pietarinen, A.-V.: Which philosophy of mathematics is Pragmaticism? In: Moore, M.E. (ed.) New Essays on the Mathematical Philosophy of C.S. Peirce, pp. 59–79. Open Court Publishing, Chicago (2010)
111. Russell, B. Sur la logique des relations avec des applications à la théorie des séries. Rev. Math./Riv. Mat. **7**, 115–148 (1901)
112. Russell, B. Principles of Mathematics. Cambridge University Press, London (1903); 2nd edn. Allen & Unwin, London (1937); 2nd American edn. W.W. Norton, New York (1938)
113. Russell, B. Letter to Lady Victoria Welby, December 27, 1904, 3 pp. ms. (1904)
114. Schröder, E. Der Operationskreis des Logikkalkuls. B. G. Teubner, Leipzig (1877)

115. Schröder, E. Rezension zu Freges Begriffsschrift. Z. Math. Phys., Historisch-literaturische Abteilung **25**, 81–93 (1880); English translation by Terrell Ward Bynum in Frege [1972], pp. 218–232 (1972); English translation by Victor H. Dudman: Translation of Schröder's Review of Begriffsschrift. S. J. Philos. **7**, 139–150 (1969)
116. Schröder, E. Vorlesungen über die Algebra der Logik (Exakte Logik), 3 Bände. B. G. Teubner, Leipzig (1890–1905)
117. Schröder, E. Signs and Symbols. Open Court **6**, 3431–3434, 3441–3444, 3663–3666 (1892)
118. Schröder, E. Über Pasigraphie, ihren gegenwärtigen Stand und die pasigraphische Bewegung in Italien. In: Rudio, F. (ed.) Verhandlungen des Ersten Internationalen Mathematiker-Kongressses in Zürich von 9. bis 11. August 1897. B. G. Teubner, Leipzig (1898)
119. Schröder, E. On pasigraphy: Its present state and the pasigraphic movement in Italy. The Monist **9**, 44–62, 320 (1898). http://www.jfsowa.com/logic/es99.pdf
120. Shalak, V.I.: Two approaches to the construction of logic. Log. Issled. **17**, 269–280 (2011)
121. Skolem, T.: Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen. Videnkapsselsk. Skr. **1** (Mathematisk-naturvidenskabelig klasse), no. 4, 1–36 (1920); English translation of Section 1 by Stefan Bauer-Mengelberg in van Heijenoort [1967a], pp. 252–263 (1967)
122. Sluga, H.-D.: Frege Against the Booleans. Notre Dame J. Form. Log. **28**, 80–98 (1987)
123. Sowa, J.F.: Commentary onCharles Sanders Peirce, Existential Graphs, MS 514 (1909). http://www.jfsowa.com/peirce/ms514.htm (2010)
124. Thiel, C.: Ernst Schröder und die Verteilung der Quantoren. Arbeitsberichte, Institut für Philosphie der Universität Erlangen-Nürnberg, no. 16 (März), 20–37 (1988); English translation: [Thiel 1990/91] (1990/91)
125. Thiel, C.: Ernst Schröder and the distribution of quantifiers. Mod. Log. **1**, 160–173 (1990/91)
126. Zum Verhältnis von Syntax und Semantik bei Frege. In: Stelzner, W. (ed.) Philosophie und Logik. Frege-Kolloquien Jena 1989/1991, pp. 3–15. Walter de Gruyter), Berlin (1993)
127. Trendelenburg, F.A.: Ueber Leibnizens Entwurf einer allgemeinen Charakteristik. In: Trendelenburg, A. (ed.) Historische Beiträge zur Philosophie, vol. III, pp. 1–31. Bethge, Berlin (1867)
128. Van Heijenoort, J. (ed.): From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931. Harvard University Press, Cambridge, MA (1967)
129. Van Heijenoort, J. Logic as calculus and logic as language. Synthése **17**, 324–330 (1967); reprinted: Cohen, R.S., Wartofsky, M.W. (eds.) Boston Studies in the Philosophy of Science 3. Proceedings of the Boston Colloquium Philosophy of Science 1964/1965, pp. 440–446. D. Reidel, Dordrecht; Sluga, H. (ed.) General Assessments and Historical Accounts of Frege's Philosophy, vol. 1: The Philosophy of Frege, pp. 72–79. Garland Publishing, New York (1993); and [van Heijenoort 1985b], pp. 11–16 (1985)
130. Van Heijenoort, J. Set-theoretic semantics. In: Gandy, R.O., Holland, J.M.E. (eds.) Logic Colloquium '76, Oxford, 1976, pp. 183–190. North-Holland, Amsterdam (1977); reprinted: [van Heijenoort 1985b], pp. 43–53 (1985)
131. Van Heijenoort, J. Absolutism and Relativism in Logic (1979); published: [van Heijenoort 1985b], pp. 75–83 (1985)
132. Van Heijenoort, J. Selected Essays. Bibliopolis, Naples (1985)
133. Van Heijenoort, J. Systéme et métasystéme chez Russell. In: The Paris Logic Group (eds.) Logic Colloquium '85, pp. 111–122. North-Holland, Amsterdam, London, New York (1987)
134. Van Heijenoort, J. Historical development of modern logic. Mod. Log. **2**, 242–255 (1992) [prepared by Irving H. Anellis from a previously unpublished typescript of 1974]; revised and corrected reprint: Log. Univers. **6**, 327–337 (2012)
135. Venn, J.: Review of G. Frege, Begriffsschrift. Mind (o.s.) **5**, 297 (1880); reprinted: [Frege 1972], pp. 234–235 (1972)
136. Venn, J.: On the various notations adopted for expressing the common propositions of logic. Proc. Camb. Philos. Soc. **4** (December 1880), pp. 36–47; reprinted as Historical Notes in Venn [1894], pp. 477–504 (1894)

137. Venn, J.: Symbolic Logic, 2nd edn. Macmillan & Co., London (1894); reprinted: Chelsea Publishing, Bronx, NY (1971)
138. Vile, A.: Logic as semiotics – Peirce's existential graphs and possibilities for making more sense in learning logic. Philos. Math. Educ. J. **10**, 17–20 (1997). http://people.exeter.ac.uk/PErnest/pome10/art5.htm. Accessed 12/9/2009 (2009)
139. Whitehead, A.N., Russell, B.: Principia Mathematica, 3 vols. Cambridge University Press, Cambridge (1910–1913)
140. Whitehead, A.N., Russell, B.: Principia Mathematica, 3 vols, 2nd edn. Cambridge University Press, Cambridge (1925–1927)
141. Wiener, N.: A Comparison Between the Treatment of the Algebra of Relatives by Schroeder and that by Whitehead and Russell. Ph.D. thesis, Harvard University (Harvard transcript and MIT transcript) (2013). Partial publication as Appendix 8 of the introduction and last chapter in Brady [2000], pp. 429–444 (2000)
142. Wittgenstein, L.: Tractatus logico-philosophicaus/Logisch-philosophische Abhandlung. Routledge & Kegan Paul, London (1922)
143. Wittgenstein, L.: In: Rush R. (ed.) Philosophische Grammatik. Suhrkamp, Frankfurt (1973)
144. Wu, J., Sen.: The problem of existential import in Dewey's theory of propositions. Ph.D. thesis, Southern Illinois University (1967)
145. Wu, J., Sen.: The problem of existential import (from George Boole to P.F. Strawson). Notre Dame J. Form. Log. **10**, 415–424 (1969)
146. Zellweger, S.: Untapped potential in Peirce's iconic notation for the sixteen binary connectives. In: Houser, N., Roberts, D.D., Van Evra, J. (eds.) Studies in the Logic of Charles Sanders Peirce, pp. 334–386. Indiana University Press, Bloomington, Indianapolis (1997)

I. H. Anellis (✉)
Peirce Edition Project, Indiana University, Indianapolis, USA
Purdue University at Indianapolis, Indianapolis, USA

The Meaning(s) of "Is": Normative vs. Naturalistic Views of Language

Ignacio Angelelli

Abstract One of the founders of modern logic, G. Frege, has insisted on the variety of meanings of the little word "is." He explicitly distinguished four such meanings (sheer predication or subsumption, identity, assertion, and existence); a fifth meaning (subordination) follows from Frege's new theory of predication. It is part of the Fregean doctrine that special symbols corresponding to different meanings of "is" are to be used. Such distinctions have been strongly challenged by J. Hintikka, in a twofold way: theoretically and historiographically. Neither challenge is regarded as successful. Behind the conflict on "is" two opposite conceptions of language may be perceived: language as culture versus language as nature ("natural language").

Keywords Frege · Hintikka · Meanings of "is" · The logic of being · Identity · Subordination · Subsumption · Existence · Assertion

Mathematics Subject Classification (2010) Primary 03B65 · Secondary 03A05

1 Introduction

In Sect. 2, the distinction of five meanings of "is" that follows from Frege's work is presented; in Sect. 3 its rejection by J. Hintikka is examined. In Sect. 4, two opposite conceptions of language are considered as possible explanation of the conflict.[1]

2 The Five Meanings of "Is" According to Frege

From Frege's work it follows that five meanings of "is" should be singled out: (1) sheer predication or subsumption (*Subsumtion*), (2) identity, (3) assertion, (4) existence, (5) subordination (*Unterordnung*). "Follows from Frege's work" as opposed to "explicitly given in Frege's work" rather applies to (5).

There is abundant textual support for the distinction between sheer predication and identity, which appears to be in fact the distinction Frege is most enthusiastic about: [14],

[1] My earlier publications on this topic include [2–6].

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_6

§57; [15] in [25], pp. 168–169; [24], letter to Linke; [16] in [25] p. 200; [22], pp. 255–256, 258–259; [17] in [25], p. 280. Frege thought of "is" as involving (at least) two uses, *Gebrauchsweisen*: as a mere copula, a "mere formal word of the enunciation," or as meaning identity. It must be observed that although not every predication statement involves identity,[2] in every identity statement predication is included. The distinction of the two uses is concealed (*verhüllt*) by ordinary language. Frege's normative approach to logic and language is at work here: the mere realization that there are two uses in ordinary German is not the end; it is necessary to adjust our notation, writing, for instance, "$A = B$" instead of "A is B."

With regard to the assertive "is" several Fregean texts are very clear: [19], p. 192; [20], pp. 201, 207, 211; [21], p. 214. The word "is" is often combined with an assertoric force (*die behauptende Kraft*). Frege observes that in logic books predicating is confused with judging (*So wird es auch in den Logiken, wie es scheint, das Prädizieren verquickt mit dem Urteilen* [21], p. 201). Frege's normative approach to language becomes active: it is necessary to adopt a special sign for assertion: "Es besteht das Bedürfnis, einen Gedanken auszudrücken, ohne ihn als wahr hinzustellen. In der Begriffsschrift habe ich ein eigenes Zeichen mit behauptender Kraft: den Urteilsstrich. In den mir bekannten Sprachen fehlt ein solches Zeichen, und die behauptende Kraft ist mit dem Indikativ in den Hauptsätzen fest verbunden" (It is necessary to express a thought without presenting it as true. In the ideography, I use a special sign to convey assertoric force: the judgment stroke. The languages known to me lack such a sign, and assertoric force is closely bound up with the indicative mood of the sentence that forms the main clause.) [21], p. 214.

The existence meaning of "is" is not at all Frege's favorite. One might conjecture that this is because the German (or English, etc.) uses of the existential "is" are unusual, contrary to Latin (*Deus est*). The reason is rather that Frege regards the existential use of "is" as uninformative, empty, leading in the end to a "deification" of the copula (*Vergötterung*, [18], p. 71, wrongly translated into "apotheosis" in the *Posthumous Writings* [23], p. 64).

The just listed four senses of "is" were known in the pre-Fregean history of logic, which does not mean that special notations were used. As Weidemann points out: "Aquinas is well aware of the difference between the 'is' of predication and the 'is' of identity" [36], p. 183. Frege's complaint that logic books mix up assertion and predication is understandable relative to what was in the air in the 19th century ("judgment" or "Urteil," properly meaning assertion but wrongly used as "proposition"); in the scholastic tradition, with which Frege was not familiar, this distinction was very familiar, continued into neoscholastic texts (for example [45], §35: the copula "is" has a "double function": "merely copulative" and "properly judicative"). The contrast between predication and existence was quite popular before Frege under the terminology *de tertio*, *de secundo adiacente*, respectively.

The fifth subordination "is," or more precisely "est," can be detected mainly in the Latin (or Greek) texts from the history of logic and philosophy, provided one looks at these texts

[2] This should not be understood as contradicting the general ontological insight according to which the ultimate truth maker of any predication is an identity (cf. [7], Sect. 9).

from the standpoint of Frege's revolutionary predication theory. Given two predicates or concepts: M and P, the sentences of the form "M est P" mean that any object that is M is also P; for example "homo est rationalis," "equus est animal" express that the concept *homo* is subordinated to the concept *rationalis*, and *equus* to *animal*. These sentences may also mean that P is part of the "essence" of M. Concepts P were referred to as marks (*Merkmalen*, *notae*) of the concept M. Pre-fregean authors, focusing on predication, the converse of *est*, said that a mark P is predicated of the concept M of which P is a mark: *rationalis praedicatur de homine*, *animalis praedicatur de equo*. This was the principal feature of the pre-Fregean predication theory (cf. [7]). This principal feature was rejected by Frege, for a very understandable reason: in his philosophy of number, he had seen that numbers "appear" as a part of our discourse about concepts. For example, number 12 occurs as part of our talking about the concept *apostle*. "Having 12 individuals" is something that we want to say of, predicate of the concept apostle; "having 12 individuals" is a *property* of the concept apostle, not a mark of it. Frege rightly realizes that the two cases are quite different. For one thing, in the case of marks we have transitivity (proclaimed by Kant as supreme rule of knowledge[3]), not so in the case of properties of the concept. Contrary to, for example, the scholastic Caietanus, who tolerates, and "lives with" the two senses of "predication" (*praedicari de praedicato contingit dupliciter*, to be predicated of a predicate happens in two ways, [7], Sect. 3), Frege takes the radical course of refusing to call "predication" the traditionally accepted, alleged predication of a mark of the concept of which it is a mark. This happens first in [14], § 53; a later text: "Man sollte mit Subjekt und Prädikat in der Logik aufräumen; oder man sollte diese Wörter auf die Beziehung des Fallen seines Gegenstandes unter einen Begriff (Subsumtion) einschränken. Die Beziehung der Unterordnung eines Begriffes unter einen Begriff ist so verschieden von jener, dass es nicht erlaubt ist, auch hierbei von Subjekt und Prädikat zu reden" (letter to Husserl Oct. 1906 in [24]).[4] Thus, when a fan of Frege reads Latin history of logic texts, she will not construe the "est" of "homo est rationalis" or "equus est animal" as subsumption, but as subordination: such is the fifth sense of "is" or rather "est" indirectly introduced by Frege, via his new predication theory. It is very hard to find examples of the German "ist" or of the English "is" in the subordination sense; a rare case occurs in English with sentences of the form of "man is P," for example, "man is rational." Normally, the Latin *est* sentences go into universal affirmative sentences, where the one or two occurrences of "is" (or "are") are of the subsumption type, or into sentences of the form "the M is P," where the new problem is to understand the nature of the strange object "the M" (in [15], p. 170, Frege suggests two ways of handling this issue).

[3] Aus dem Angeführten erkennet man, daß die erste und allgemeine Regel aller bejahenden Vernunftschlüsse sei: Ein Merkmal vom Merkmal ist ein Merkmal der Sache selbst (nota notae est etiam nota rei ipsius). [43], § 2, pp. 601–602. English translation: "From the aforesaid one sees that the first and universal rule of all affirmative consequences is: a mark of the mark is a mark of the thing itself." The last portion in the German text simply repeats the principle for the Latin "nota" instead of the German "Merkmal."

[4] One should get rid of subject and predicate in logic, or one should restrict these words to the relation of an object falling under a concept (subsumption). The relation of subordination of a concept under a concept is so different from it, that it is not admissible to talk, in its case, also of subject and predicate.

3 Hintikka's Twofold (Theoretical and Historiographical) Revolt Against Frege

This chapter is restricted to the consideration of J. Hintikka's criticism of Frege's distinction of meanings of "is." Such a criticism is only part of a general "revolt" against Frege [32].

In the particular issue of the semantics of "is" Hintikka's anti-Frege campaign is twofold: *theoretical* (no need to distinguish meanings of "is" in English) and *historical* (the distinction of meanings of "is" does not help in our study of texts from the history of philosophy, and should not be retrospectively applied to them.) The following brief examination of this double "revolt" is based on [26–40]. The literature on the meanings of "is" includes many other valuable and often quite challenging essays [9, 10, 12, 13, 42, 46, 47] not discussed in this chapter.

3.1 Theoretical Revolt

It is not easy to determine what is exactly the theoretical opposition to the distinctions of "is." There is on the one hand, a recognition, on the part of Hintikka, of semantical discrepancies among the occurrences of "is." He observes that the "is" of "Jack is a boy" is clearly predicative, while the "is" of "Jack is John Jr." is clearly one of identity (*Self-Profile*, [11], p. 36). On the other hand, however, these discrepancies are, as it were, minimized in the sense that the context suffices to resolve them, which entails, according to Hintikka, that no special notations are required (no special sign for identity, for example).

The main problem is that beyond the five Fregean meanings, reduced to the status of "discrepancies," there arises in Hintikka's writings a sixth component in the semantics of "is." Using the Greek "esti" instead of the English "is," Hintikka refers to "the basic meaning of esti" ([37], end of Sect. 3), "the basic semantical force of esti" ([37], p. 86), "the basic semantical meaning of esti" ([37], p. 87). The nature of this semantic nucleus of the "is" remains unexplained, and seems to be unexplainable.

The hints at understanding the "basic meaning" of "is" with the help of the old *analogia entis* only obscure the issue further, since then instead of "the basic meaning" we have phrases such as "analogical focal meaning theory of 'is'" ([38], p. xiv), "focal meaning" ([37], p. 108), "analogous term" [36], p. xiii).

In sum, it seems unavoidable to conclude that the "revolt" against Frege boils down to the rejection of special symbols to express the various meanings, senses, uses, forces... semantic discrepancies of "is." The revolt ends up being merely "symbolic," i.e. not a real revolt, and just limited to the prohibition of new symbols. Imagine that the first thesis of Wittgenstein's *Tractatus:* "Die Welt ist alles, was der Fall ist," is read both by Frege and by Hintikka, and that a discussion follows. Frege would point out that the first "ist" is of identity, while the second is predicative (copula, cf. his April 1920 letter to Wittgenstein [41]). If Frege recommended the rewriting of the thesis in the following form: "Die Welt = alles, was der Fall ist," Hintikka would disapprove – but only because (for him) the context suffices to determine that the "ist" between "Welt" and "alles" means identity.

3.2 The Historiographical Campaign

Hintikka's intention in editing, with S. Knuuttila, the impressive collective volume *The Logic of Being* [36] may be described as follows. The lack of a distinction of meanings of "is" in English or in the pre-Fregean tradition is not to be regarded as a defect (as a "logical howler" [36], p. x) but as a feature to be praised since it reflects a better semantics of natural language ([36], p. 82). Scholars interested in revealing the good things of ancient philosophy should no longer feel obligated to show that classical authors paved the way for the "Frege–Russell thesis"[5] – such a precursorship would be, in Hintikka's judgment, a "non-honor" ([36], p. 257).

Four important claims made by Hintikka's [37] on Aristotle are the following. (1) "Aristotle does not recognize the Frege-Russell ambiguity of 'is'"'(such is the title of Sect. 1), that however "the nonambiguity of esti does not preclude purely existential uses" (title of Sect. 2), or "purely identificatory" uses (Sect. 3, last paragraph). That they are uses rather than meanings merely indicates, in Hintikka's approach, that there is no need for special notations for each one of them: the context will tell. (2) Moreover, Hintikka admits in Aristotle a "simultaneous presence of several Frege meanings in esti" ([37], p. 85) and suggests that this should help to understand such texts as *Metaphysics Z 6*, where Aristotle affirms the identity of each thing with its essence. (3) While for those of us who have been corrupted by Frege, the solution of the famous Coriscus fallacy ("you know Coriscus, Coriscus is the man who is approaching, hence you know the man who is approaching [whereas in fact you do not know the man who is approaching]") is in terms of a failure of the substitutivity of identicals (cf. [1]), for Hintikka the Aristotelian solution consists in distinguishing two types of predication: a transitive and a nontransitive one ([37], p. 83). (4) Hintikka writes: "it is clear that there is no Frege–Russell type difference in meaning for Aristotle between the different occurrences of "is" in 'Socrates is a man' and 'a man is an animal' ([37], p. 84)."

The following are comments on points (2)–(4). With regard to (2), pace Hintikka, it is hard to see how by just letting "is" display its identity meaning, the traditional puzzles about the famous passage are going to be solved, given that such puzzles originate, precisely, from an alleged identity. With regard to (3), it is inconsistent to reject the distinction of meanings in "is" while allowing it in the case of "predicated of" which is, after all, the converse. With regard to (4), surely there is no difference between the two, or potentially three "is" occurring in the two sentences: all of them express subsumption. But the Latin (or Greek) sentences of the logico-philosophical tradition are "Socrates est homo, homo est animal," where the difference between the two "est" cannot be ignored: the former goes from an individual to a concept (subsumption), the latter from a concept to another concept (subordination).

Hintikka celebrates that there is not, in Kant, the Fregean distinction subsumption–subordination: "As for the fourth alleged sense of 'is' apud Frege and Russell, Kant's assimilation of it to other senses (especially to the 'is' of predication) is seen from his failure (or refusal) to distinguish the subsumption of one concept to another from the

[5] Hintikka often adds Russell to Frege in referring to the logicians who launched the thesis of the various meanings of "is."

application of a concept to a particular (in other words, this particular's falling under the concept)." [31], p. 258. Crucial in this passage is the parenthetical suggestion that Kant "refused" to distinguish subsumption and subordination, presupposing that he *knew* the distinction. Unfortunately, Hintikka does not elaborate on this point.

As for some of the other contributions to *The Logic of Being* [36] the following remarks can be made. Kahn distinguishes seven ways of understanding the Greek *esti*. One of these is the "veridical meaning," curiously the same as the assertive meaning emphasized by Frege but overlooked by Hintikka in the inventory of senses of "is." Kahn's paper, if anything, helps Frege rather than Hintikka's historiographical program, pace the editors' claim in the Introduction (p. ix): "Kahn has argued that Frege and Russell's thesis that verbs for being, such as 'esti', are multiply ambiguous is ill suited for the purpose of appreciating the actual conceptual assumptions of Greek thinkers." Mates' contribution, while highlighting the uniqueness of "the primitive is" (which is as mysterious as Hintikka's "basic meaning" of the little word) appears to acknowledge at least some of the Fregean different meanings under the guise of special axioms stipulated for the primitive "is"; less than clear what is Mates' help to the "frontal attack" on Frege. Dancy's contribution is helpful in an amazing way. Dancy displays a lethal reason to discourage the friends of the "Frege–Russell thesis" both from projecting it on the old Greek texts and from applying it to English: the little word "is" has no meaning at all. No stronger argument can be imagined against the Fregean thesis on the ambiguity of "is" – the little word is declared to be... meaningless, but then, at the same time, Hintikka's "basic meaning" of "is" vanishes too. The relevance of Knuuttila's paper to *The Logic of Being*'s main purpose is the observation that Aquinas being is analogous. It seems that this is part of the assumption that the analogy of being is a weapon against the Frege–Russell thesis, a wrong assumption in my view. In sum, the otherwise excellent papers included in [36] are not helpful at all in the historiographical part of the "revolt" against Frege's distinctions.

4 Language as Culture Versus Language as Nature

The controversy on "is" seems to hinge, in the final analysis, on whether it makes sense to talk of a "natural" language which, like other phenomena of nature, should be left alone or, on the contrary, language is to be regarded as a cultural product, a tool which can and should be improved by its users, that is, with regard to which one may proceed in a normative way.

It is obvious on which side stands Frege. As he writes to Husserl: *Die Logik soll Richterin sein über die Sprachen*, logic must be the judge of languages [24], p. 103. This normativism with regard to language became weaker or even went lost towards the second half of the 20th century, being replaced by a linguistic naturalism (often strangely combined with a huge amount of axiomatic set theory, the latter however not intended to "change nature" in the least but to discover its secrets). The normative view survived in individual cases (e.g. Lorenzen, cf. [44]).

If one adopts a normative approach to language, and one is interested in enhancing the clarity and precision of the linguistic tools, one will recommend, with Frege, the use of a special symbol, say "=," instead of "is" in sentences such as "Venus is the morning star" and one will find, with Weyl [48], §8, that the "confusion of the copula with existence and with equality" is "grotesque." The predicament of the naturalistic conception of language lies in failing to realize that ordinary (not natural) language is itself a product of many special notations introduced (normatively or carelessly) in past centuries. There is no "natural" language.[6]

It is interesting that in such a comprehensive volume [36] there seems to be just one single hint at the normative approach to language. According to Jacobi, Abelard thought that the role of the science of argumentative discourse was "to reflect on and explain language and yet not to shy away from correcting it and bringing to it a greater precision" (p. 156).

References

1. Angelelli, I.: Friends and opponents of the substitutivity of identicals in the history of logic. In: M. Schirn (ed.) Studien zu Frege, vol. II, pp. 141–166. Frommann Verlag, Stuttgart (1976).
2. Angelelli, I.: En torno a la "lógica del ser" de Hintikka. In: Proceedings of Congreso Internacional Extraordinario de Filosofía (Córdoba, Argentina, 1987), vol. 2, pp. 881–888. Univ. de Córdoba, Córdoba (1988)
3. Angelelli, I.: Lógica y lenguaje en la historia de la filosofía. Los sentidos del verbo 'ser' (Hintikka vs. Frege). In: Analesde la Academia de Ciencias de Buenos Aires, XXXVI (1) pp. 115–137. 2002. http://www.ciencias.org.ar/user/files/angelelli02.pdf, also at http://www.utexas.edu/cola/depts/philosophy/faculty/iaa4774
4. Angelelli, I.: Review of Hintikka, J.: A hundred years later: the rise and fall of Frege's influence in language theory. Synthese **59**(1), 27–49 (1984). Math. Rev. 86d:03007a
5. Angelelli, I.: Review of Günthner, F.: Comments on Hintikka's "A hundred years later: the rise and fall of Frege's influence in language theory". Math. Rev. 86d:03007b
6. Angelelli, I.: Review of Hintikka, J.: A second generation epistemic logic and its general significance. Math. Rev. 2005e:03030
7. Angelelli, I.: Predication theory: classical and modern. In Hochberg, H., Mulligan, K. (eds.) Relations and Predicates, pp. 55–80. Ontos Verlag, Frankfurt (2004)
8. Apel, K.O.: Die Idee der Sprache in der Tradition des Humanismus von Dante bis Vico. Archiv für Begriffsgeschichte, vol. 8. Bouvier Verlag, Bonn (1963)
9. Bacon, J.: The meaning of 'is'. In: Manuscrito, Campinas, Sao Paulo, 1988, pp. xi, 1
10. Bürsmeyer, K.:The verb and existence. The new Scholast. **40**, 45–162 (1986)
11. Bogdan, R.J. (ed.): Jaakko Hintikka. Reidel, W. Koebner (1987)
12. Deely, J.: The supplement of the copula. Rev. Metaphys. **46**, 251–278 (1992)
13. Frede, M.: Prädikation und Existenzaussage. Vandenhöck und Ruprecht, Göttingen (1967). This essay was to be included in: In: Hintikka, J., Knuuttila, S. (eds.) The Logic of Being, cf. p. ix. Reidel, Dordrecht (1986)
14. Frege, G.: Die Grundlagen der Arithmetik. Breslau (1884)
15. Frege, G.: Über Begriff und Gegenstand. Vierteljahrschr. Wiss. Philos. **16**, 192–205 (1892). In: Frege, G.: Kleine Schriften, 2nd edn., ed. by Angelelli, I. Olms Verlag, Hildesheim (1990)
16. Frege, G.: Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik. In: Frege, G.: Kleine Schriften, 2nd edn., ed. by Angelelli, I. Olms Verlag, Hildesheim (1990)
17. Frege, G.: Was ist eine Funktion? In: Festschrift Ludwig Boltzmann gewidmet zum sechzigsten Geburtstage, pp. 656–666. Barth, Leipzig (1904). In: Frege, G.: Kleine Schriften, 2nd edn., ed. by Angelelli, I. Olms Verlag, Hildesheim (1990)

[6] On the history of the concept of "natural" language cf. [8].

18. Frege, G.: Dialog mit Pünjer über Existenz. In: Hermes, H., Kambartel, F., Kaulbach, F. (eds.) Nachgelassene Schriften. Meiner Verlag, Hamburg (1969)
19. Frege, G.: Über Schoenflies: Die logischen Paradoxien der Mengenlehre. In: Hermes, H., Kambartel, F., Kaulbach, F. (eds.) Nachgelassene Schriften. Meiner Verlag, Hamburg (1969)
20. Frege, G.: Einleitung in die Logik. In: Hermes, H., Kambartel, F., Kaulbach, F. (eds.) Nachgelassene Schriften. Meiner Verlag, Hamburg (1969)
21. Frege, G.: Kurze Übersicht meiner logischen Lehren. In: Hermes, H., Kambartel, F., Kaulbach, F. (eds.) Nachgelassene Schriften. Meiner Verlag, Hamburg (1969)
22. Frege, G.: Nachgelassene Schriften, ed. by Hermes, H., Kambartel, F., Kaulbach, F. Meiner Verlag, Hamburg (1969)
23. Frege, G.: Posthumous Writings. Transl. by Long, P., White, R., University of Chicago Press, Chicago (1979)
24. Frege, G.: Wissenschaftlicher Briefwechsel, ed. by Gabriel, G. et al. Meiner Verlag, Hamburg (1976)
25. Frege, G.: Kleine Schriften, 2nd edn., ed. by Angelelli, I. Olms Verlag (1990)
26. Haaparanta, L.: Frege's doctrine of being. Acta Philos. Fenn. **39**, 182 (1985)
27. Haaparanta, L.: Frege on existence. In: Haaparanta, L., Hintikka, J. (eds.) Frege Synthesized, p. 155–174. D. Reidel Pub. Co, Dordrecht (1986)
28. Haaparanta, L.: On Frege's concept of being. In: Hintikka, J., Knuuttila, S. (eds.) The Logic of Being. Reidel, Dordrecht (1986)
29. Hintikka, J.: Quantifiers in logic and quantifiers in natural languages. In: Saarinen, E. (ed.) Game-theoretical Semantics, p. 27–47. Reidel, Dordrecht (1979)
30. Hintikka, J.: "Is" semantical games, and semantical relativity. J. Philos. Log. **8**, 433–468 (1979). Reprinted with "minor changes" in: Hintikka, J., Kulas, J. (eds.): The Game of Language. Reidel, Dordrecht (1983)
31. Hintikka, J.: Kant on existence, predication, and the ontological argument (1981). Reprinted in: Hintikka, J., Knuuttila, S. (eds.) The Logic of Being. Reidel, Dordrecht (1986)
32. Hintikka, J.: Semantics: a revolt against Frege. In: Floistad, G., von Wright, G.H. (eds.) Contemporary Philosophy, a New Survey, vol. 1, pp. 57–82. Nijhoff, The Hague (1981)
33. Hintikka J.: Semantical games, the alleged ambiguity of "is" and Aristotelian categories. Synthese **54**, 443–467 (1983). Reprinted, revised, shorter title in: Hintikka, J., Kulas, J. (eds.): The Game of Language. Reidel, Dordrecht (1983)
34. Hintikka, J., Kulas, J. (eds.): The Game of Language. Reidel, Dordrecht (1983)
35. Hintikka J.: A hundred years later: the rise and fall of Frege's influence in language theory. Synthese **59**, 27–49 (1984)
36. Hintikka, J., Knuuttila, S. (eds.): The Logic of Being. Reidel, Dordrecht (1986). This volume includes, aside from an introduction by the editors, the following contributions: Kahn, C.H.: Retrospect on the verb 'to be' and the concept of being; Mates, B.: Identity and predication in Plato; Dancy, R.M.: Aristotle and existence; Hintikka, J.: The varieties of being in Aristotle; Ebbesen, S.: The chimera's diary; Jacobi, K.: Peter Abaelard's investigations into the meaning and functions of the speech sign 'est'; Weidemann, H.: The logic of being in Thomas Aquinas; Knuuttila, S.: Being qua being in Thomas Aquinas and John Duns Scotus; Alanen, L.: On Descartes's argument for dualism and the distinction between different kinds of being; Hintikka, J.: Kant on existence, predication, and the ontological argument; Haaparanta, L.: On Frege's concept of being
37. Hintikka, J.: The varieties of being in Aristotle. In: Hintikka, J., Knuuttila, S. (eds.) The Logic of Being. Reidel, Dordrecht (1986)
38. Hintikka, J.: Introduction. In: Hintikka, J., Knuuttila, S. (eds.) The Logic of Being. Reidel, Dordrecht (1986)
39. Hintikka, J.: Annotated bibliography. In: Bogdan, R.J. (ed.) Jaakko Hintikka. Reidel, Dordrecht (1987)
40. Hintikka, J., Sandu, G.: A Revolution in Logic? Nordic J. Philos. Log. **1**(2), 169–183 (1996). http://www.hf.uio.no/ifikk/forskning/publikasjoner/tidsskrifter/njpl/
41. Janik, A., Berger, P.: Gottlob Frege: Briefe an Ludwig Wittgenstein. Grazer Philos. Stud. **33/34**, 5–33 (1989)
42. Hubien, H.: Logiciens medievaux et logique d'aujourd'hui. Rev. Phil. de Louvain **75**, 219–233 (1977)

43. Kant, I.: Die falsche Spitzfindigkeit der vier syllogistischen Figuren. In: Werke, Band 2. Frankfurt am Main (1977). http://www.zeno.org/nid/2000918791X
44. Lorenzen, P.: Normative Logic and Ethics. Bibliographisches Institut, Mannheim (1969)
45. Maritain J.: Eléments de philosophie. II. L'ordre des concepts. I. Pétite Logique, 15th edn. Téqui, Paris (1946). English translation: An Introduction to Logic. London (1937)
46. Mendelsohn, R.L.: Frege's two senses of 'is'. Notre Dame J. Form. Log. **28**, 139–160 (1987)
47. Mignucci, M.: Aristotle's theory of predication. In: Angelelli, I., Cerezo, M. (eds.) Studies on the History of Logic (Proceedings of the III Symposium on the History of Logic, Pamplona, May 1993). Walter de Gruyter, Berlin, New York (1996)
48. Weyl, H.: Philosophy of Mathematics and Natural Science. Princeton University Press, Princeton (1949)

I. Angelelli (✉)
Philosophy Department, The University of Texas at Austin, Austin, USA
e-mail: angelelli@utexas.edu

The Relation Between Logic, Set Theory and Topos Theory as It Is Used by Alain Badiou

Régis Angot-Pellissier

Abstract Topos theory plays, in Alain Badiou's philosophical model, the role of inner logic of mathematics, given its power to explore possible mathematical universes; whereas set theory, because of its axiomatics, plays the role of ontology. However, in category theory, which is a vaster theory, topos theory embodies a particular axiomatic choice, the fundamental consequence of which consists in imposing an internal intuitionist logic, that is a non-contradictory logic which gets rid of the principle of excluded middle. Category theory shows that the dual axiomatic choice exists, namely the one imposing, this time, a logic of the excluded middle which accepts true contradictions without deducing from them everything, and this is called a paraconsistent logic. Therefore, after recalling the basics of category and topos theory necessary to demonstrate the categorical duality of paracompleteness (i.e. intuitionism) and paraconsistency, we will be able to introduce into Badiou's thought category theory seen as a logic of the possible ontologies, a logic which demonstrates the strong symmetry of the axioms of excluded middle and of non-contradiction.

Keywords Alain Badiou · Category theory · Set theory · Topos · Intuitionist logic · Paraconsistent logic · Heyting algebra · Brouwer algebra · Internal logic · Open subsets · Closed subsets · Categorical duality · Closed Cartesian category · Closed co-Cartesian category

Mathematics Subject Classification (2000) Primary 00A30 · Secondary 03G30, 18A15

1 The Place of Mathematical Theories (Set, Category, Topos) in Alain Badiou's Philosophy

Mathematics plays a central role in the theory of the French philosopher Alain Badiou[1], given that in his view they embody ontology. However, his conception of mathematics has evolved progressively within his increasingly closer study of new mathematical theories

[1] This paper is the written version of a talk given, in front of Alain Badiou, at the postgraduate congress "Autour de la pensée d'Alain Badiou", organised by Alessio Moretti on behalf of the "Association des Thésards en Philosophie de Nice (ATP)" at the University of Nice, France, 19th March 2004 (Alain Badiou was an Invited Speaker at this event).

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_7

such as category and topos theory. As he writes in his paper "Platon et/ou Aristote-Leibniz. Théorie des ensembles et théorie des topos sous l'oeil du philosophe" (1995) ([1])[2], mathematics is no longer just ontology, it is rather onto-logical. Here the hyphen "-" is very important insofar it relates two diametrically opposed conceptions of mathematics.

The "onto-" prefix expresses the purely ontological part of mathematics, following a vision inherited by Plato, according to Badiou's analysis. In fact, according to the French thinker mathematics is thinking and, as such, it thinks "being", which makes it an ontology (i.e. a study of being). Set theory, as inherited from Cantor at the end of the 20th century embodies this "Platonic" conception, at least in the interpretation given by Badiou to this adjective (Platonic), which is situated on the opposite side with respect to what is commonly called "mathematicians' Platonism". As it happens, by virtue of its axiomatic treatment of the notion of "set", which is left undefined by set theory, this theory consists in a *decision* about being, the "essence" of which is seen in the pure multiplicity symbolised by the notion of set ("set" means nothing other than pure multiplicity). It must be remarked that the logic immanent in set theory is classical logic (i.e. a logic with excluded middle and non-contradiction).

The "-logic" suffix concerns a more Aristotelian conception of mathematics, where the latter is seen from the viewpoint of the formal consistency of its arguments. However, Badiou goes farther. Following Leibniz, mathematical logic is not, for Badiou, a linguistic phenomenon, it is on the contrary a logic of the possible ontologies. This new aspect of mathematics, far from the Hilbertian formal rigors, is rediscovered by Badiou in the more recent theories of categories and of topoi. As it happens, category theory and, more particularly (inside it), topos theory expound, classify and explore the possible mathematical universes. Topos theory makes explicit the conditions that must be verified by a universe in order to make it possible to practice in it mathematics. However, the power of this theory does not stop here; it also puts into light the fact that the internal properties of a universe do induce a particular kind of logic in it. Consequently, the theorems of topos theory say nothing other than the following thing: to any given axiomatic decision relative to a universe corresponds, for that universe, a particular logic. This means, in philosophical terms, that with any decision about being, i.e. with any ontological decision, derives a logic that follows from that choice. Whence we find that the choice of a local, and not a qualitative, difference, joined with the choice of the unicity of void (which are the fundamental choices of set theory) entails that the logic there will be classic.

Now, topos theory, a development of category theory that appeared after World War II, although developed by classical mathematicians,[3] has shown by irony of history that the logic underlying a topos is, in general, an intuitionist logic and that classical logic is only a very particular case. Let us recall here that inside the mathematics of topoi classical logic is nothing else than a particular intuitionist logic, given that it shares the same axiomatic basis but verifies, additionally, a few more axioms, among which the axiom of excluded middle. This is the reason why Badiou, considering topoi as the only possible universes, writes in the aforementioned paper ([1] page 78):

[2] "Plato and/or Aristotle–Leibniz. Set theory and topos theory as seen by the philosopher's eye".

[3] By classical mathematicians I mean mathematicians who are not partisans or followers of intuitionism. They constitute the overwhelming majority of the mathematical community.

> It is remarkable to acknowledge that everything, in the classification of Topos, turns around the excluded middle. For the book gamma of *Metaphysics* shows well that in between the principle of non-contradiction (which is the issue of the "war" with the Sophists) – a principle which any Topos verifies – and the principle of excluded middle, which is the founding issue of the stability of the possible attributes of a substance (for they belong to it *or* they do not belong to it), we have made a change: we stepped from general logic to a special logic, which is appropriate to the ontological choices of Aristotle.

This is why we find in Badiou, the traditional *asymmetry* between the axiom of non-contradiction and the axiom of excluded middle. My purpose in this paper consists in showing that Badiou's analysis of the logic of mathematics can (and must) be pushed farther. Relying on the axiomatics of category and topos theory, as well as on the practice of mathematics on the level of contemporary research, I propose to establish category theory as the logic of mathematics, which for philosophy has three main consequences, which are intrinsically related. Firstly, this will put topos theory in its place, namely that of an ontological choice, already made (i.e. no longer open), of the possible mathematical universes; a choice with respect to which sets constitute a particular case. Secondly, the asymmetry between non-contradiction and excluded middle will no longer any reason to exist, mathematically speaking, given that it will dissolve into a perfect categorical duality; a duality only veiled by the choice of topoi as the only possible mathematical universes. Thirdly, paraconsistent logics, that is the logics that accept true contradictions without getting trivial[4], will be rehabilitated as logics dual of intuitionist (i.e. paracomplete) logic, which is the one that gets rid of the principle of excluded middle without getting trivial.

Here philosophy stops, and mathematics begins. Nevertheless, my discourse bearing mainly on the formalism and on the analysis of modern mathematical praxis is directed both to the neophyte and to the experimented mathematical reader. Consequently, this paper can be read according to different levels of reading and it can, at least I hope so, either initiate or bring some light to the readers of Badiou on the elements of category and topos theory, which are central in his theory starting from the *Court traité d'ontologie transitoire* ([2]).[5]

2 Topoi Axiomatics

2.1 Categories

Contrary to set theory which deals only with collections of mathematical objects, category theory has been created in order to deal with mathematical structures. However, these structures are less carried by the objects themselves than by the relations between them, whose main property is the preservation of the structure. For example, what is the difference between a mathematical group and the set of its elements? The group is endowed with morphisms of the group, which are relations between elements of the group

[4] One says that a logic is trivialised when it deduces everything, i.e. when inside it everything becomes a theorem (so that the notion of "theoremhood" no longer means anything).

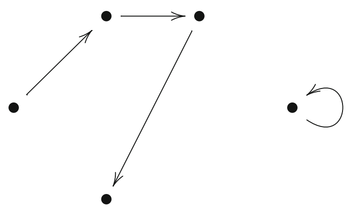
[5] They will become even more central in his book *Logiques des mondes* (2006).

and which preserve the group structure. This is why category theory deals with objects and their relations.

The formal definition of a category is very simple and short, nevertheless the subsequent theory has a power that has no pair in any branch of mathematics, even nowadays. It is given below.

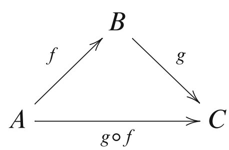
A category consists of a collection of objects and for every couple of objects a set of oriented arrows between these objects. We could see that the undefined notions of object and arrow (also called morphisms) are the primary notions of category theory, whereas the notions of set and membership are the primary notions of set theory. Furthermore we could remark that, by definition, a category can easily deal with an inconsistent multiplicity of objects (we talk of a class of objects). This is the case of the category of sets whose objects are the sets and whose arrows are the sets applications, yet we know quite well from Russell's paradox that the collection of all sets cannot be a set. At this stage of the definition, a category is like an oriented graph whose vertices are the objects of the category and whose paths are the arrows of the category, like in the following example:

$C =$



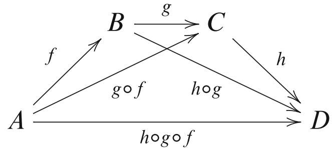
Nevertheless, to be a true category a minimal number of axioms must be checked. In terms of graphs, they ensure that the graph can be intuitively travelled by following the directed paths.

First of all, a category is endowed with a composition operation:



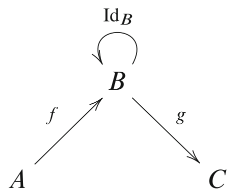
If we go from A to B by f, then if we go from B to C by g, we can always go straightaway from A to C by following the composed path $g \circ f$, which runs along f and then g.

This composition verifies the associativity axiom:



If we go from A to B by f, then from B to C by g and then from C to D by h, we can either go from A to C by $g \circ f$, then from C to D by h or go from A to B by f, then from B to D by $h \circ g$. In both cases we go from A to D passing through B and C by the same path noted $h \circ g \circ f$.

This composition also verifies the unitary axiom:



This axiom tells us that, in any vertex of the graph, we can stay on the spot. Here the staying on the spot B is noted Id_B and is called the identity morphism for B. Its properties are the following: if we go from A to B by f and then we stay on B, we are in fact going from A to B by f. In the same manner, if we first stay on B and then go from B to C by g, we directly go from B to C by g. Here we must insist on the difference between paths from B and going back to B with the identity morphism for B: it is the only path from B to B that stays on the spot. In fact, it is the morphism representing the object B: it is the object B as an identical identity of B with itself[6].

In short, a category is a collection of objects and oriented arrows between these objects endowed with an associative and unitary composition for arrows and that is all. Despite this imaged presentation of category theory, this theory is not a visual game but an extremely powerful mathematical concept; indeed each mathematical structure forms a category whose objects are the studied objects endowed with the structure and the arrows the applications or relations preserving the structure. In doing so, we have a category of sets with their applications, a category of groups with their morphisms of groups, a category of vector spaces with their linear applications, a category of topological spaces with their continuous functions and so on...

Furthermore, the definition of a category is so poor that category theory also supports non-mathematical structures; there exist logical categories, linguistic theories categories, lambda-calculus categories...

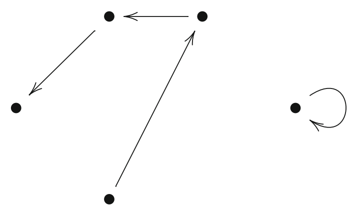
Last but not least, category theory, unlike set theory, studies not only categories and graphs but it can also study itself, as we could expect from it! As category theory deals with all mathematical structures and as categories form a mathematical structure, there must exist a category of categories, even if some technical size limitations must be made in order to cut Russell's paradox short. That's why, in category theory, language and metalanguage are at the same level, with the consequence that, in topoi theory, logical syntax and logical semantics are of the same nature and then lose their profound difference.

The last element of category theory that we would like to present here is duality, which deals with the orientation of the arrows. As we have seen, a category must be seen as an oriented graph. However, the orientation of the graph is arbitrary and it is also natural to consider the same graph with all arrows in opposite direction; we call it the opposite category and, if we take our former example of category C, we obtain the following opposite

[6] *Cf* Chap. XVI of the *Court traité d'ontologie transitoire* ([2]).

category:

$C^{op} =$



Here also, we are not in a visual game but in the heart of the notion of categorical duality, a fundamental notion of category theory. Categorical duality is the fact that to each category corresponds an opposite category with the same objects but with all arrows in the opposite direction. It is with this notion that important tools as sheaves or colimits are defined.

We must remark at this point that as orientation in graphs is arbitrary, so is the choice between a category and its dual. Furthermore, the dual of the opposite category is the category itself. Then it is categorically impossible to distinguish, between a category and its dual, which is the primordial category as in the question about which came first, the hen or the egg.

Besides, some known examples of duality contain a profound mathematical significance, like open/closed topological subsets or affine schemes/rings. We shall return to one of them later.

To conclude this presentation of category theory, let us recall that this simple and graphic notion that transcends mathematical structures is due to MacLane between the two World Wars. From it, the notion of topos was born in the 1960s. This notion is the subject of the following section.

2.2 Topoi and Internal Logics

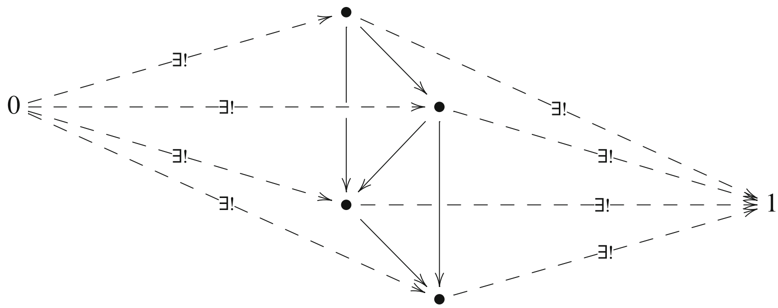
The notion of topos[7] is more complex than the notion of category and needs difficult categorical tools. Therefore, here we shall not give topos axiomatics but we shall comment on its definition: a closed Cartesian category endowed with a subobject classifier. For an axiomatic approach, the reader is referred to reference books and for a philosophical approach to the paper of Alain Badiou quoted above or Lavendhomme's book *Lieux du sujet. Psychanalyse et mathématiques* ([5]).

Nevertheless, it is important to understand what is meant by closed Cartesian category. It means that to be a topos a category must contain some objects with a typical configuration in terms of a graph.

First of all, a closed Cartesian category must possess an initial object and a terminal object, two dual categorical notions. An initial object is an object with a unique arrow to each object of the category, we denote it by 0. If we view arrows as order relations, it is like a minimal object of the category. Dually, a terminal object is like a maximal object

[7] from old Greek "topos", place, whose old-Greek plural is "topoi".

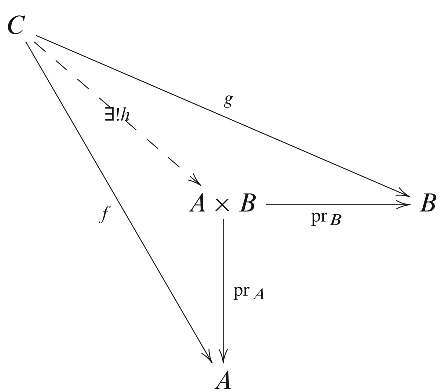
denoted by 1.

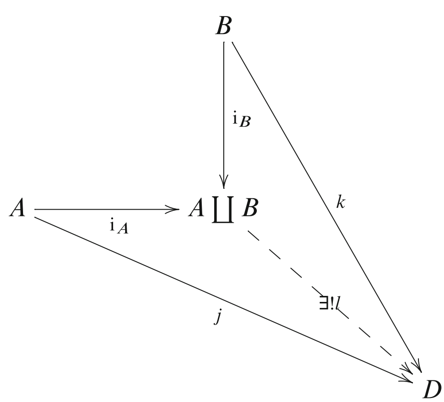


The above parallelogram must be considered as the heart of our category, whose objects 0 and 1 are, respectively, the initial and terminal objects. We say that a diagram commutes when, in the above graph, for each triangle the third arrow is the composition of the two other arrows.

Let us examine the well-known category of sets: the initial object is the empty set, which is included in only one manner in each set, and the final object is the point because all sets go to it in only one manner, which consists of identifying all their elements. This example shows that although the empty set is the smallest element, the point is not the greatest element and, furthermore, it can go to all non-empty sets. Generally, it is possible that some objects go to the initial object or that the terminal object go to some objects!

The second requirement for a category to be a closed Cartesian category is the existence of a product and of a coproduct, two more dual notions. The product of two objects A and B is denoted $A \times B$ and is the greatest object of the category that goes to both A and B, the arrows of the product to A or B are called projections. The expression "the greatest" signifies that, for each object C going to A by f and B by g, there exists a unique arrow h from C to $A \times B$ factorizing f and g through the projections. The coproduct, dual of the product, denoted $A \coprod B$ is the smallest object to which both A and B go.





Here the diagrams commute, which is the visual translation of the factorization.

In the category of sets, the product is the Cartesian product; for example, the product of two real lines $I\!R$ is the real plane $I\!R^2$ with its two projections on the coordinate axes. The coproduct is the disjointed union of sets; the coproduct of two real lines is a set consisting in two real lines.

The last and central element of the notion of closed Cartesian category is the existence of exponential objects. This is the most subtle notion and the richest from a logical point of view. For each pair of objects B and C, there exists an exponential object C^B verifying the following property:

$$A \times B \longrightarrow C \qquad \Leftrightarrow \qquad A \longrightarrow C^B$$

To each arrow of the product $A \times B$ to C there corresponds a unique arrow from A to C^B and conversely.

A more concrete vision is given by the example of the category of sets. An arrow from $A \times B$ to C is, in fact, a two-variable function $f(x, y)$ where x lies in A, y in B and the result of the function lies in C. The exponential object C^B is the set of arrows from B to C. In particular, for each element a in A, $f(a, y)$ is a one-variable function from B to C. Therefore, for each element a in A, we can associate a function $f(a, y)$ that is an element of C^B. Then we have constructed a function from A to C^B that represents the two-variable function $f(x, y)$. In another way, it is clear that the function $g(a)$ from A to C^B gives for each a in A a function from B to C denoted $g(a)(.)$. That means that for each y in B, $g(a)(y)$ is an element of C. We can construct a two-variable function as follows: $g(x, y) = g(x)(y)$, which represents the function g. This result, called adjunction, between product and exponential object in the category of sets bespeaks a crucial mathematical phenomenon: to each two-variable function there corresponds a unique parametric collection of one-variable functions and conversely. We must be careful not to generalize this result to each category equipped with exponential objects! Let us remark that exponential objects and products are really relied by the categorical adjunction. In the case of sets, the parametric collection of one-variable functions of the exponential object represents the two-variable functions whose source is the product of sets, which is, in fact, the set of the pairs of variables.

Where does the logic lie in this axiomatics of closed Cartesian category? The answer is quick if we consider the mathematical point of view of deductive systems, which is a set of formulas formed with variable letters and logical connectors verifying some axioms. Here in the axiomatics of closed Cartesian category we have, in fact, the axiomatics of positive (i.e. without negation) intuitionist propositional calculus.

We can consider that the objects of the category are the formulas of the language and the arrows are the deductions between formulas. The false $\perp$ is the initial object 0 because it goes to each object in the same manner that from the false one can deduce all formulas. Dually, the true $\top$ is the terminal object 1 because true is deduced from all formulas. We can clearly see that all objects going to 0 correspond to false formulas and all objects to which 1 goes correspond to true formulas.

As for logical connectors, the product is the conjunction "and" $\wedge$ and the coproduct is the inclusive disjunction "or" $\vee$. We know, in fact, that from $A \wedge B$ we can deduce both A and B and that if from C we can deduce both A and B, then from C we can also deduce

$A \wedge B$; $A \wedge B$ is then the greatest formula from which we can deduce both A and B and, therefore, it is the product $A \times B$. We can demonstrate by duality that the coproduct axiom is the very definition of the connector $\vee$.

For the implication, we shall use the following definition: $\Rightarrow$ is the binary connector verifying the following property: from A and B we can deduce C if and only if from A we can deduce $B \Rightarrow C$. By this definition, we can recognize in the implication the notion of exponential object. Furthermore, applying the definition of the exponential object to the objects 1, B and C gives us back the deduction theorem, as follows. To each arrow from 1 to C^B corresponds a unique arrow from $1 \times B$ to C and conversely. However, we can easily show by the above axioms that $1 \times B$ is, in fact, B. Therefore, we have that $B \Rightarrow C$ is true (i.e. from 1 we deduce C^B) if and only if from B we can deduce C.

Finally we can complete our positive intuitionist propositional calculus in a complete intuitionist propositional calculus by defining the negation connector $\neg$ as follows: $\neg A$ is by definition $A \Rightarrow 0$ (this is the most frequent definition of intuitionist negation). Categorically 0^A is the negation of A. However, applying the definition of the exponential object to the trivial deduction "from $\neg A$ we can deduce $\neg A$" (which is true because each category must have one identity arrow for each of its object) gives us:

$$0^A \xrightarrow{\text{Id}} 0^A \qquad \Leftrightarrow \qquad 0^A \times A \longrightarrow 0$$

Therefore, the adjoint arrow to the identity arrow of 0^A is the arrow from $0^A \times A$ to 0, which means that the formula $\neg A \wedge A$ is false, i.e. that the axiom of non-contradiction holds. Category theory shows us that the axiom of non-contradiction is, in fact, adjoint to the identity arrow of $\neg A$ and, moreover, that this fact is a direct consequence of the axiomatics of a closed Cartesian category endowed with an intuitionist negation. This is why it is natural that the internal logic of a topos is intuitionist, as this fact is inscribed in its axiomatics.

This last result is fundamental because category theory shows not only that the axiomatics of topos in itself induces an intuitionist logic (in fact, it is true for a closed Cartesian category as well) but also that this choice of axiomatics carries in its notion of exponential object the source of non-contradiction. We can conclude, therefore, that the non-contradiction is purely a choice induced by an axiomatics of mathematical nature, and the categorical duality shows us that this choice is purely arbitrary and that we can, in fact, choose to accept contradiction and preserve the axiom of excluded middle, which is the object of the following section.

2.3 Cotopoi and Paraconsistent Logic

Categorical duality enables us to define a "cotopos" as a closed co-Cartesian category with quotient classifier.[8] However, let us insist upon co-Cartesian closure. As dual of the

[8] This notion and the analysis of its internal logics are explored by William James and Chris Mortensen in their book *Categories, Sheaves and Paraconsistent Logic*.

notion of Cartesian closure, we find in it all dual constructions: to be a closed co-Cartesian category a category must have initial and terminal objects, products and coproducts as a closed Cartesian category because these constructions are dual one from another.

However, it does not need to have exponential objects, which is the heart of a Cartesian closure. On the contrary, it must have a co-exponential object. For all pairs of objects A and B, there must exist a co-exponential object A_B with the following property:

$$A_B \longrightarrow C \qquad \Leftrightarrow \qquad A \longrightarrow B \coprod C$$

To each arrow from the co-exponential object A_B to C there corresponds a unique arrow from A to the coproduct $B \coprod C$ and conversely. As an exponential object is the adjoint to a product, a co-exponential object is, dually, the adjoint to a coproduct.

To illustrate this notion of co-exponential object, we cannot examine the case of the category of sets, which is not a closed co-Cartesian category. However, its restriction to the category of subsets of a fixed set with inclusions for arrows is a closed co-Cartesian category. Here the co-exponential object is the set difference $A - B$ formed by all elements of A not included in B. In fact, if $A - B$ is included in C, as A may contain either elements of A not included in B or elements of A included in B, then A is included in $B \bigcup C$, which is a coproduct. Reciprocally, if A is included in $B \bigcup C$, elements of A are either in B or in C, therefore all elements of A that are not in B are by force in C, then we can conclude that $A - B$ is included in C. Here we can clearly see that the co-exponential object is tied with the set difference and this notion of set difference is itself very closely tied to the excluded middle, because the preceding proof is based on the fact that A is the union of $A \bigcap B$ and $A - B$. Here we again recognize the categorical adjunction because the set difference $A - B$ is tied to the coproduct $(A \bigcap B) \bigcup (A - B)$, which is, in fact, A.

What sort of logic is induced by the axiomatics of a closed co-Cartesian category? Again, we consider the category objects as the formulas of the language and the arrows as the deductions. 0 is again false, 1 is true and products and coproducts are, respectively, the connectors "and" $\wedge$ and "or" (inclusive or) $\vee$. A formula is true when there exists an arrow from 1 to the object which is the formula and it is false when there exists an arrow from the object which is the formula to 0.

The most delicate notion is the co-exponential object, which naturally does not represent the implication but the dual connector called pseudo-difference and denoted by $-$. This connector plays the role that was usually played by the implication. As implication is adjoint to conjunction "and", pseudo-difference is adjoint to inclusive disjunction "or" from $A - B$ we can deduce C if and only if from A we can deduce B or (inclusive or) C. As for implication, there is a deduction theorem: $A - B$ is false if and only if from A we can deduce B. This theorem is a mere consequence of the definition of co-exponential object applied to objects A, B and 0: to each arrow from A_B to 0 there corresponds a unique arrow from A to $B \coprod 0$ and conversely. However, we can easily prove that $B \coprod 0$ is, in fact, B.

In so doing, we have deduced from the axiomatics of a closed co-Cartesian category a propositional calculus dual to the intuitionist positive propositional calculus based on the pseudo-difference which is dual to the implication and, as the latter, intimately tied to deduction.

The natural negation connector for this calculus will be defined by the pseudo-difference true, as intuitionist negation was implication to false: $\neg A$ is then by definition $\top - A$. This negation is, in fact, paraconsistent and with it the preceding calculus becomes a paraconsistent propositional calculus. A logical system is called paraconsistent when contradiction may happen but the logical system does not collapse. This means that, in these systems, it is no more true that from $A \wedge \neg A$ we can deduce all formulas.[9] This is the case of our calculus, which is dual to the intuitionist propositional calculus: intuitionist calculus forbids contradiction and accepts middle terms, therefore by categorical duality our calculus permits contradictions and forbids middle terms.

The fact that the excluded middle axiom is valid in this system (and not the non-contradiction axiom) is derived directly from the definition of co-exponential object applied to the identity arrow of 1_A:

$$1_A \xrightarrow{\text{Id}} 1_A \qquad \Leftrightarrow \qquad 1 \longrightarrow A \coprod 1_A$$

Therefore, the adjoint arrow to the identity arrow of 1_A is the arrow from 1 to $A \coprod 1_A$, which means that the formula $A \vee \neg A$ is true, i.e. that the axiom of excluded middle holds. As non-contradiction is the arrow adjoint to the identity arrow of the negation object defined by exponential and initial objects, the excluded middle is the arrow adjoint to the identity arrow of the negation object defined by co-exponential and terminal objects.

Finally, as the excluded middle axiom is not induced by the axiomatics of a closed Cartesian category (in those categories this means that for each object A of the category, there should exist an arrow from 1 to $A \vee \neg A = A \coprod 0^A$), closed Cartesian category and topoi can be either classical or intuitionist, depending on whether or not this axiom is added to the axiomatics. Dually, the non-contradiction axiom is not induced by the axiomatics of a closed co-Cartesian category (in these latter categories this means that for each object A, there should exist an arrow from $A \wedge \neg A = A \times 1_A$ to 0). Therefore, contradiction is optional: a closed co-Cartesian category can be either classical or paraconsistent. Moreover, if there is contradiction, the logic does not collapse; if a contradiction $A \times 1_A$ exists, this means that there is no arrow from $A \times 1_A$ to 0, but if this contradiction had induced a collapse of the whole logic, all formulas would have been deduced from the contradiction and then there would be arrows from $A \times 1_A$ to all objects of the category and in particular to 0, which cannot be the case. We have then proved that the calculus dual to the intuitionist propositional calculus is paraconsistent because no existing contradiction can induce the collapse of the whole logic.

All this shows that, from a purely formal point of view, the choice of non-contradiction and the choice of excluded middle are strictly dual because they are both induced by axiomatic choices bearing upon structures of mathematical universes that are themselves categorically dual. Therefore, category theory shows us that the intuitionist logics, which are internal to topoi and more generally to closed Cartesian categories, are not central

[9] These logic systems were discovered in 1963 by da Costa in Brazil. It is important to remark that all these systems possess a paraconsistent implication, whereas our calculus does not possess implication but possesses a pseudo-difference, because it is the dual of the intuitionist propositional calculus. This special paraconsistent system was discovered more recently by the Brazilian school (da Silva de Queiroz 2001) ([3]).

logics but symmetric to paraconsistent logics, internal to closed co-Cartesian categories and dual to closed Cartesian categories. The reason for this is that these logics are, in fact, contained in the axiomatics defining these categorical structures. Then a question arises: why do topoi have such an axiomatics, or, which is the same, why are the topoi of an intuitionist nature? The next section gives an answer to these questions.

3 Intuitionism and Paraconsistency: a Duality of the Type Open–Closed Subsets

3.1 Heyting Algebras and Open Subsets of a Topological Space

A known mathematical result says that each intuitionist propositional calculus is equivalent to a Heyting algebra. Despite this barbarous name, a Heyting algebra is a simple tool: it is an ordered set endowed with a minimal element 0, a maximal element 1, for each pair of elements (a, b) an element inferior to both a and b, denoted by $\inf(a, b)$, and an element superior to both a and b, denoted by $\sup(a, b)$, but also an element $a \Rightarrow b$ such as, for all elements c, we have $\inf(c, a) \leq b$ if and only if $c \leq a \Rightarrow b$. The identification with intuitionist logic is as follows: formulas are the elements of the set, deduction is the order on the set, $\perp$ is the minimal element, $\top$ the maximal element, $a \wedge b$ the element inferior to both a and b, $a \vee b$ the element superior to both a and b, for implication itself, and the negation of an element a is defined as $a \Rightarrow \perp$.

The reader will recognize the structure to be of the closed Cartesian category. Heyting algebra is, in fact, a category whose objects are the elements of the set and the arrows the order relations; there exists a unique arrow from a to b if and only if a is smaller than b. However, as a category, Heyting algebra is a closed Cartesian category: 0 is the initial object, 1 the terminal object, $\inf(a, b)$ the product of a and b, $\sup(a, b)$ the coproduct of a and b and $a \Rightarrow b$ the exponential object for a and b. The categorical definition of Heyting algebra is simpler than its definition in set theory: Heyting algebra is an ordered set that is, as a category, a closed Cartesian category.

However, a very important result says that each Heyting algebra corresponds to a unique category of open subsets and conversely so. This result implies general notions of topology and it is not within the scope of this paper to detail topological matters. Nevertheless, we can present the main lines of topology. A topological space is a set endowed with two families of subsets: the open subsets family and the closed subsets family, one being the set complement of the other. The open subsets family verifies some properties: the empty set and the total set are open subsets, infinite unions of open subsets and finite intersections of open subsets are open subsets. Dually, the closed subsets family verifies the following properties: the empty set and the total set are closed subsets, finite unions and infinite intersections of closed subsets are closed subsets. Open subsets of a topological space endowed with inclusions form a category that is a Heyting algebra and reciprocally to each Heyting algebra corresponds a subcategory of the category of open subsets of a topological space. Properties of open subsets show us that the empty set is the initial or minimal object, the total set (the topological space itself) is the terminal or maximal object, the intersection is the product or inf, the union the coproduct or sup. To define the exponential object or implication, we must use the notion of interior: the interior of a subset of a topological set is the greatest open subset included in the subset. Implication $U \Rightarrow V$ will be defined as interior of the intersection of V with the set-complement of U. The induced negation will only be the interior of the set-complement.

Let us take an example to better understand these definitions. The real line $I\!R$ has a natural topology whose open subsets are unions of intervals of type $]a, b[$, a or b could even be ∞. Therefore, the subset of strictly positive reals $]0, +\infty[$ is an open subset and then a formula of our language. Let us denote it by A. $\neg A$ is by definition the interior of the set-complement of A. The set-complement of A is clearly the subset of negative reals $]-\infty, 0]$, but this subset is not open. (Let us note here that 0 is both negative and positive in the French school of mathematics). Its interior, the greatest open subset that it contains, is in fact $]-\infty, 0[$, the subset of strictly negative reals. However, $A \vee \neg A$, which is the union, contains all strictly positive reals and all strictly negative reals, but 0 is missing! Therefore, the axiom of excluded middle is not valid because $A \vee \neg A$ is not the total set, the whole real line.

Thus open subsets of a topological space sustain the intuitionist propositional calculus; as mathematical objects, they follow the rules of intuitionist propositional calculus. Furthermore, each intuitionist propositional calculus has a mathematical translation in open subsets of a topological space. This is a fundamental result tying logic and mathematics. Moreover, it is also the reason why topoi are intuitionist. To understand this, let us recall some mathematics history.

The notion of topos was introduced after World War II by the French school of algebraic geometry of Grothendieck. Their motives were the study of geometric objects such as varieties and schemes by the local–global duality. The credo of this duality is: all which is globally true is locally true and if some locally valid things can “glue” together (i.e. are compatible on intersections), then those things are induced by a global something, which is globally valid. From these considerations the notion of sheaves on a topological space was born: a sheaf is a gluing together of elementary and local bricks which are, in fact, open subsets of a topological space. A Grothendieck topos is only a category of sheaves. However, the internal logic of a Grothendieck topos, bearing by its subobject classifier, is the algebra of the elementary and local bricks glued together to form the sheaves of this topos, and it is, in fact, the algebra of the open subsets of a topological space. Thus it is natural that Grothendieck topoi are intuitionist because they were created for a mathematical method of study of geometry by the local–global duality. Moreover, the notion of local is the essence of the open subsets in topology and is tied with the gluing together principle, which builds the global from compatible locals, which is also a typical intuitionist principle. However, the irony is that all this mathematical history is of a most classical nature!

From this notion of Grothendieck topos, some mathematicians and logicians such as Lawvere defined the notion of elementary topos, the basis of topoi theory, with what is common to sheaves and sets. Then in the notion of topos they resumed the whole axiomatics necessary for a universe of a generalized type of sets. Therefore, topos theory is the pendant of category theory because for mathematicians the second one explores all mathematical structures and the first one only the generalized sets as sheaves. Furthermore, these generalized sets universes can only be intuitionist or even classical if the excluded middle axiom holds, because all are based on the gluing together principle, which builds the global from compatible locals.

Nevertheless this is not the only one type of mathematics. As we saw, the notion of closed Cartesian category, the heart of the topos notion, with its internal intuitionist logic has a dual notion of closed co-Cartesian category with paraconsistent logic. In the next section, we shall explore this paraconsistent logic and what is behind it.

3.2 Co-Heyting Algebra and Closed Subsets of a Topological Space

The categorical duality permits us to define co-Heyting algebras, also called Brouwer algebras although they are paraconsistent, as dual categories of the Heyting algebras viewed as categories. Their structure is very simple: they are ordered sets that are closed co-Cartesian categories. Therefore, these algebras possess a minimum, a maximum, inf and sup, and also differences defined by $a - b \leq c$ if and only if $a \leq \sup(b, c)$. As each intuitionist propositional calculus is equivalent to a Heyting algebra, each paraconsistent propositional calculus dual of an intuitionist calculus is equivalent to a co-Heyting algebra, formulas being objects, deductions being order relations, false being the minimum and true the maximum, conjunctions being inf and inclusive disjunctions being sup, differences being themselves and negations being differences with true.

As we saw before, a very important theorem identifies Heyting algebras with algebras of open subsets of a topological space, letting open subsets be mathematical incarnations and essences of intuitionist propositional calculus. Categorical duality then gives us a dual theorem: co-Heyting algebras corresponds to dual categories of open subsets algebras viewed as categories. Yet, concretely what are these algebras? The answer is simple and elegant: the dual category of an open subsets algebra viewed as a category is the algebra of the corresponding closed subsets! We recall here that closed subsets are exactly the set-complements of open subsets in a topological space.

The co-Heyting algebra of closed subsets of a topological space is then the category of these closed subsets with inclusions for arrows. The minimum is the empty set and the maximum the total set (the topological space itself), which are both closed subsets by the definition of a topology. Inf are intersections and sup are unions, closed subsets being stable by infinite intersections and finite unions. In order to define difference, we shall use the topological notion of closure: the closure of a subset of a topological space is the smallest closed subset containing the subset. Difference $F - G$ will then be defined as closure of the set difference, i.e. as the closure of the intersection of F with the set-complement of G. Negation will be simply the closure of the set-complement.

Let us return to our example of the real line $I\!R$. 0 was the non-excluded middle of the intuitionist propositional calculus of open subsets because it is not a member of the open subset A of strictly positive reals nor a member of its negation $\neg A$, the open subset of strictly negative reals. Thus the proposition $A \vee \neg A$ is not the total set (the whole real line) because 0 is missing and, consequently, this proposition is not true. On the contrary, here in the closed subsets propositional calculus, 0 bears a contradiction without a collapse of the whole calculus. The closed subsets are unions of intervals of type $[a, b]$, $]-\infty, b]$ or $[a, +\infty[$. Let be B the closed subset of positive reals $[0, +\infty[$ (here again recall that 0 is both positive and negative for the French school of mathematics), we can remark that it contains 0, which is a positive number. Its negation $\neg B$ is the closure of its set-complement. The set-complement of B is the open subset of strictly negative reals $]-\infty, 0[$, but it is easy to see that the smallest closed sunset containing it is the closed subset of negative reals $]-\infty, 0]$, which contains 0, which is a negative number. Therefore, the intersection $B \wedge \neg B$ is the closed subset of reals, both positive and negative, thus it is the singleton (a set with only one element) $[0, 0]$. We can conclude that a contradiction holds for 0. So, $B \wedge \neg B$ is not included in the empty set, which means that from $B \wedge \neg B$ we cannot deduce false and, furthermore, that the contradiction $B \wedge \neg B$ is not false and that the logic does not collapse. In fact, we have even the very important following result: from $B \wedge \neg B$ we can deduce C if and only if the closed subset C contains 0[10]! This propositional calculus of closed subsets of a topological space is then paraconsistent because contradictions can exist and the logic does not collapse.

We can thus view closed subsets as the essence of the paraconsistency dual of intuitionism. As open subsets correspond to a mathematical method, it is the same for the closed subsets. We have seen that following Grothendieck, French algebraic geometry studies geometric objects with the point of view of the duality global–local: global data are valid on the local level and each gluing together of compatibles local data forms a global data. However, at the beginning of the 20th century, the Italian school of algebraic geometry studied geometric objects, varieties by example, from the point of view of their equations. These equations being polynomials, it is sufficient to know their decompositions in more simple polynomials to know what the subvarieties of a given variety are. This method is called *splitting* and gives many important results in algebra about rings and rings of polynomials, mathematical structures forming categories but not topoi. Theses varieties are closed sets because they are formed by all points verifying their equation. More generally each closed subset corresponds to a set of solutions of one or more equations or one or more not strict inequations. Closed subsets then correspond to an equational and polynomial way to study algebraic geometry. This method is neither exclusive nor a rival of the global–local method: it is simply dual and complementary.

Two anecdotes complete this part. The first one is that this formal duality of the global–local and polynomial method in mathematics is symbolized in a very beautiful mathematical result. Schemes, geometric structures of a sheaf type studied by Grothendieck and his school, have affine schemes for elementary bricks. However, the dual category of affine schemes is the category of (commutative unitary) rings whose extensions, the rings of

[10] This contradicts the principle of classical logic known as Pseudo-Scot or *ex falso quodlibet*, which says that from a contradiction such as $B \wedge \neg B$ we can deduce anything.

polynomials, are the worlds where the equations defining varieties live, geometric structures studied by classic algebraic geometry from the time of the Italian school.

The second anecdote is based on the preceding result. Bourbaki, an association of French mathematicians, which has made a refoundation of classic mathematics from classical logic and set theory to more complex theorems by reconstructing one after another all mathematical notions, also wanted to deal with recent results on sheaves, schemes and topoi, which are intuitionist notions. However, in order to do that, the Bourbaki group would have had to rework the set foundations of mathematics to have a categorical foundation, because category theory is the basis of these notions. That would be too much work and a new beginning! Therefore, on applying the categorical duality of affine schemes and rings, Bourbaki was able to demonstrate very difficult results of Grothendieck algebraic geometry only from the point of view of rings, making, without knowing it, paraconsistent mathematics.

This last example shows very well the complementarity of two mathematical approach of geometry, the global–local approach and the polynomial or equational approach. Therefore, we shall end our paper on examples where intuitionism and paraconsistency are not only dual but complementary.

3.3 Subsets, Subgraphs: Universes Uniting Intuitionism and Paraconsistency

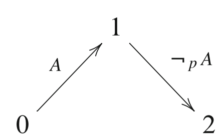
Up to now categorical duality has been viewed from a Manichean approach: either we are in a category or we are in its dual category. Nevertheless, some categories are their own duals: they are auto-dual. That is the case by example of the category of subsets of a given set and also of the category of subgraphs of a given graph. However, the category of sets and the category of graphs are, in fact, topoi and then their internal logic is intuitionist and lies in their subobjects classifier, which assures that each category of subobjects of a given object is a Heyting algebra. Therefore, categories of subsets of a given set or those of subgraphs of a given graph are Heyting algebras, and then closed Cartesian categories with a propositional intuitionist calculus. However, these categories are, in fact, auto-dual because they are their own dual. As dual to themselves, they are also co-Heyting algebras and consequently closed co-Cartesian categories with a propositional paraconsistent calculus. We then have two types of categories, both closed Cartesian and closed co-Cartesian categories.

When a category is both closed Cartesian and closed co-Cartesian, it possesses an initial object 0, a terminal object 1, products $\times$, coproducts $\coprod$, exponential objects $\Rightarrow$ and co-exponential objects $-$, respectively, adjoint to products and to coproducts. The deduction of B from A corresponds both to the fact that $A \Rightarrow B$ is true and that $A - B$ is false. We can then define, for each element A, two negations: an intuitionist negation $\neg_{\mathrm{i}} A = A \Rightarrow 0$ and a paraconsistent negation $\neg_{\mathrm{p}} A = 1 - A$. The intuitionist negation verifies the non-contradiction axiom and with the exponential objects gives a propositional intuitionist calculus to the category. The paraconsistent negation verifies the excluded middle axiom and with the co-exponential objects gives a propositional paraconsistent calculus to the category. The fundamental result is the following: if the two negations coincide then the logic of the category will be classical and conversely so.

This phenomenon occurs in particular in the category of subsets of a given set which is both a Heyting and co-Heyting algebra; the morphisms are the inclusions, 0 is the empty set, 1 the total set, $\times$ is the intersection, $\coprod$ the union, $A \Rightarrow B$ is the union of B with the set-complement of A, hence $\neg_{\mathrm{i}} A$ is the set-complement of A, and $A - B$ is the difference of subsets, i.e. the intersection of A with the set-complement of B, hence $\neg_{\mathrm{p}} A$ is again the set-complement of A. Therefore, the two negations coincide and the logic is classical: we have here a Boolean algebra. A very well-known result says that each Boolean algebra can be identified with a category of subsets and conversely. As the classical propositional calculus is equivalent to a Boolean algebra, we arrive at the well-known fact that subsets of a given set are the mathematical representatives of classical logic, whereas open subsets of a topological space represent intuitionist logic and closed subsets of a topological space represent paraconsistent logic. Moreover, we can remark that each set can be viewed as a topological space with all its subsets for open and closed subsets; we call this type of topology discrete topology. We have been able to see clearly see that categories of subsets of a given set are both categories of open subsets and categories of closed subsets of the set viewed as a discrete topological space; intuitionism and paraconsistency, coinciding in particular on their negations, form a classical logic because the first one brings the non-contradiction axiom and the second one brings the excluded middle axiom!

As for the categories of subgraphs of a given graph with their inclusions, we also have both Heyting and co-Heyting algebras but this time the two negations do not coincide. The categorical structure is as follows: 0 is the empty graph, 1 the whole given graph, $\times$ is intersection and $\coprod$ union, $A \Rightarrow B$ and $A - B$ are defined by their adjunction properties. As these properties are too complex too be defined in this paper, let us explore directly the induced negations. On the one hand, by the topos structure of the category of graphs, categories of subgraphs inherit of an intuitionist negation $\neg_{\mathrm{i}} A$ is the greatest subgraph whose intersection with A is empty, which ensures the validity of the non-contradiction axiom. On the other hand, the mathematical practice defines the pseudo-complement of a subgraph A as follows: it is the smallest subgraph whose union with A is equal to the whole given graph. We can show that it is, in fact, the paraconsistent negation $\neg_{\mathrm{p}} A$ of the category of subgraphs of a given graph, whose definition ensures the validity of the excluded middle axiom. However, in this case, these two negations do not coincide, which gives to the category of subgraphs of a given graph two distinct propositional calculi, one being intuitionist and the other paraconsistent. In order to better see this particular phenomenon of coexistence of two distinct propositional calculi, let us take the following example:

$\Gamma =$



Here the given graph is Γ, the subgraph A is the first arrow. The intuitionist negation $\neg_{\mathrm{i}} A$ of A, being the greatest subgraph whose intersection with A is empty, is then the vertex

2. The excluded middle axiom does not hold for this negation because the second arrow is missing in the union of A with 2 if we want this union to be equal to the whole graph Γ. The paraconsistent negation $\neg_p A$ of A is the second arrow because it is, in fact, the smallest subgraph whose union with A is the whole graph Γ. However, the vertex 1 is both an element of A and of its paraconsistent negation $\neg_p A$. Thus this contradiction is no longer false and at the same time the whole logic does not collapse because only the subgraphs containing 1 can be deduced from this contradiction; by example neither the empty graph nor the vertex 2 can be deduced from this contradiction.

The case of subgraphs is very interesting at two levels. On the one hand, it shows that in mathematical practice, the choice of intuitionist logic when classical logic does not hold is not so natural as we can think at first sight. Although graphs form a topos, it is a paraconsistent negation (pseudo-complement of a subgraph) that is used by mathematicians and not the intuitionist one induced by the topos structure. This example confirms the fact viewed in the preceding part that geometry can study either the intuitionist point of view of global–local or the paraconsistent point of view of polynomials. On the other hand, with this example the circle is complete because we started from categories, which can be viewed as graphs, in order to explore different mathematical universes and their internal logic and we have arrived at the fact that the logic of subgraphs is a very subtle one combining intuitionism and paraconsistency. All this confirms the complexity and variety of logics inhabited category theory, in contrast to the monotony of topos logics which are only intuitionist or classical (even if by playing on topology, we can obtain sheaves logics with interesting modal interpretations: localized logics, fuzzy logics, temporal logics ...[11]) This is why it seems very simplistic to us, in light of mathematical notions and practices, to consider that topoi with their intuitionist logics are the only possible mathematical universes. Therefore, we suggest to go further than Alain Badiou and view in category theory the basis of the logic of possible ontologies.

4 Conclusion

In Badiou's metaphysics mathematics plays a central role. It takes place in it as onto-logic, which means that it offers at the same time a decisional face and a formal vision. Ontology, as a decision over being, is represented inside mathematics, by set theory with its axiomatics, which bears on the undefined notion of set, sets being in turn the representatives of being as pure multiplicity. The logic of ontology, in Badiou's eyes, is represented by topos theory, conceived as the theory that explores the possible mathematical universes and which, by the way, makes explicit the link between the properties of a universe and the underlying logics. The fact that this theory shows that these universes are intuitionist, or in some cases classic, confirms the Aristotelian asymmetry holding between the axiom of non-contradiction, vital for logic, and the axiom of the excluded middle, which is purely optional for a logic.

[11] These logics of sheaves parameterized by topology are one the main themes explored in Alain Badiou's book *Logiques des mondes*.

Now, as we have tried to show in this paper, which we must admit is quite mathematical, the framework of topos theory corresponds only to the different possible set-theoretical universes, these universes having – because of their very axiomatics – an intuitionist internal logic. The theory that has been created in order to explore the different possible universes, namely the different possible mathematical structures, is category theory. It is the latter that enabled us to see, on the one hand, which ontological choices have been made for the topoi, inducing intuitionism on them; but also, on the other hand, that a dual choice exists, which leads to paraconsistent universes where contradiction exists without thereby trivialising the logic, and where it is the excluded middle that is vital. Therefore, the axioms of non-contradiction and of the excluded middle are formally symmetric, they are dual in the categorical meaning of the word, and the choice of taking either one or the other corresponds to mathematical practices of localisation and of polynomial equations, which are each one as mathematically natural as the other and which, most of all, are complementary.

This is why the categorical-theoretical working mathematician, as well as the reader of Alain Badiou, would lay category theory as the logic of the possible ontologies, for the topoi already constitute, among the different categorical universes, an ontological choice, one generalising the choice of sets. This position has the advantage of going farther than Badiou in the direction of the desacralisation of the Aristotelian axioms, given that category theory demonstrates the duality, the logical symmetry of non-contradiction and of the excluded middle. If we can, as the intuitionists, touch almost unpunishably the excluded middle, there is no reason, neither on the formal nor on the practical level of mathematics, for having to avoid touching, as the paraconsistents do touch, non-contradiction, provided that occurring contradictions do not lead to a trivialisation of the logic! In Badiou's sense, all this is only an ontological choice between the different possible ontologies, even if it is not the most easily tenable choice in front of the challenge of making a decision about being. However, in the end, is not the most comfortable choice, philosophically speaking, the one of the axiomatics of set theory, which makes out of the category of sets the best of the possible mathematical universes, in the sense of Leibniz?...

Finally, if category theory is really the logic of ontologies, then it is natural to ask ourselves the following question: for that given mathematical structure, what were the ontological choices, and what underlying logic derives from that? This work has just begun in mathematics and the notion of topos is only a first draft of the answer to that question. We are persuaded that the pursuit of fundamental research in this direction will make it possible not only to shed light on the current mathematical practices bearing on these structures, but will also have many implications for the conception of logic and, more generally, for philosophy.

References

1. Badiou, A.: Platon et/ou Aristote. Théorie des ensembles et théorie des topoi sous l'œil du philosophe. In: Panza M., Salanskis, J.M. (eds.) L'Objectivité mathématique. Masson, Paris (1995)
2. Badiou, A.: Court traité d'ontologie transitoire. Seuil, Paris (1998)

3. da Silva de Queiroz, G.: Sobre a dualidade entre intuicionismo e paraconsistência. PhD thesis, Brasil (1998)
4. Lambek J., Scott, P.J.: Introduction to higher order categorical logic. Cambridge Studies in Advanced Mathematics 7. Cambridge University Press, Cambridge (1986)
5. Lavendhomme, R.: Lieux du Sujet. Psychanalyse et Mathématiques. Seuil, Paris (2001)

R. Angot-Pélissier (✉)
Le Mans, France
e-mail: rpelliss@yahoo.com

Potentiality and Contradiction in Quantum Mechanics

Jonas R. Becker Arenhart and Décio Krause

Abstract Following J.-Y.Béziau in his pioneer work on nonstandard interpretations of the traditional square of opposition, we have applied the abstract structure of the square to study the relation of opposition between states in superposition in orthodox quantum mechanics in [1]. Our conclusion was that such states are *contraries* (i.e., both can be false, but both cannot be true), contradicting previous analyzes that have led to different results, such as those claiming that those states represent *contradictory* properties (i.e., they must have opposite truth values). In this chapter, we bring the issue once again into the center of the stage, but now discussing the metaphysical presuppositions which underlie each kind of analysis and which lead to each kind of result, discussing in particular the idea that superpositions represent potential contradictions. We shall argue that the analysis according to which states in superposition are contrary rather than contradictory is still more plausible.

Keywords Contradiction · Superposition · Potentiality · Contrariety · Opposition

Mathematics Subject Classification (2000) Primary 03B53 · Secondary 81P05

1 Introduction

J.-Y. Béziau has advanced the thesis that the square of opposition is a general framework that may be profitably employed for conceptual analysis (e.g., in [4]). Almost any kind of opposition between propositions may be profitably studied by the conceptual machinery furnished by the square, so that the proper relationships between the propositions in question may be brought to light and further analyzed. In this sense, the traditional opposition between Aristotelian categorical propositions is one of the many interpretations of the abstract structure of the square.

Bearing these multiple interpretations in sight, one of the possible uses of the square concerns application in the case of quantum superpositions. States in quantum mechanics such as the one describing the famous Schrödinger cat – which is in a superposition between the states “the cat is dead” and “the cat is alive” – present a challenge for our understanding which may be approached via the conceptual tools provided by the square.

Dedicated to Jean-Yves Béziau on his 50th birthday

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_8

According to some interpretations, such states represent contradictory properties of a system (for one such interpretation see, e.g., [6]). On the other hand, we have advanced the thesis that states such as “the cat is dead” and “the cat is alive” are contrary rather than contradictory (see [1, 2]).

To keep the paper self-contained, we now recall the traditional definitions of the oppositions which shall be employed in this paper and that are used in the discussions of applications of the square:

Contradiction: Propositions α and β are contradictory when both cannot be true and both cannot be false.
Contrariety: Propositions α and β are contrary when both cannot be true, but both can be false.
Subcontrariety: Propositions α and β are subcontraries when both can be true, but both cannot be false.
Subaltern: Proposition α is subaltern to proposition β if the truth of β implies the truth of α.

Subalterns are not in any sense in a relation of opposition, but rather a kind of implication. Anyway, they are part of the traditional discussions of the square and so are usually included in the nonstandard approaches to the square.

In this chapter we shall once again engage in the discussion by presenting some further differences between the two approaches to superposition, *viz.*, the one that considers them as contradictions and the one that considers them as contraries. In the next section we present in outline both approaches, the one which considers quantum states of a superposition as contradictory and as contrary, so that our discussion can be self-contained. In Sect. 3, we discuss the relation between contradiction and potential properties. It is said that the analysis of superposition as contradictory holds when such a state is thought as a superposition of potential properties, so we analyze this claim. In Sect. 4 we investigate whether a concept of potentiality can sit comfortably with a notion of contradiction. In view of the previous discussions, we conclude by defending the idea that contrariety is still a more adequate way to understand superpositions.

2 Contradictions and Contrariety in Superpositions

The proper understanding of superpositions is an open challenge typical of most of quantum mechanical conceptual innovations. Typically, to address the problem an interpretation of the theory is offered, and along with it, the hopes that the difficulties that are generated by superpositions in quantum mechanics get a proper explanation. As is well known, Dirac [9, p.12] claimed that superposition is not reducible to any classical notion, it is a *sui generis* feature of quantum mechanics: it lies behind most of the applications of the theory, and represents somehow the essence of the novelties brought by quantum mechanics.

In a recent paper, da Costa and de Ronde [6] proposed to deal with superpositions by adhering to a completely innovative approach to the superpositions and, as a result of such move, by adjusting the underlying logic with which we discuss such issues: according to their proposal, a superposition involves in general contradictory properties, and the underlying logic for the discussion of those issues is a paraconsistent logic. Indeed, it is usually said that paraconsistent logics deal with contradictions without allowing trivialization.[1]

To take an example that illustrates the main thesis, consider a spin-$\frac{1}{2}$ system which is in the state $|\uparrow_z\rangle$. Now, when we change the direction and consider the x axis, this state is in a superposition between $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$. Each of these states corresponds to a projection operator $|\uparrow_x\rangle\langle\uparrow_x|$ and $|\downarrow_x\rangle\langle\downarrow_x|$, respectively, with each projection representing a propriety of the system (in this case, "to have spin up in the x direction" and "to have spin down in the x direction," respectively). Now, according to da Costa and de Ronde [6, p. 848], these properties "which constitute the superposition and must be considered simultaneously are in general *contradictory properties*."

So, it seems that the idea is very simple. Two properties such as "to have spin up" and "to have spin down" (from now on, the context should make it clear that they are being taken in the same spatial direction), when their corresponding states are in a superposition, are said to be contradictory. Furthermore, they are not actual properties of the system, but rather possible or potential properties (more about this in the next section, see also [8]). Then, potentially, the system has contradictory properties (see da Costa and de Ronde *loc. cit.*). According to them, this fact must be dealt with by a paraconsistent approach, that is, we must adopt a paraconsistent logic as the underlying logic (see [6, Sect. 5]).

Obviously, if two properties are to be thought of as contradictory, then one must make clear what is the meaning of contradiction. We attempted such an explanation in [1]. Indeed, if "spin up" and "spin down," or else (taking into account Schrödinger's cat) "cat dead" and "cat alive" are to be understood as contradictory, then this contradictoriness must be spelled out. To spell out this fact by using the traditional definition of contradiction of the square of opposition results in the fact that such attribution of properties must always have well-defined truth values: recalling that contradictory propositions have opposed truth values, the system must have spin up *and* not have spin down, or else have spin down *and* not have spin up. However, this was obviously not the idea behind the claim that we must "consider simultaneously" those properties. It is simply impossible to take them simultaneously and still keep them as contradictories when "contradiction" is understood in the terms provided by the traditional definition of the square.

Then, an analysis through the square seems to provide for an obstacle for the consideration of contradictory properties in a superposition. What alternative do we have? One could, perhaps, insist in a paraconsistent approach along the following lines. Let us consider properties such as "to have spin up" and "to have spin down" as being somehow one *paraconsistent negation* of the other. Then, following the suggestion of the paraconsistent

[1] In a nutshell, in classical logic, in the presence of a contradiction any proposition whatever may be said derivable, and the resulting system is called *trivial*. In paraconsistent systems, on the other hand, even if expressions *formally representing* contradictions are derivable, not every proposition is also derivable, so that the system is not trivial. For the details see [7], and also the discussion in Béziau [3, 5].

approach, we could explain how both can be taken as true of a system. Indeed, when the system is in a superposition we would have something of the form $\alpha \wedge \neg\alpha$, a true contradiction. However, this move would not do. First of all, an expression such as $\alpha \wedge \neg\alpha$ is not a contradiction in a paraconsistent logic, strictly speaking. As philosophical analyzes have made it clear (see, e.g., [3] and [5]), paraconsistent negations represent subcontrariety, that is, when $\neg$ is taken as paraconsistent negation, α and $\neg\alpha$ are subcontraries, not really contradictories. This move would at best amount to a change in terminology (see also [2]). Second, subcontrariety requires that both propositions involved can be true but both cannot be false. So, by adopting this view, one would be obliged to accept that in a superposition, at least one of the properties corresponding to the states in superposition always hold. This, however, is still weaker than the requirement that both always be the case. Furthermore, it is not clear whether this assumption does not violate some form of no-go theorems that prohibit such kind of property attribution in quantum mechanics.

We have suggested an alternative route in [1]: to take common wisdom seriously and claim, following Dirac, that a superposition represents a new state of the system, one in which the system, as far as we know, does not have any of the properties involved. To explain how to account for superpositions according to this analysis, let us consider once again the case of the spin-$\frac{1}{2}$ system, as an electron in the $|\uparrow_z\rangle$ state. When we inquire about what happens in the x direction, then the system is in a superposition between $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$. Now, instead of allowing that the system has both associated properties $|\uparrow_x\rangle\langle\uparrow_x|$ and $|\downarrow_x\rangle\langle\downarrow_x|$, we say that the system does not have any of the properties. In this sense, it can be false that the system has "spin up" and it can be false that the system has "spin down." However, if it is the case that the system has one of the properties (e.g., spin up), then it does not have the other (spin down). This situation describes precisely *contrary* propositions. So, the case of such superpositions involves an opposition, but it is not contradiction, but rather contrariety.

Notice that this view does not rule out interpretations such as versions of the modal interpretation (see [11]). It could be the case that the system, even if it is in a superposition, *does* have one of the associated properties (modal interpretations, recall, break the eigenstate–eigenvalue link). However, as we mentioned, in this case, contrariety is still preserved, because only one of the properties is the case, while the other is not. In fact, modal interpretations seem to be incompatible with any kind of approach to superposition in which the states in superposition correspond to properties that must be taken simultaneously.

Now, even though this seems to be very plausible (to our minds, at least), the approach that considers that states in superposition are contrary does seem to take into account some assumptions about property instantiation that a paraconsistentist may deny. Indeed, de Ronde [8] has approached the subject following these lines. According to this line of reasoning, to consider that states in superposition are contraries but not contradictories involves assuming an orthodox metaphysical view, which includes the assumption that quantum mechanics describes entities and how they bear properties. The paraconsistent approach, on the other hand, should be understood as taking properties as potentialities, following a completely different metaphysical approach. We shall explore the different metaphysical views underlying the paraconsistent approach in the next two sections.

3 Potentiality and Contradiction

As we have mentioned by the end of the previous section, perhaps two distinct kinds of metaphysical assumptions underlie the two analyzes proposed for the case of quantum superposition. As we have suggested, there may be a different set of metaphysical presuppositions making the job in each case, and these presuppositions may well be incompatible. So, in order to make things clearer, in this section we shall discuss a little more the suggestion that distinct modes of being underlie each kind of analysis, mainly by trying to bring to light the idea of potentiality that accompanies the paraconsistent approach.

As da Costa and de Ronde [6] and de Ronde [8] have suggested, systems described by states in a superposition are such that they have only potentially or possibly the properties associated with each state. More than that, those properties, as we have already quoted, must be taken simultaneously and are thought of as contradictory. Now, the main question is: what are possible properties or potencies and how can they be contradictory?

Let us begin by exploring the idea of a contradiction in the potential realm. Once we admit that reality is divided in two spheres, the actual and the potential (or possible), both equally real, we may have contradictions in both. Given that superpositions are existent only in the potential realm, we may concern ourselves only with this case. The first point we shall raise concerns is terminology: possibility and potentiality are treated as synonymous, it seems, by the paraconsistent approach. However, "possible" here has two distinct senses. In the most straightforward sense, it is said of a proposition that it is possible, while in the intended sense we are discussing, it is said of a property that it is possible. It is obviously the second sense that is being used here: possibility regarding properties. However, in its traditional use, a possible property, also called a *modal property*, is understood as a property that an object *does not have, but could have*. It is a useful notion, for instance, in the metaphysical discussion concerning the numerical difference (or identity) of a statue and the piece of clay of which it is made: if modal properties are admitted as legitimate properties in this case, then the statue is different from the piece of clay. Clearly, the statute *could not* be squashed and still be the same, while the piece of clay *could* be squashed and still be the same (see [12]).

Now, if this analysis of a possible property is correct, then we are entitled to understand potential properties as metaphysicians have traditionally understood modal properties. However, if the properties corresponding to the states in a superposition are modal properties of the system, then there can be no contradiction between them. Indeed, once again consider the case of the electron in $|\uparrow_z\rangle$. When we are concerned with the x direction, the system is in a superposition between $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$, and both properties $|\uparrow_x\rangle\langle\uparrow_x|$ and $|\downarrow_x\rangle\langle\downarrow_x|$ are possible (modal) properties. However, if it is correct to associate possible properties with modal properties, then these are properties the system does not actually have. In this case, then, it is difficult to understand how they can be contradictory; both simply fail to be properties of the system, and as a bonus, our suggestion that they are merely contraries applies (recall Sect. 2).

So, modal properties, even though they may be related to potential properties, will not work. Let us leave modal properties behind for now and keep with the same line of inquiry, but now turning our attention to "potential property." It happens that one can advance a very similar argument. Traditionally, *potential* is understood as being in straight opposition to *actual* (more on this in the next section). However, once again, a potential property is one which is not actually possessed by the system (by definition). In a superposition, when we identify the properties corresponding to the states in the superposition as potential properties in this sense (i.e., as opposed to actual), then, those are properties *not* possessed by the system. Again, there is no contradiction in this case, but only mere contrariety.

Perhaps through these analyzes we are still considering the potential and the possible too closely related to actuality and actualization. What if we consider potentiality as a separated realm completely independent of any entity in which it exerts its actualization? Could this independence somehow help us in attributing some sense to the idea that a contradiction is real in this realm? It is this suggestion that we now investigate.

The first point of the suggestion seems to be that properties are potential all by themselves. In this sense, they are independent of their possible attribution to a system. That is, a property is not possible because it is the modal or potential property of any specific system, but rather it is an independent power, existent by itself. In this sense, powers like "spin up" and "spin down" are contradictory. But it is even more difficult to make sense of contradiction here.

As far as contradiction is defined by the square of opposition, or even in other contexts, it involves some kind of affirmation and some kind of negation. Also, it involves truth and falsehood. Both, broadly understood, are missing in the completely independent understanding of contradictory properties.

First of all, a property, taken by itself as a power (a real entity not actual), is not affirmed nor denied of anything. To take properties such as "to have spin up in the x direction" and "to have spin down in the x direction" by themselves does not affirm nor deny anything. To say "to have spin up in the x direction" is not even a statement, it is analogous to speak "green" or "red hair." To speak of a contradiction, it seems, one must have complete statements, where properties or relations are attributed to something. That is, one must have something like "spin up is measured in a given direction," or "Mary is red haired," otherwise there will be no occasion for truth and falsehood, and consequently, no occasion for a contradiction. So, the realm of the potential must also be a realm of attribution of properties to something if contradiction is to enter in it. However, this idea of attribution of properties seems to run counter the idea of a merely potential realm. On the other hand, the idea of a contradiction seems to require that we speak about truth and falsehood.

Second, perhaps we can make clearer the idea of contradictory properties by analyzing the formal approach to contradictions advanced by da Costa and de Ronde. In [6, p. 855] it is provided for a paraconsistent set theory ZF_1 in which superpositions are formalized. Let us consider a system S which is in a superposition of states s_1 and s_2, both "classically incompatible." A predicate symbol $K(S, s_1)$ is introduced in the language of ZF_1 by da Costa and de Ronde to represent the predicate that "S has the superposition predicate associated with s_1." The same reading holds with obvious adaptation for $K(S, s_2)$ and similarly for $\neg K(S, s_1)$ and $\neg K(S, s_2)$, where $\neg$ is a paraconsistent negation. Now, with the help of these predicate symbols the *Postulate of Contradiction* is introduced: when S

is in a superposition of s_1 and s_2, we have

$$K(S, s_1) \wedge \neg K(S, s_1) \wedge K(S, s_2) \wedge \neg K(S, s_2)\,.$$

This postulate somehow represents the situation in a superposition.

Now, let us apply this postulate to our former example of an electron (our system S) in $|\uparrow_z\rangle$ which in the x direction is in a superposition of the states $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$ (our states corresponding to s_1 and s_2). The Postulate of Contradiction now reads

$$K(S, |\uparrow_x\rangle) \wedge \neg K(S, |\uparrow_x\rangle) \wedge K(S, |\downarrow_x\rangle) \wedge \neg K(S, |\downarrow_x\rangle)\,.$$

But as we have remarked in [1], the contradiction now comes from the postulation of $K(S, |\uparrow_x\rangle) \wedge \neg K(S, |\uparrow_x\rangle)$, not from any relation between $|\uparrow_x\rangle\langle\uparrow_x|$ and $|\downarrow_x\rangle\langle\downarrow_x|$. Anyway, let us concede this point and accept that a superposition is inconsistent because it involves things like $K(S, |\uparrow_x\rangle) \wedge \neg K(S, |\uparrow_x\rangle)$. There are some difficulties with this move.

To begin with, this attribution of contradiction requires that there is a system S which both has "spin up" and does not have "spin up." This is clearly an actualist reading of property attribution, in the sense that the system has *actually* the property and *actually* does not have it. Indeed, this is a contradiction in the actual world. So, in this sense, the Postulate violates the requirement that the properties in a superposition be only potential in order to build a contradiction. This move clearly goes against the main motivation to consider the properties of a superposition as potential.

But what if we count such an attribution as merely potential? We could try to read $K(S, s_1)$ as "the system S potentially has the predicate associated with s_1." Now, an ambiguity enters the stage in the case of negation. If this suggestion is correct, when we try to read $\neg K(S, s_1)$ we have two options. The first one reads "the system S does not have potentially the predicate associated with s_1," understood as meaning that it is not potentially that the system has this property, it rather has it actually. The second reading is "it is not the case that the system S has potentially the predicate associated with s_1," that is, it is false that S is potentially s_1. Both readings are problematic.

The first reading is obviously troublesome when we consider the Postulate of Contradiction: it says that S has the predicates associated with s_1 and s_2 both potentially and actually. However, as the case of the electron illustrates, the system would have to be both "spin up" and "spin down" in the potential realm and in the actual realm, which, at least in the last case, is impossible. The second reading is also problematic for the approach. Indeed, since the motivation for approaching superpositions as contradictions was by considering properties in a superposition simultaneously, there seems to be no reason to consider them as holding simultaneously, even if potentially, and then deny that they hold simultaneously, even if potentially. That is, the introduction of a negation read as denying that the system has a given property potentially simply does not make sense if the idea was to represent a superposition as attributing both properties to the system (even if potentially). In this sense, the Postulate of Inconsistency does not seem to represent the intuition behind the paraconsistency approach to superpositions.

However, it may be the case that a paraconsistent set theory is just an inadequate formalism to capture the idea of a potential property, and a language involving an operator whose role is to represent potential property attribution could help us better in this task. Our next section shall investigate the prospects of this move.

4 Potentiality and Oppositions

What if the difficulties presented above come from distinct senses of "contradiction"? Perhaps the contradiction as represented in the traditional square is not the same as a contradiction for potential properties. In this section, we shall discuss a little more the idea of a potential property by employing the square of opposition. Throughout this section, we shall introduce a special operator $\Diamond$ to represent potentiality. In this sense, $\Diamond p$ means that "p is potential." Now, of course we must deal with p as representing potential property attribution *to something*, in general, to an already given system.

First of all, to establish the terminology, one could begin by distinguishing two distinct ways the operator may represent potential property attribution that could, both, represent the case of superposition. Let us suppose that s_1 and s_2 are in a superposition and let us concede in using s_1 and s_2 ambiguously both for the states as for the statements that the system is in the corresponding state. In a first reading of the situation, we could understand this situation as represented by the formula $\Diamond(s_1 \wedge s_2)$ or as $(\Diamond s_1 \wedge \Diamond s_2)$. We shall argue that only the second reading is a sensible reading of a superposition, and that it is difficult to understand any of both statements as contradictory.[2]

Consider first the second statement, $(\Diamond s_1 \wedge \Diamond s_2)$. If this is the intended meaning of the claim that a superposition involves potentiality, then clearly there is no contradiction in it. There is a simple analog in classical modal logic with the statement $(\Diamond p \wedge \Diamond\neg p)$. Notice that it is possible for some proposition to be the case and it is possible for it not to be the case, and this is different from the contradictory statement $(\Diamond p \wedge \neg\Diamond p)$. It is this last statement we would need to represent a contradiction, but this is clearly not the case in a superposition, that is, we do not have $(\Diamond s_1 \wedge \neg\Diamond s_1)$. Indeed, no one would claim that a state in a superposition stands for a potential property of the system and then try to represent such thing by including the information that it also *does not stand for a potential property of the system*. That move, we believe, renders the project of introducing potential properties in the discussion senseless.

However, one may complain that this reading of potentiality is not the intended meaning at all. This meaning, it could be argued, behaves too close to classical potentiality, it does represent at best the relation of a potential property coming to actuality, and is not the real quantum potentiality (see [8]), which is pure potentiality independently of actuality. In this case, it is the first reading of $\Diamond$, as it appears in $\Diamond(s_1 \wedge s_2)$ that represents a superposition. If this is the case, then some further difficulties arise.

[2] We have our doubts about the possibility of representing a superposition this way, in particular, in reading the $+$ sign of a superposition as a conjunction, but we shall do that for the purposes of argumentation in this section (see also [1]).

First of all, taken by itself, this does not represent a contradiction in the sense of the square. Indeed, one needs two statements in order to have such a contradiction. Perhaps the meaning of contradiction is different, as we have already suggested in the beginning of this section, in the sense that the paraconsistentist wishes that this formula be a logical contradiction, a formula that is always false (as defined in traditional logic textbooks). But what is the use in quantum mechanics of a formula that is always false? Furthermore, is it really the case that $\Diamond(s_1 \wedge s_2)$ is always false? Not really, in any standard version of normal modal logic this formula would represent a contingent statement, while in nonnormal modal logic it would be always true in the nonnormal worlds, for instance, due to the peculiar semantical understanding of $\Diamond$ in those worlds (see [10]). So, to tackle this issue seriously one must provide the axioms or rules for the operator $\Diamond$, which we have been understanding only informally till now.

However, instead of advancing a formal analysis of the operator $\Diamond$, we shall now investigate different meanings it could have and see how well they fare in relation to the idea that $\Diamond(s_1 \wedge s_2)$ is a potential contradiction, or, a contradiction in the potential realm. To fix our ideas, let us keep with the case of the electron whose state is a superposition of $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$. So, if the idea is that $\Diamond(s_1 \wedge s_2)$ represents a contradiction in the potential realm, in our example that means that potentially our system has both spin up and spin down. Obviously, no system has *actually* both spin up and spin down.

Now, this raises some important questions. First of all, what is the relation of the operator $\Diamond$ with actuality? Let us introduce for simplicity an operator @ for actuality. One could think at first that actuality and potentiality are mutually incompatible: when something is potential it is not actual, and when it is actual, then it is not potential. In this sense, the statements $\Diamond p$ and @p are *contradictory*, in the sense of the square. However, if this is a sensible reading, notice that even though $\Diamond$(spin up and spin down), we never have @(spin up and spin down). In this sense, the postulated contradiction in the potential realm never gets actualized, it does not work for actual entities. That leaves the contradiction in the potential realm unmotivated, and makes the reading of a superposition as ($\Diamond$ spin up and $\Diamond$ spin down) much more plausible, without the need for a contradiction, once again. That is, since the alleged contradiction is not doing any physical and any metaphysical work, we may plausibly leave it behind. Some people call it an application of Priest's razor, the metaphysical principle according to which we should not populate the world with contradictions beyond necessity (see [14] and [1] for a related discussion of this principle in the same context).

For a second possibility, let us consider that $\Diamond$ and @ are not contradictory, but are somehow compatible, in the sense that potential properties may also be actual and vice versa. In this sense, $\Diamond p$ and @p represent rather *subcontrary* statements, not contradictory statements. Indeed, according to this reading, any property must be either actual or potential (both cannot be false), but some properties can be both actual and potential (both can be true). Our main difficulty with this interpretation is once again the lack of motivation for introducing the claim that $\Diamond$(spin up and spin down) represents a contradiction and to prefer this reading of a superposition instead of ($\Diamond$ spin up and $\Diamond$ spin down), which is not contradictory in any sense. Because even though some simple properties may be understood as being both actual and potential, this is clearly not the case for quantum properties such as spin up and spin down: that is, even if one accepts that potentially the system can be both simultaneously, actually that never occurs. So, this reading would provide for a distinct understanding of the relation between $\Diamond$ and @ which saves the postulation of a contradiction in potentiality, but it still does not help motivating the postulation.

Obviously, our proposal is not that both $\Diamond p$ and $@p$ could both be false, generating a kind of limbo between actuality and potentiality, but rather that *if* one is going to concede that some kind of potentiality must be introduced to account for superpositions (notice the conditional), *then* perhaps the best way to understand that potentiality in relation to superposition is by leaving $\Diamond(s_1 \wedge s_2)$ behind and sticking to $(\Diamond s_1 \wedge \Diamond s_2)$. In this sense, we can still grant that a superposition represents a state in which not both properties are actual, that is, it is not the case that $(@s_1 \vee @s_2)$ must be necessarily true. The case is left open whether one or the other obtains, but the issue depends on some interpretational details that we shall present very briefly in the next section.

5 Conclusion: Contrariety Again

We hope we have made it clear that even though a paraconsistent approach to quantum superpositions is viable and defensible, it is still hard to see it as well motivated by the theory. Furthermore, there are some difficulties related with the very idea that a paraconsistent logic deals with contradictions *stricto sensu* as well as with the idea that superpositions are indeed contradictory (for the first difficulty, see again [3, 5] and for the second, see Sect. 2 and [1]).

Now, to consider that a superposition is better understood as a contrariety still leaves some issues open. It is a position compatible with very weak requirements on quantum mechanics. Let us make the issue clearer. We take it as a rather reasonable assumption that whenever a quantum system is in an eigenstate, then it really does have the property associated with the corresponding eigenvalue. This is a *Minimal Property Ascription Condition*, in a formulation taken from Muller and Saunders [13, p. 513]:

The minimal property attribution condition: If a system is in an eigenstate of an operator with eigenvalue **v**, then the system has the qualitative property corresponding to such value of the observable.

Notice that this puts a fairly weak condition for us to attribute properties to quantum systems: when in an eigenstate, we can surely say the system has the associated property. This is only half of the famous eigenstate–eigenvalue link. But what happens when the system is not in an eigenstate, when it is in a superposition? The condition is silent about that. One can complement the minimal condition in a variety of ways, for instance, by claiming that when not in an eigenstate the system does not have any of the properties associated with the superposition. This option is compatible with the claim that states in a superposition are contraries: both fail to be the case. Or instead of adopting this position, one can assume another interpretation, such as modal interpretations, and hold that even in a superposition one of the associated properties hold, even if not in an eigenstate (see [11]). Following this second option, notice, the understanding of superpositions as contraries still holds: even when one of the properties in a superposition holds, the other must not be the case.

So, the idea that states in a superposition are contrary rather than contradictory are compatible with a variety of interpretational moves. It is compatible with assuming only the minimal property attribution condition. The paraconsistent approach, on the other hand, introduces a further interpretational postulate, a kind of converse for the minimal condition according to which every superposition corresponds to properties attributed simultaneously to the system (see [1]):

Paraconsistent property attribution: When in a superposition, the system does have the properties related to the vectors forming the superposition, and they are contradictory.

This move, obviously, closes some alternative interpretation which are also plausible candidates. So, it seems, this is a further advantage of dealing with superpositions as contraries: one leaves open some important issues that are still hot issues of interpretation in quantum mechanics, while the paraconsistent property attribution seems to put *a priori* constraints on the theory and its future developments.

References

1. Arenhart, J.R.B., Krause, D.: Contradictions, Quantum Mechanics, and the Square of Oppositions. Preprint available at http://philsci-archive.pitt.edu/10582/ (2014)
2. Arenhart, J.R.B., Krause, D.: Oppositions and quantum mechanics: Superposition and identity. In: Béziau, J.-Y., Gan-Krzywoszynska, K. (eds.) New Dimensions of the Square of Opposition, Philosophia Verlag, Munich (2015), in press
3. Béziau, J.-Y.: New Light on the square of oppositions and its nameless corners. Log. Investig. **10**, 218–232 (2003)
4. Béziau, J.-Y.: The Power of the hexagon. Log. Univers. **6**(1–2), 1–43 (2012)
5. Béziau, J.-Y.: Contradiction and Negation. Talk delivered at the XVII EBL 2014 (2014)
6. da Costa, N.C.A., de Ronde, C.: The paraconsistent logic of superpositions. Found. Phys. **43**, 854–858 (2013)
7. da Costa, N.C.A., Krause, D., Bueno, O.: Paraconsistent logic and paraconsistency. In: Jacquette, D. (ed.) Philosophy of Logic, pp. 655–775. Elsevier, Amsterdam (2006)
8. de Ronde, C.: A Defense of the Paraconsistent Approach to Superpositions (Answer to Arenhart and Krause). Available at http://arxiv.org/abs/1404.5186 (2014)
9. Dirac, P.A.M.: The Principles of Quantum Mechanics, 4th edn. Oxford University Press, Oxford (2011)
10. Hughes, G.E., Cresswell, M.J.: A New Introduction to Modal Logic. Routledge, London (1996)
11. Lombardi, O., Dieks, D.: Modal interpretations of quantum mechanics. In: Zalta, E. (ed.) The Stanford Encyclopedia of Philosophy, Spring 2014 Edition. http://plato.stanford.edu/archives/spr2014/entries/qm-modal/ (2014)
12. Lowe, E.J.: A Survey of Metaphysics. Oxford University Press, Oxford (2002)
13. Muller, F.A., Saunders, S.: Discerning Fermions. Br. J. Philos. Sci. **59**, 499–548 (2008)
14. Priest, G.: In Contradiction. Nijhoff, Dordrecht, The Netherlands (1987)

J. R. Becker Arenhart (✉) · D. Krause
Department of Philosophy, Federal University of Santa Catarina, Florianópolis, Brazil
e-mail: jonas.becker2@gmail.com

Two, Many, and Differently Many

Diderik Batens

Abstract This paper is a modest contribution to a universal logic approach to many-valued semantic systems. The main focus is on the relation between such systems and two-valued ones. The matter is discussed for usual many-valued semantic systems. These turn out to exist for more logics than expected. A new type of many-valued semantics is devised and its use has been illustrated. The discussion that involves truth functionality and the syntactic rendering of truth-values leads to philosophical conclusions.

Keywords Many-valued semantics · Universal logic · Paraconsistency

Mathematics Subject Classification (2000) Primary 03A05 · Secondary 03B50

Like the volume in which it appears, this paper is a tribute to Jean-Yves Béziau. We met a long time ago on the side of those who fought then widespread prejudices on paraconsistency. Later, I admired his efforts for "universal logic" – papers, books, and the journal *Logica Universalis*. My paper is a modest contribution toward universal logic with a strong paraconsistent flavor.

1 Aim of This Chapter

The general aim of this chapter concerns the relation between two-valued and many-valued semantic characterizations of Tarski logics – reflexive, transitive and monotonic functions that map sets of closed formulas (the premises) to sets of closed formulas (the consequences). Suszko [38] has shown that logics (in this sense of the term) have a two-valued semantics.[1] This casts doubts on the use of "many-valued" as an attribute of logics. Moreover, it suggests that, at least for some logics, there must be an interesting relation between the values of their two-valued semantics and their many-valued semantics. We shall see in Sect. 3 that this relation is usually of a specific kind. In the present chapter, three-valued and four-valued logics will be introduced that are rather unusual. For example, the

[1] Throughout this chapter, "semantics of a logic **L**" should be read as "characteristic **L**-semantics."

The author is indebted to Joke Meheus for comments on a former draft.

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_9

relation between their values and the values from the two-valued semantic characterization will be very different from the similar relation for usual many-valued logics.

A two-valued semantic characterization of a (so-called) many-valued logic connects the logical symbols to truth preservation, as Suszko's proof underlines. We shall consider several relations that connect the values of the two-valued characterization to values of a many-valued characterization. Varying the relation brings one from the same two-valued semantics to different many-valued semantics and to different interpretations.

A new such relation will be introduced in this chapter. Its use will be illustrated in terms of a specific problem in adaptive logics, viz. the identification of flip–flop logics – Sect. 6. Readers that have no specific insight in adaptive logics should not worry; the flip–flop problem will be easy to understand. In comparison to **CL** (Classical Logic), some Tarski logics allow for gluts or gaps – see Sect. 2 – with respect to certain logical symbols. Delineating the precise points at which gluts or gaps originate offers important insights for solving the flip–flop problem. As we shall see, the delineation also provides an revealing interpretation of the nonextreme values.

Needless to say, the new type of many-valued logics are not meant to replace the usual ones. Also, they are not superior in any sense of the term. They are helpful, however, to reveal the presuppositions that lurk underneath widespread views on many-valued logics, often confusing technical features with ontological ones. Rather central presuppositions concern truth-functionality.

2 Preliminaries

A logic is defined over a language schema $\mathcal{L}$, of which $\mathcal{F}$ is the set of formulas and $\mathcal{W}$ is the set of closed formulas. I shall need some names for sets of schematic letters for nonlogical symbols: $\mathcal{S}$ (sentential letters), $\mathcal{P}^r$ (predicates of each rank $r \in \{1, 2, \ldots\}$), $\mathcal{C}$ (individual constants), and $\mathcal{V}$ (individual variables).

To handle quantifiers in the semantics, extend $\mathcal{L}$ to the pseudolanguage schema $\mathcal{L}_\mathcal{O}$. This is just like $\mathcal{L}$ except that the role of $\mathcal{C}$ is played by $\mathcal{C} \cup \mathcal{O}$, in which $\mathcal{O}$ is a set of pseudoconstants. Strictly speaking, we introduce a pseudolanguage schema $\mathcal{O}$ for each model M, requiring that $\mathcal{O}$ has at least the cardinality of the domain of M.[2] $\mathcal{F}_\mathcal{O}$ and $\mathcal{W}_\mathcal{O}$ are the sets of formulas and closed formulas of $\mathcal{L}_\mathcal{O}$, respectively.

A model is a tuple. One of the elements of the tuple is the domain D, which is a nonempty set. Another element is the assignment v, which relates certain linguistic entities to the other elements of the tuple. Next, for each model M, the meaning of the logical symbols of $\mathcal{L}$ is fixed by v_M, the valuation function determined by M.

In the usual **CL**-semantics, v relates the nonlogical symbols to the model. More specifically $v\colon \mathcal{S} \to \{0, 1\}$. In order to turn indeterministic two-valued semantic systems into deterministic ones – see below – I shall generalize this to $v\colon \mathcal{W}_\mathcal{O} \to \{0, 1\}$.[3]

[2] The pseudolanguage schema $\mathcal{L}_\mathcal{O}$ is not a language schema whenever its set of symbols is nondenumerable. The resulting style of semantics – examples follow in the text – offers a means to quantify over nondenumerable sets.

[3] This move is independent of the reference to the pseudolanguage schema $\mathcal{L}_\mathcal{O}$. To combine the move with a different semantic style, restrict it to $v\colon \mathcal{W} \to \{0, 1\}$.

In the case of the **CL**-semantics, the pseudolanguage schemas $\mathcal{L}_{\mathcal{O}}$ extend the standard predicative language schema $\mathcal{L}_s$ (which I do not describe here). A **CL**-model $M = \langle D, v\rangle$, in which D is a nonempty set and v an assignment function, is an interpretation of $\mathcal{W}_{\mathcal{O}}$. The assignment function v is defined by[4]

C1 $v\colon \mathcal{W}_{\mathcal{O}} \to \{0, 1\}$
C2 $v\colon \mathcal{C} \cup \mathcal{O} \to D$ (where $D = \{v(\alpha) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$)
C3 $v\colon \mathcal{P}^r \to \wp(D^r)$.

The valuation function determined by M, $v_M\colon \mathcal{W}_{\mathcal{O}} \to \{0, 1\}$, is defined as follows:

C$\mathcal{S}$ where $A \in \mathcal{S}$, $v_M(A) = v(A)$
C$\mathcal{P}^r$ $v_M(\pi^r\alpha_1, \ldots, \alpha_r) = 1$ iff $\langle v(\alpha_1), \ldots, v(\alpha_r)\rangle \in v(\pi^r)$
C= $v_M(\alpha = \beta) = 1$ iff $v(\alpha) = v(\beta)$
C¬ $v_M(\neg A) = 1$ iff $v_M(A) = 0$
C⊃ $v_M(A \supset B) = 1$ iff $v_M(A) = 0$ or $v_M(B) = 1$
C∧ $v_M(A \wedge B) = 1$ iff $v_M(A) = 1$ and $v_M(B) = 1$
C∨ $v_M(A \vee B) = 1$ iff $v_M(A) = 1$ or $v_M(B) = 1$
C≡ $v_M(A \equiv B) = 1$ iff $v_M(A) = v_M(B)$
C∀ $v_M(\forall\alpha A\,(\alpha)) = 1$ iff $\{v_M(A(\beta)) \mid \beta \in \mathcal{C} \cup \mathcal{O}\} = \{1\}$
C∃ $v_M(\exists\alpha A(\alpha)) = 1$ iff $1 \in \{v_M(A(\beta)) \mid \beta \in \mathcal{C} \cup \mathcal{O}\}$

$M \Vdash A$ (**CL**-model M verifies A) iff $v_M(A) = 1$. M is a model of Γ iff $M \Vdash A$ for all $A \in \Gamma$. $\Gamma \vDash_{\mathbf{CL}} A$ iff all models M of Γ verify A. $\vDash_{\mathbf{CL}} A$ (A is valid) iff all models verify A.

The metalanguage is classical and will be so in the rest of this chapter. All identities that occur in the semantics outside the expression $v_M(\alpha = \beta)$ are metalinguistic identities and hence are classical.

I still need to illustrate the use of clause C1. The logic **CLuN** is like **CL** except in that it allows for gl*u*ts with respect to *N*egation – in some **CLuN**-models M there are A such that $v_M(A) = v_M(\neg A) = 1$.[5] A semantics of **CLuN** is obtained from the **CL**-semantics by replacing the equivalence C¬ by an implication

C$\neg^{iu}$ if $v_M(A) = 0$, then $v_M(\neg A) = 1$.

The resulting logic is paraconsistent (e.g., $p, \neg p \nvDash_{\mathbf{CLuN}} q$),[6] invalidates Replacement of Equivalents (for example $\vDash_{\mathbf{CLuN}} p \equiv (p \vee p)$ and $\vDash_{\mathbf{CLuN}} \neg p \supset \neg p$ but $\nvDash_{\mathbf{CLuN}} \neg p \supset \neg(p \vee p)$), invalidates Replacement of Identicals (e.g., $a = b, \neg Pa \nvDash_{\mathbf{CLuN}} \neg Pb$), and invalidates many other rules validated by **CL** (e.g., Disjunctive Syllogism, Contraposition, Modus Tollens, ...).

This **CLuN**-semantics is indeterministic. Indeed, consider a **CLuN**-model $M = \langle D, v\rangle$ in which $v(p) = 1$ and hence $v_M(p) = 1$. In view of C$\neg^{iu}$, both $v_M(\neg p) = 0$ and

[4] The restriction in C2 ensures that $\langle D, v\rangle$ is only a **CL**-model if every element of D is named by a constant or pseudoconstant. In C3, $\wp(D^r)$ is the power set of the rth Cartesian product of D.
[5] The indeterministic propositional semantics was first formulated in [8]; the deterministic predicative semantics in [10].
[6] Technically speaking, a logic **L** is paraconsistent iff $A, \neg A \vdash_{\mathbf{L}} B$ does not hold generally. Interesting discussions of the underlying philosophical questions are available, for example, by Béziau [18, 19].

$v_M(\neg p) = 1$ are possible. To be more precise, the **CLuN**-semantics is bound to contain a copy of this M in which $v_M(\neg p) = 0$ and another copy in which $v_M(\neg p) = 1$. Both copies of M need to belong to the **CLuN**-semantics because a semantics is required to exhaust the logical possibilities – in this case **CLuN**-possibilities.

That a semantics is indeterministic is somewhat annoying. Models are supposed to exhaust the logical possibilities. That variants have to be taken into account – actually a nondenumerable set of variants for each model – introduces a complication that is not matched by any advantage.[7] Fortunately, it is possible to devise a deterministic semantics for **CLuN** [10] and the result was later generalized to other gluts and gaps – the best survey paper on the matter [15] is only electronically published at this moment. In order to obtain the deterministic **CLuN**-semantics one replaces C$\neg^{iu}$ by

C$\neg^{u}$ $\quad v_M(\neg A) = 1$ iff $v_M(A) = 0$ or $v(\neg A) = 1$.

The first disjunct guarantees that Excluded Middle holds, the second disjunct introduces gluts for some A.

This approach is easily generalized, first to gaps with respect to negation and to both gluts and gaps with respect to negation, and next to gluts and gaps with respect to other logical symbols – details are in another paper [15]. Just to give you the flavour, the logic **CLaN** allows for gaps (not gluts) with respect to negation. Its indeterministic semantics requires

C$\neg^{ia}$ $\quad$ if $v_M(A) = 1$, then $v_M(\neg A) = 0$

and its deterministic semantics is delivered by

C$\neg^{a}$ $\quad v_M(\neg A) = 1$ iff $v_M(A) = 0$ and $v(\neg A) = 1$.

CLoN allows for both gaps and gluts with respect to negation. Its indeterministic semantics is obtained by dropping the negation clause altogether. Its deterministic semantics is obtained by

C$\neg^{o}$ $\quad v_M(\neg A) = v(\neg A)$.

Restoring Replacement of Identicals in **CLuN** is easy. Given a **CLuN**-model M, define, for each $A \in \mathcal{W}_{\mathcal{O}}$, an equivalence class $[\![A]\!]$: (i) $A \in [\![A]\!]$ and (ii) if $\alpha, \beta \in \mathcal{C} \cup \mathcal{O}$, A is $B(\alpha)$, and $v(\beta) = v(\alpha)$, then $B(\beta) \in [\![A]\!]$. Note that $[\![A]\!] = \{A\}$ if $A \in \mathcal{S}$. Next, replace C$\neg^{u}$ by

C$\neg^{uR}$ $\quad v_M(\neg A) = 1$ iff $v_M(A) = 0$ or $v(\neg B) = 1$ for a $B \in [\![A]\!]$.

and analogously for C$\neg^{a}$ and C$\neg^{o}$. Note that the semantics characterizes the same logic if "for a $B \in [\![A]\!]$" is replaced by "for all $B \in [\![A]\!]$" – for every model of the one semantics there is a model of the other semantics that verifies exactly the same members of $\mathcal{W}$ – recall that verification depends on v_M and not on v.

The present approach to gluts and gaps leads to rather basic logics. Thus, the propositional fragment of **CLuN** was shown [32] to be the intersection of all propositional logics

[7] Still, indeterministic semantic systems, have been around at least since the 1970s and led to interesting studies, for example, in work by Arnon Avron and associates [5–7].

that allow for negation gluts but do not allow for any other gluts nor for any gaps. Obviously, some logics extend **CLuN** and are nevertheless paraconsistent, and similarly for other gluts and for gaps. A very popular paraconsistent extension of **CLuN** is a logic that I prefer to call **CLuNs** because its propositional version was first proposed by Kurt Schütte [36].[8] The idea is that de Morgan properties and all similar negation-reducing properties are restored. I consider at once a version that validates Replacement of Identicals. Let $\mathcal{F}^a_{\mathcal{O}}$ be the set of atomic (or primitive) members of $\mathcal{F}_{\mathcal{O}}$ – those not containing any logical symbols other than identity – and $\mathcal{W}^a_{\mathcal{O}}$ the set of atomic members of $\mathcal{W}_{\mathcal{O}}$. The deterministic **CLuNs**-semantics is obtained by replacing C$\neg$ in the **CL**-semantics by the following clauses:

C$\neg^{us}$ where $A \in \mathcal{W}^a_{\mathcal{O}}$, $v_M(\neg A) = 1$ iff $v_M(A) = 0$ or $v(\neg B) = 1$ for a $B \in [\![A]\!]$
C$\neg\neg$ $v_M(\neg\neg A) = v_M(A)$
C$\neg\supset$ $v_M(\neg(A \supset B)) = v_M(A \wedge \neg B)$
C$\neg\wedge$ $v_M(\neg(A \wedge B)) = v_M(\neg A \vee \neg B)$
C$\neg\vee$ $v_M(\neg(A \vee B)) = v_M(\neg A \wedge \neg B)$
C$\neg\equiv$ $v_M(\neg(A \equiv B)) = v_M((A \vee B) \wedge (\neg A \vee \neg B))$
C$\neg\forall$ $v_M(\neg\forall\alpha A(\alpha)) = v_M(\exists\alpha\neg A(\alpha))$
C$\neg\exists$ $v_M(\neg\exists\alpha A(\alpha)) = v_M(\forall\alpha\neg A(\alpha))$.

The **CLuNs**-semantics enables me to illustrate a method to restore Replacement of Identicals that does not refer to equivalence sets. Let v assign to predicates a couple of extensions rather than a single one: $v\colon \mathcal{P}^r \to \langle \Sigma_1, \Sigma_2\rangle$ with $\Sigma_1 \cup \Sigma_2 = D^r$. Identity is handled as a binary predicate with the special characteristic that $\Sigma_1 = \{\langle o, o\rangle \mid o \in D\}$. For all predicates, including identity, one defines $v^+(\pi^r) = \Sigma_1$ and $v^-(\pi^r) = \Sigma_2$. Finally, one replaces $v(\pi^r)$ by $v^+(\pi^r)$ in C$\mathcal{P}^r$, and one replaces C$\neg^{us}$ by two clauses:

- where $A \in \mathcal{S}$, $v_M(\neg A) = 1$ iff $v_M(A) = 0$ or $v(\neg A) = 1$
- $v_M(\neg\pi^r\alpha_1 \ldots \alpha_r) = 1$ iff $\langle v(\alpha_1), \ldots, v(\alpha_r)\rangle \in v^-(\pi^r)$.

This may be called the $\pm$-semantics of **CLuNs**.[9]

The words "glut" and "gap" were used rather intuitively until now. Actually, these words were used with several meanings in the literature. For example, Georg Henrik von Wright [40] says that there is an overlap (rather than glut) when a formula is true together with its negation and that there is a gap if a formula is false together with its negation. In the same place, von Wright calls a formula false iff its negation is true. So he also says that there is a glut (or overlap) if a formula is both true and false, and a gap if it is neither.

I shall use the terms differently. I already mentioned that the metalanguage of this chapter is fully classical. Unlike von Wright, I shall consider "true" and "false" as exhaustive and exclusive within a classical metalanguage. Finally, "glut" and "gap" will be used in a way that is directly contingent on the clauses of the **CL**-semantics. Any formula in $\mathcal{W}_{\mathcal{O}} - \mathcal{W}^a_{\mathcal{O}}$ has a specific logical form F, determined by its central logical term $*$. The **CL**-semantics contains a specific clause for F. This clause may be seen as the conjunction of

[8] **CLuNs** is apparently the most popular paraconsistent logic. It is known under a mutiplicity of names. Further useful references to studies of **CLuNs** and of its fragments are [2–4, 8, 16, 22, 24–29, 33, 34, 37]. Proofs of some results are in [13, 16].
[9] Another version of the approach, requiring only a single clause, is illustrated in a paper under review [39].

two implications, one specifying when $v_M(\mathsf{F}) = 1$ and one specifying when $v_M(\mathsf{F}) = 0$. Consider a model M of a logic **L**. If in M the condition is fulfilled for $v_M(A) = 1$ in **CL**-models but $v_M(A) = 0$, then this situation is said to cause a $*$-gap. If in M the condition is fulfilled for $v_M(A) = 0$ in **CL**-models but $v_M(A) = 1$, then this is said to cause a $*$-glut. Thus that $v_M(p \vee q) = v_M(\neg r) = 1$ and $v_M((p \vee q) \wedge \neg r) = 0$ causes a $\wedge$-gap. Similarly, that $1 \notin \{v_M(A(\beta)) \mid \beta \in \mathcal{C} \cup \mathcal{O}\}$ and $v_M(\exists \alpha A(\alpha)) = 1$ causes a $\exists$-glut.

Let us make this more precise. Define, for each $A \in \mathcal{W}_{\mathcal{O}}$, the set of *direct subformulas* of A, dsub(A), as follows: (i) dsub$(\neg A) = \{A\}$, (ii) where $* \in \{\vee, \wedge, \supset, \equiv\}$, dsub$(A * B) = \{A, B\}$, and (iii) where $\alpha \in \mathcal{V}$ and $* \in \{\forall, \exists\}$, dsub$(*\alpha A(\alpha)) = \{A(\beta) \mid \beta \in \mathcal{C} \cup \mathcal{O}\}$. If $A \notin \mathcal{W}^a_{\mathcal{O}}$, then, in the **CL**-semantics, $v_M(A)$ is a function of the valuation values of the members of dsub(A). In a semantics of a different logic, $v_M(A)$ will be said to be a glut or a gap if it departs from that function. Consider a **CLuN**-model in which $v_M(p) = 1$, $v_M(\neg p) = 1$ and $v_M(\neg\neg p) = 0$. The former two cause a $\neg$-glut. The combination of the first and third do not cause a gap, notwithstanding the fact that, within the **CL**-semantics, $v_M(\neg\neg A) = 1$ whenever $v_M(A) = 1$.

This seems the best point to mention a few simple technicalities. The set of *subformulas* of $A \in \mathcal{W}_{\mathcal{O}}$, sub$(A)$ is the smallest set such that (i) dsub$(A) \subseteq$ sub(A) and (ii) if $B \in$ sub(A), then sub$(B) \subseteq$ sub(A). Note that $A \notin$ sub(A).

We shall need the set of *first subformulas* of $A \in \mathcal{W}_{\mathcal{O}}$, fsub$(A)$. This is the smallest set such that (i) $A \in$ fsub$(\neg A)$, (ii) where $* \in \{\vee, \wedge, \supset, \equiv\}$, $A \in$ fsub$(A * B)$, (iii) where $\alpha \in \mathcal{V}$ and $* \in \{\forall, \exists\}$, $A(a) \in$ fsub$(*\alpha A(\alpha))$, and (iv) if $A \in$ fsub(B), then fsub$(A) \subseteq$ fsub(B).[10] The crucial distinction with the set sub(A) is in (ii) and (iii). We actually need fsub(A) to define the set of *first superformulas* of A, viz. fsup$(A) = \{B \in \mathcal{W}_{\mathcal{O}} \mid A \in$ fsub$(B)\}$.

Another concept we shall need is that of a (finite) *pseudopartition*. $\{\Sigma_1, \ldots, \Sigma_n\}$ is a pseudopartition of Σ iff (i) $\Sigma_i \cap \Sigma_j = \emptyset$ for all different $i, j \in \{1, \ldots, n\}$ and (ii) $\bigcup\{\Sigma_1, \ldots, \Sigma_n\} = \Sigma$ (but it is not required that the members of the pseudopartition are nonempty).

3 The Usual Many-Valued Approach

The **CL**-semantics is deterministic: the valuation value of every formula is determined by the model; in the presence of $\mathcal{L}_{\mathcal{O}}$ it is determined by the assignment values of the nonlogical symbols that occur in the formula and its subformulas. The **CL**-semantics is also recursive: there is a complexity function such that, for every nonatomic formula $A \in \mathcal{W}_{\mathcal{O}}$, $v_M(A)$ depends only on valuation values of formulas that are less complex than A. The **CL**-semantics is also truth-functional: there is a function that connects the valuation value of every nonatomic formula $A \in \mathcal{W}_{\mathcal{O}}$ to the valuation values of subformulas of A; only the value of atomic formulas is directly determined by assignment values.

[10] In (iii), a is the alphabetically first individual constant, which is used here as a metalinguistic name of itself.

A presupposition of this type of semantics is that a distinction can be made between two things. On the one hand, there is the model itself: $M = \langle D, v\rangle$. This represents a state of the world. All nonlogical symbols receive their meaning here – the assignment assures that they do. On the other hand, there is the realm of the logical symbols. These are required to formulate statements about the world. Indeed, by merely concatenating atomic statements one cannot express that an atomic statement is false or that one of two atomic statements is true but not necessarily both. Still, truth-functionality makes complex statements parasitic on atomic statements in that the valuation values of the complex statements are fully determined by the valuation values of atomic statements.

To be sure, a semantics is defined with respect to a language schema. In this sense, its models represent at best structural states of the world. In order to transform the models into representations of actual states of the world, one needs to replace the language schema by a language that is covered by the schema. So a semantics involves a hypothesis about the structure of the language in which the world, or some parts or aspects of it, may be adequately described. Needless to say, adequacy comes in degrees and the estimated adequacy may be poor due to the present state of our knowledge.[11]

When many-valued logics came around, new valuation values were added next to "true" and "false." There were two intuitions behind the new values. One idea was that some sentences do not have a truth-value, but are indeterminate in one of several senses. According to the other idea there are further truth-values, which are sometimes seen as expressing degrees of partial truth. The logical symbols of those logics were still truth-functions, namely, with respect to the extended set of valuation values. Apparently, the architects of many-valued logics first had the idea of additional "truth-values" and next devised truth-functional operators in terms of them. They apparently did not imagine, and possibly could not imagine, that a logical symbol would not be truth-functional.[12]

Within a many-valued semantics, semantic consequence is defined in terms of designated and nondesignated values. This shows the way from the many-valued semantics to the two-valued one. It is instructive to consider also the opposite road. This road was explored a long time ago for the propositional case, among others by me [9], and the generalization to the predicative level is obvious. The idea is that bivalent values of several formulas are "melted together" into a many-valued value of a single formula. The approach works fine for some paraconsistent logics, for example, for **CLuNs**. Consider, for any $A \in \mathcal{W}_{\mathcal{O}}$, the couple $\langle v_M(A), v_M(\neg A)\rangle$ in the above two-valued semantics. The possible couples are $\langle 1, 0\rangle$, $\langle 1, 1\rangle$, and $\langle 0, 1\rangle$. These may be handled as three valuation values and, if they are so handled, it is convenient to rename them to T, I, and F, which correspond to "consistently true," "inconsistent" and "consistently false" respectively. The resulting three-valued semantics is truth-functional, as I show below. The matter is utterly simple for the propositional case. For predicative models, the easiest approach requires that the assignment is redefined, for example, as in the next paragraph.

[11] Carnap [20, 21] clearly saw the linguistic relativity of the semantic enterprise. Apparently, many have forgotten his insight and seem to presume that they can talk about states of the world in an absolute way.
[12] I do not intend to refer, for example, to a worlds semantics but rather to a nontruth-functional semantics such as the ones from Sect. 2, the indeterministic as well as the deterministic ones.

In a three-valued **CLuNs**-model $M = \langle D, V \rangle$, defined over the language $\mathcal{L}_{\mathcal{O}}$, the domain D is a set and the assignment V has the following four properties. (i) $V : \mathcal{S} \rightarrow \{T, I, F\}$.[13] (ii) $V : \mathcal{C} \cup \mathcal{O} \rightarrow D$ (where $D = \{V(\alpha) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$). (iii) $V : \mathcal{P}^r \rightarrow \langle \Sigma_1, \Sigma_2, \Sigma_3 \rangle$ such that $\{\Sigma_1, \Sigma_2, \Sigma_3\}$ is a pseudopartition of $\wp(D^r)$. To simplify the notation, consider V as composed in this case of the three functions V^T, V^I, and V^F, with $V^T(\pi^r) = \Sigma_1$, $V^I(\pi^r) = \Sigma_2$, and $V^F(\pi^r) = \Sigma_3$.[14] (iv) Identity is handled as a binary predicate with the special characteristic that $V^T(=) \cup V^I(=) = \{\langle o, o \rangle \mid o \in D\}$.

The valuation function $V_M : \mathcal{W}_{\mathcal{O}} \rightarrow \{T, I, F\}$ is defined as follows:

C$\mathcal{S}$ where $A \in \mathcal{S}$, $V_M(A) = V(A)$

C$\mathcal{P}^r$ for $X \in \{T, I, F\}$, $V_M(\pi^r \alpha_1 \ldots \alpha_r) = X$ iff $\langle V(\alpha_1), \ldots, V(\alpha_r) \rangle \in V^X(\pi^r)$

Ccon

	$\neg$
T	F
I	I
F	T

$\supset$	T	I	F
T	T	I	F
I	T	I	F
F	T	T	T

$\wedge$	T	I	F
T	T	I	F
I	I	I	F
F	F	F	F

C$\forall$ $V_M((\forall \alpha)A(\alpha)) = T$ iff $V_M(A(\beta)) = T$ for all $\beta \in \mathcal{C} \cup \mathcal{O}$
$V_M((\forall \alpha)A(\alpha)) = F$ iff $V_M(A(\beta)) = F$ for at least one $\beta \in \mathcal{C} \cup \mathcal{O}$
$V_M((\forall \alpha)A(\alpha)) = I$ otherwise.

Define $M \Vdash A$ (a three-valued **CLuNs**-model M verifies A) iff $V_M(A) \in \{T, I\}$; and so on.

The other logical symbols are defined explicitly: $A \vee B =_{df} \neg(\neg A \wedge \neg B)$, $A \equiv B =_{df} (A \supset B) \wedge (B \supset A)$, and $(\exists \alpha)A(\alpha) =_{df} \neg(\forall \alpha)\neg A(\alpha)$.

This three-valued **CLuNs**-semantics is equivalent to the two-valued **CLuNs**-semantics from Sect. 2 in that their semantic consequence relations coincide – this is easily shown by slightly modifying the proof of Theorem 1 in [16].

So this is the usual approach to many-valued logics. There is a n-tuple of functions $\langle f_1, \ldots, f_n \rangle$ such that $f_i : \mathcal{W}_{\mathcal{O}} \rightarrow \mathcal{W}_{\mathcal{O}}$ for each i; the n-tuple of bivalent values $\langle v_M(f_1(A)), \ldots, v_M(f_n(A)) \rangle$ functions as the many-valued value $V_M(A)$. Expressed somewhat crudely, the (bivalent) values of formulas containing A are pushed into the (many-valued) value of A.

The attractiveness of the approach seems related to the fact that the many-valued valuation values appear to be a kind of truth-values that are more sophisticated than the bivalent valuation values and that are introduced for sound philosophical reasons. The view on logical symbols is simply the traditional view: they are truth functions. They differ from the classical logical symbols as a result of the modified set of valuation values.

Let us proceed more carefully. We already knew that a logic that has a many-valued semantics also has a two-valued semantics. It seems obvious that any many-valued semantics can be described as obtained by pushing the bivalent values of a tuple of formulas into the many-valued value of an atomic formula. From a technical point of view, the two semantics are on a par. Still, there is the philosophical question which semantics is ontologically correct. Are there many truth-values or are these merely tuples of binary

[13] There is no need to assign a three-valued assignment value to all members of $\mathcal{W}_{\mathcal{O}}$.

[14] The three functions determine for which r-tuples the predicate is true, inconsistent, and false respectively.

truth-values? Consider again **CLuNs**. Is the truth of $\neg A$ a consequence of the fact that A has the truth-value I or is saying that A has the value I merely a statement summarizing that A and $\neg A$ are both true?

The truth-values of the bivalent **CL**-semantics may be seen as "expressed" within the standard **CL**-language by A and $\neg A$, respectively. Similarly, one may (explicitly) define n logical symbols $\mathsf{V}_1, \ldots, \mathsf{V}_n$ within the language of Łukasiewicz's n-valued logics $\mathbf{Ł}_n$ such that, for all $i \in \{1, \ldots, n\}$, $M \Vdash \mathsf{V}_i A$ iff $v_M(A) = i$ – see [35, 39]. It is not possible to do so for **CLuNs**; a definable symbol does so correspond to the semantic value I, but no definable logical symbol so corresponds to T or to F.

Some will see this as an argument to consider the truth-values of the two-valued **CL**-semantics and those of the n-valued $\mathbf{Ł}_n$-semantics as real truth-values, at least with respect to the presuppositions of those logics, but will not consider the valuation values of the three-valued **CLuNs**-semantics as real truth-values.[15] They might argue that, if statements may have three different truth-values, then the logical symbols of your language should enable you to express, for each of the truth-values, that it pertains to a statement. Whether you may get to *know* the actual truth-value of a statement is altogether a different matter.[16]

Others will be less demanding and consider the fact that a logic has an adequate many-valued semantics in which all its logical symbols are truth-functions as a sufficient reason to consider those values as truth-values. The view apparently presupposes that some logics do not have such a semantics, but is that correct?

It seems unlikely a priori that **CLuN** has an adequate semantics in which all logical symbols are truth-functions. Apart from some transparent exceptions, such as $(p \wedge \neg p) \wedge \neg(p \wedge \neg p) \vdash_{\mathbf{CLuN}} p \wedge \neg p$, inconsistencies seems to be independent of each other within **CLuN**. Consider, for example, the set $\Delta = \{\neg p, \neg(p \wedge p), \neg(p \wedge (p \wedge p)), \neg(p \wedge (p \wedge (p \wedge p))), \ldots\}$. Each member of Δ is **CL**-equivalent to $\neg p$. However, for every $\Delta' \subset \Delta$, there is a **CLuN**-model M such that $M \Vdash A$ for all $A \in \{p\} \cup \Delta'$, whereas $M \nVdash A$ for all $A \in \Delta - \Delta'$. And there is more. Let $\mathcal{W}_p$ comprise the formulas in which no other nonlogical term than p occurs. For every $\Delta' \subseteq \Delta$, infinitely many members of $\mathcal{W}_p$ are verified by some **CLuN**-models of $\{p\} \cup \Delta'$ and falsified by other **CLuN**-models of $\{p\} \cup \Delta'$. Here are some examples of members of $\mathcal{W}_p$ for which this holds: $\neg B$ for any $B \in \Delta'$; $\neg((p \wedge p) \wedge p)$ and all similar results of commuting two different conjuncts in a conjunctive subformula of a $B \in \Delta'$; all $\neg(p \wedge B) \in \mathcal{W}_p - \Delta'$ such that the considered models verify B; and so on. Notwithstanding all this, all logical symbols are truth-functions in an infinite-valued **CLuN**-semantics. This is shown in Sect. 4.

If even the negation of **CLuN** is a truth-function in a many-valued semantics, one wonders whether being a truth-functional logic (with respect to some many-valued semantics) is a distinctive feature and, if it were distinctive, whether there is anything interesting about it. According to a truth-functional logic, the truth-value of every formula is fully determined by the truth-value of its atomic subformulas. This hardly means anything if it

[15] Similarly for the valuation values of the three-valued **LP**-semantics [34] and for the valuation values of four-valued semantics for relevant logics [1].

[16] As was already pointed out by Viktor Kraft [31], nothing warrants that the syntactically atomic sentences of a language are also epistemologically atomic.

is realized by coding the binary valuation value of every complex formula into the many-valued valuation value of an atomic formula.

Independent of philosophical worries concerning truth-functionality, one may wonder whether many-valued logics may not be cast in a way that is different from the usual approach. I shall be looking for an unusual mould in Sect. 5.

4 The Case of the Logic CLuN

In this section, it is shown that even **CLuN** is a many-valued logic on the usual many-valued approach. It seems instructive to describe the way in which this result was obtained, both in order to make the argument transparent and in order to illustrate the way in which the result may be extended to other logics.

The usual approach requires that all information required to fix the valuation value of a formula A is contained in the valuation value of the atomic subformulas of A. It follows that infinitely many bits of information have to be contained in a single valuation value. Recall indeed that the following holds for the two-valued **CLuN**-semantics: if $v_M(p) = 1$, then $v_M(\neg p)$ may be 1 or 0, depending on the value of $v(\neg p)$; if $v_M(p) = v_M(\neg p) = 1$, then the same applies to $v_M(\neg\neg p)$; and so on. So the information contained in $v(p), v(\neg p), v(\neg\neg p), \ldots$ in a two-valued **CLuN**-model needs to be compressed in the valuation value of p within the corresponding many-valued **CLuN**-model – I shall write this valuation value as $V_M(p)$.[17] The situation is even more complex. If $v_M(p \wedge q) = 1$, then $v_M(\neg(p \wedge q))$ may be 1 as well as 0, depending on the value of $v(\neg(p \wedge q))$, and this information too must be contained in $V_M(p)$ or in $V_M(q)$. Note also that, in the considered case, $v(\neg(p \wedge q))$ and $v(\neg(p \wedge r))$ are not only independent of each other, but also of $v(\neg p)$, $v(\neg q)$, and $v(\neg r)$.

So, if the approach can be applied to **CLuN**, then the valuation value $V_M(A)$ contains all information contained in an infinite list $\langle v_M(A), v(B^1), v(B^2), \ldots\rangle$ in which $B^1, B^2, \ldots$ are formulas of which A is a subformula. In view of this, it seems fitting to identify $V_M(A)$ with an infinite sequence of 0s and 1s. In order for the approach to be viable, several difficulties have to be resolved.

Consider $V_M(\neg(p \wedge q))$. This should be a truth-function of $V_M(p)$ and of $V_M(q)$. So, in terms of the two-valued semantics, the information on $v(\neg(p \wedge q))$ needs to be contained in $V_M(p)$ or $V_M(q)$. For every binary logical term $*$, I shall store the information on $v(\neg(A * B))$ in $V_M(A)$. This is obviously a conventional matter and there are several alternatives. The information on the two-valued assignment value $v(\neg\forall x P x)$ should also be contained within the many-valued valuation value of an atomic formula. I shall store it in $V_M(Pa)$. In view of these conventions, I defined (at the end of Sect. 2), for every $A \in \mathcal{W}_O$, the set fsub(A) of "first subformulas" of A and the set fsup(A) of "first superformulas" of A. For

[17] In the text I use the name M for both models although they are not only different but even different in kind. Where it matters, I shall obviously introduce different names.

every $B \in \text{fsup}(A)$, $v(B)$ (from the two-valued model) will be stored in $V_M(A)$ (from the many-valued model).[18]

Fact 1 For all $A \in \mathcal{W}_{\mathcal{O}}$, $\text{fsub}(A)$ is finite and decidable.

Fact 2 For all $A \in \mathcal{W}_{\mathcal{O}}$, there is exactly one $B \in \text{fsub}(A) \cap \mathcal{W}^a_{\mathcal{O}}$.

Fact 3 If $A, B \in \mathcal{W}^a_{\mathcal{O}}$ are different, then $\text{fsup}(A) \cap \text{fsup}(B) = \emptyset$.

Fact 4 If $A \in \text{fsub}(B)$, then $\text{fsup}(B) \subseteq \text{fsup}(A)$.

Fact 5 If $A \notin \text{fsub}(B)$ and $B \notin \text{fsub}(A)$, then $\text{fsup}(A) \cap \text{fsup}(B) = \emptyset$.

So the present state of our plot is to identify a many-valued valuation value with $\langle v_M(A), v(B^1), v(B^2), \ldots \rangle$, where $\langle B_1, B_2, \ldots \rangle$ is an ordering of $\text{fsup}(A)$. However, notwithstanding Fact 3, $\text{fsup}(A)$ is uncountable if $\mathcal{L}_{\mathcal{O}}$ is uncountable and this is always the case if model M is uncountable. In that case, however, there is no list $\langle v_M(A), v(B^1), v(B^2), \ldots \rangle$. Fortunately, the Löwenheim–Skolem theorem enables one to restrict a semantics to its countable models. So let us do this and consider only countable pseudolanguage schemas $\mathcal{L}_{\mathcal{O}}$ in the sequel of the present section.[19]

Fact 6 For all $A \in \mathcal{W}_{\mathcal{O}}$, $\text{fsup}(A)$ is infinite, enumerable, and decidable.

Consider a Gödel numbering and let $G(A)$ be the Gödel number of A. Let $L_A = \langle B_1, B_2, \ldots \rangle$ be such that (i) $\{B_1, B_2, \ldots\} = \text{fsup}(A)$ and (ii) $i \leq j$ iff $G(B_i) \leq G(B_j)$. Where $B \in \text{fsup}(A)$, let $\#_A(B)$ be the place of B in L_A – so if $L_A = \langle B_1, B_2, \ldots \rangle$, then $\#_A(B_i) = i$.

Fact 7 For all $A \in \mathcal{W}_{\mathcal{O}}$, L_A is a recursive list.

Fact 8 If $B \in \text{fsup}(A)$, then all members of L_B occur in the same order in L_A.

Fact 9 If $B \in \text{fsup}(A)$, then there is a computable function f such that $L_B = f(L_A)$.

Let $\mathfrak{S}$, the set of valuation values, be the set of infinite sequences of 0s and 1s; let $\mathfrak{s}$ be a variable for members of $\mathfrak{S}$; let $\mathfrak{S}_D = \{\langle d_0, d_1, \ldots \rangle \in \mathfrak{S} \mid d_0 = 1\}$ (the set of designated values). Where $V_M(A) = \langle d_0, d_1, \ldots \rangle$, define $V_M(A)[A] = d_0$ and define, for all $B \in \text{fsup}(A)$, $V_M(A)[B] = d_{\#_A(B)}$.[20]

[18] If B does not have the form $\neg C$, then $v(B)$ does not play any role within the **CLuN**-semantics. So one might just as well decide not to store the value of $v(B)$ in $V_M(A)$ for such $B \in \text{fsup}(A)$. While the disadvantage of the approach followed in the text is that some digits of $V_M(A)$ are irrelevant, the advantage is that the approach is more general, as is the case for the assignment function of the two-valued semantics itself. That the advantage outweighs the disadvantage will be shown in Sect. 5.

[19] These are actually language schemas. Still $\mathcal{L}_{\mathcal{O}}$ need to be different from $\mathcal{L}$ in order to allow for models that are not ω-complete.

[20] So, if $L_A = \langle B_1, B_2, \ldots \rangle$ and $V_M(A) = \langle 1011 \ldots \rangle$, then $V_M(A)$ contains the information that in the corresponding two-valued model M' holds: $v_{M'}(A) = 1$, $v(B_1) = 0$, $v(B_2) = 1$, $v(B_3) = 1$, and so on.

In a $\mathfrak{S}$-valued **CLuN**-model $M = \langle D, V\rangle$ (defined over the countable pseudolanguage schema $\mathcal{L}_{\mathcal{O}}$), the domain D is a countable set and the assignment V has the following four properties. (i) $V : S \to \mathfrak{S}$. (ii) $V : C \cup \mathcal{O} \to D$ (where $D = \{V(\alpha) \mid \alpha \in C \cup \mathcal{O}\}$). (iii) Where $\mathfrak{S} = \{\mathfrak{s}^1, \mathfrak{s}^2, \ldots\}$, $V : \mathcal{P}^r \to \langle \Sigma_{\mathfrak{s}^1}, \Sigma_{\mathfrak{s}^2}, \ldots\rangle$ such that (a) $\{\Sigma_{\mathfrak{s}^1}, \Sigma_{\mathfrak{s}^2}, \ldots\}$ is a pseudopartition of $\wp((C \cup \mathcal{O})^r)$ and (b) if $V(\beta) = V(\alpha_i)$ ($1 \leq i \leq r$), $\langle \alpha_1, \ldots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \ldots, \alpha_r\rangle \in \Sigma_{\mathfrak{s}^j}$, and $\langle \alpha_1, \ldots, \alpha_{i-1}, \beta, \alpha_{i+1}, \ldots, \alpha_r\rangle \in \Sigma_{\mathfrak{s}^k}$, then $\mathfrak{s}^j, \mathfrak{s}^k \in \mathfrak{S}_D$ or $\mathfrak{s}^j, \mathfrak{s}^k \in \mathfrak{S} - \mathfrak{S}_D$. To simplify the notation, consider V as composed in this case of uncountably many functions $V^{\mathfrak{s}^1}$, $V^{\mathfrak{s}^2}$, ..., with $V^{\mathfrak{s}^1}(\pi^r) = \Sigma_{\mathfrak{s}^1}$, $V^{\mathfrak{s}^2}(\pi^r) = \Sigma_{\mathfrak{s}^2}$, and so on. (iv) Identity is handled as a binary predicate with the special characteristic that $\bigcup\{V^{\mathfrak{s}}(=) \mid \mathfrak{s} \in \mathfrak{S}_D\} = \{\langle \alpha, \beta\rangle \mid \alpha, \beta \in C \cup \mathcal{O}; V(\alpha) = V(\beta)\}$. The valuation function $V_M : \mathcal{W}_{\mathcal{O}} \to \mathfrak{S}$ is defined as follows:

CS where $A \in S$, $V_M(A) = V(A)$
C$\mathcal{P}^r$ $V_M(\pi^r\alpha_1 \ldots \alpha_r) = \mathfrak{s}$ iff $\langle \alpha_1, \ldots, \alpha_r\rangle \in V^{\mathfrak{s}}(\pi^r)$
C$\neg$ Where $L_{\neg A} = \langle C^1, C^2, \ldots\rangle$, $V_M(\neg A) = \langle \max(1 - V_M(A)[A], V_M(A)[\neg A]), V_M(A)[C^1], V_M(A)[C^2], \ldots\rangle$.
C$\wedge$ Where $L_{A\wedge B} = \langle C^1, C^2, \ldots\rangle$, $V_M(A \wedge B) = \langle \min(V_M(A)[A], V_M(B)[B]), V_M(A)[C^1], V_M(A)[C^2], \ldots\rangle$.
C$\vee$ Where $L_{A\vee B} = \langle C^1, C^2, \ldots\rangle$, $V_M(A \vee B) = \langle \max(V_M(A)[A], V_M(B)[B]), V_M(A)[C^1], V_M(A)[C^2], \ldots\rangle$.
C$\supset$ Where $L_{A\supset B} = \langle C^1, C^2, \ldots\rangle$, $V_M(A \supset B) = \langle \max(1 - V_M(A)[A], V_M(B)[B]), V_M(A)[C^1], V_M(A)[C^2], \ldots\rangle$.
C$\equiv$ Where $L_{A\equiv B} = \langle C^1, C^2, \ldots\rangle$, $V_M(A \equiv B) = \langle \min(\max(1 - V_M(A)[A], V_M(B)[B]), \max(1 - V_M(B)[B], V_M(A)[A])), V_M(A)[C^1], V_M(A)[C^2], \ldots\rangle$.
C$\forall$ Where $L_{\forall\alpha A(\alpha)} = \langle C^1, C^2, \ldots\rangle$, $V_M(\forall\alpha A(\alpha)) = \langle \min\{V_M(A(\beta)) \mid \beta \in C \cup \mathcal{O}\}, V_M(A(a))[C^1], V_M(A(a))[C^2], \ldots\rangle$.
C$\exists$ Where $L_{\exists\alpha A(\alpha)} = \langle C^1, C^2, \ldots\rangle$, $V_M(\exists\alpha A(\alpha)) = \langle \max\{V_M(A(\beta)) \mid \beta \in C \cup \mathcal{O}\}, V_M(A(a))[C^1], V_M(A(a))[C^2], \ldots\rangle$.

Define $M \Vdash A$ (a $\mathfrak{S}$-valued **CLuN**-model M verifies A) iff $V_M(A) \in \mathfrak{S}_D$; and so on. To avoid confusion, let $\Gamma \vDash^{\mathfrak{S}}_{\mathbf{CLuN}} A$ denote that A is a **CLuN**-semantic consequence of Γ on the $\mathfrak{S}$-valued semantics.

Fact 10 The $\mathfrak{S}$-valued **CLuN**-semantics is recursive.

Fact 11 All logical symbols are truth-functions in the $\mathfrak{S}$-valued **CLuN**-semantics.

Lemma 4.1 If $M = \langle D, v\rangle$ is a two-valued **CLuN**-model, then there is a $\mathfrak{S}$-valued **CLuN**-model $M' = \langle D, V\rangle$ such that, for all $A \in \mathcal{W}_{\mathcal{O}}$, $V_{M'}(A)[A] = v_M(A)$ and, where $L_A = \langle B_1, B_2, \ldots\rangle$, $V_{M'}(A)[B_i] = v(B_i)$ for all $i \in \{1, 2, \ldots\}$.

Proof From M construct M' as follows. For all $\alpha \in C \cup \mathcal{O}$, $V(\alpha) = v(\alpha)$. If $A \in S$ and $L_A = \langle B_1, B_2, \ldots\rangle$ then $V(A) = \langle v_M(A), v(B_1), v(B_2), \ldots\rangle$. For all $\alpha_1, \ldots, \alpha_r \in C \cup \mathcal{O}$, $\langle \alpha_1, \ldots, \alpha_r\rangle \in V^{\mathfrak{s}}(\pi^r)$ iff, where $L_{\pi^r\alpha_1\ldots\alpha_r} = \langle B_1, B_2, \ldots\rangle$, $\mathfrak{s} = \langle v_M(\pi^r\alpha_1 \ldots \alpha_r), v(B_1), v(B_2), \ldots\rangle$.

From this, one establishes the lemma for $A \in \mathcal{W}^a_{\mathcal{O}}$. With that as a basis, one establishes the lemma by an obvious induction. □

Lemma 4.2 If $M' = \langle D, V\rangle$ is a $\mathfrak{S}$-valued **CLuN**-model, then there is a two-valued **CLuN**-model $M = \langle D, v\rangle$ such that, for all $A \in \mathcal{W}_{\mathcal{O}}$, $v_M(A) = V_{M'}(A)[A]$ and, where $L_A = \langle B_1, B_2, \ldots\rangle$, $v(B_i) = V_{M'}(A)[B_i]$ for all $i \in \{1, 2, \ldots\}$.

Proof Obvious in view of the converse of the transformation described in the proof of Lemma 4.1. □

Theorem 4.3 $\Gamma \vDash_{\mathbf{CLuN}} A$ iff $\Gamma \vDash^{\mathfrak{S}}_{\mathbf{CLuN}} A$.
Proof $\Rightarrow$ Suppose that $\Gamma \nvDash^{\mathfrak{S}}_{\mathbf{CLuN}} A$. So there is a $\mathfrak{S}$-valued **CLuN**-model M that verifies all members of Γ and falsifies A. By the transformation described in the proof of Lemma 4.1, there is a two-valued **CLuN**-model M' that verifies exactly the same formulas as M.

$\Leftarrow$ Suppose that $\Gamma \nvDash_{\mathbf{CLuN}} A$. In view of the relevant Löwenheim–Skolem theorem, a countable two-valued **CLuN**-model M verifies all members of Γ and falsifies A. By the transformation described in the proof of Lemma 4.2, there is a $\mathfrak{S}$-valued **CLuN**-model M' that verifies exactly the same formulas as M. □

Let $\mathfrak{T}$ be the set of all finite sequences of 0s and 1s. The logic **CLuN** is compact. This enables us to replace the $\mathfrak{S}$-valued semantics by a semantics that takes its values from $\mathfrak{T}$.

As a first step, we represent some infinite sequences of 0s and 1s by finite sequences. This is easy: remove trailing ones from the sequences, reducing for example the infinite sequence $\langle 110100011\ldots\rangle$, in which the ellipsis represents ones only, by the finite sequence $\langle 1101000\rangle$. Of course, the first member of a sequence is always retained. So the finite sequences are $\langle 0\rangle, \langle 1\rangle, \langle 00\rangle, \langle 10\rangle, \langle 000\rangle, \langle 010\rangle, \langle 100\rangle, \langle 110\rangle, \ldots$, all but one of which end with a 0. The remaining infinite sequences are those in which there is no last 0. We simply remove them.

The set $\mathfrak{T}$ of valuation values is the set of sequences $\langle d_0, \ldots, d_n\rangle$ such that (i) $n \geq 0$, (ii) $d_0, \ldots, d_n \in \{0, 1\}$, and (iii) $d_n = 0$ if $n > 0$. Let $\mathfrak{t}$ be a variable for members of $\mathfrak{T}$. Let $\mathfrak{T}_D = \{\langle d_0, \ldots, d_n\rangle \in \mathfrak{T} \mid d_0 = 1\}$ (the set of designated values). Where $V_M(A) = \langle d_0, \ldots, d_n\rangle$, define $V_M(A)[A] = d_0$ and define, for all $B \in \text{fsup}(A)$, $V_M(A)[B] = d_{\#_A(B)}$ if $\#_A(B) \leq n$ and $V_M(A)[B] = 1$ otherwise.[21] Where $\langle d_0, \ldots, d_n\rangle$ is a sequence of 0s and 1s, let $\langle d_0, \ldots, d_n \rhd\rangle$ be the result of removing trailing 1s as long as the sequence counts more than one digit. So $\langle d_0, \ldots, d_n \rhd\rangle \in \mathfrak{T}$.

Fact 12 $\mathfrak{T}$ is denumerable (infinite and enumerable) and decidable.

After these preliminaries, let us turn to the semantics. In a $\mathfrak{T}$-valued **CLuN**-model $M = \langle D, V\rangle$ (defined over the countable pseudolanguage schema $\mathcal{L}_{\mathcal{O}}$), the domain D is a countable set and the assignment V has the following four properties. (i) $V: S \to \mathfrak{T}$. (ii) $V: C \cup \mathcal{O} \to D$ (where $D = \{V(\alpha) \mid \alpha \in C \cup \mathcal{O}\}$). (iii) Where $\mathfrak{T} = \{\mathfrak{t}^1, \mathfrak{t}^2, \ldots\}$,

[21] So, if $L_A = \langle B_1, B_2, \ldots\rangle$ and $V_M(A) = \langle 1110\rangle$, then $V_M(A)$ contains the information that in the corresponding two-valued model M' holds: $v_{M'}(A) = 1$, $v(B_1) = 1$, $v(B_2) = 1$, $v(B_3) = 0$, and $v(B_i) = 1$ whenever $i > 3$; similarly, $V_M(A) = \langle 0\rangle$ then contains the information that in the corresponding two-valued model M' holds: $v_{M'}(A) = 0$ and $v(B_i) = 1$ for all $i \in \{1, 2, \ldots\}$.

$V{:}\mathcal{P}^r \to \langle\Sigma_{t^1}, \Sigma_{t^2}, \ldots\rangle$ such that (a) $\{\Sigma_{t^1}, \Sigma_{t^2}, \ldots\}$ is a pseudopartition of $\wp(D^r)$ and (b) if $V(\beta) = V(\alpha_i)$ $(1 \leq i \leq r)$, $\langle\alpha_1, \ldots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \ldots, \alpha_r\rangle \in \Sigma_{t^j}$, and $\langle\alpha_1, \ldots, \alpha_{i-1}, \beta, \alpha_{i+1}, \ldots, \alpha_r\rangle \in \Sigma_{t^k}$, then $t^j, t^k \in \mathfrak{T}_D$ or $t^j, t^k \in \mathfrak{T} - \mathfrak{T}_D$. To simplify the notation, consider V as composed in this case of the denumerably many functions V^{t^1}, V^{t^2}, ..., with $V^{t^1}(\pi^r) = \Sigma_{t^1}$, $V^{t^2}(\pi^r) = \Sigma_{t^2}$, and so on. (iv) Identity is handled as a binary predicate with the special characteristic that $\bigcup\{V^t(=) \mid t \in \mathfrak{T}_D\} = \{\langle\alpha, \beta\rangle \mid \alpha, \beta \in \mathcal{C} \cup \mathcal{O}; V(\alpha) = V(\beta)\}$.

In the clauses below, let $V_M(A)$ count n_A members and let $V_M(A(a))$ count $n_{A(a)}$ members. The valuation function $V_M{:}\,\mathcal{W}_\mathcal{O} \to \mathfrak{T}$ is defined as follows:

C$\mathcal{S}$ where $A \in \mathcal{S}$, $V_M(A) = V(A)$
C$\mathcal{P}^r$ $V_M(\pi^r\alpha_1 \ldots \alpha_r) = t$ iff $\langle\alpha_1, \ldots, \alpha_r\rangle \in V^t(\pi^r)$
C$\neg$ Where $L_{\neg A} = \langle C^1, C^2, \ldots\rangle$, $V_M(\neg A) = \langle\max(1 - V_M(A)[A], V_M(A)[\neg A]),$ $V_M(A)[C^1], \ldots, V_M(A)[C^{n_A}] \rhd\rangle$.
C$\wedge$ Where $L_{A\wedge B} = \langle C^1, C^2, \ldots\rangle$, $V_M(A \wedge B) = \langle\min(V_M(A)[A], V_M(B)[B]),$ $V_M(A)[C^1], \ldots, V_M(A)[C^{n_A})] \rhd\rangle$.
C$\vee$ Where $L_{A\vee B} = \langle C^1, C^2, \ldots\rangle$, $V_M(A \vee B) = \langle\max(V_M(A)[A], V_M(B)[B]),$ $V_M(A)[C^1], \ldots, V_M(A)[C^{n_A})] \rhd\rangle$.
C$\supset$ Where $L_{A\supset B} = \langle C^1, C^2, \ldots\rangle$, $V_M(A \supset B) = \langle\max(1 - V_M(A)[A], V_M(B)[B]),$ $V_M(A)[C^1], \ldots, V_M(A)[C^{n_A})] \rhd\rangle$.
C$\equiv$ Where $L_{A\equiv B} = \langle C^1, C^2, \ldots\rangle$, $V_M(A \equiv B) = \langle\min(\max(1 - V_M(A)[A],$ $V_M(B)[B]), \max(1 - V_M(B)[B], V_M(A)[A])), V_M(A)[C^1], \ldots, V_M(A)[C^{n_A})] \rhd\rangle$.
C$\forall$ Where $L_{\forall\alpha A(\alpha)} = \langle C^1, C^2, \ldots\rangle$, $V_M(\forall\alpha A(\alpha)) = \langle\min\{V_M(A(\beta)) \mid \beta \in \mathcal{C} \cup \mathcal{O}\},$ $V_M(A(a))[C^1], \ldots, V_M(A(a))[C^{n_{A(a)}}] \rhd\rangle$.
C$\exists$ Where $L_{\exists\alpha A(\alpha)} = \langle C^1, C^2, \ldots\rangle$, $V_M(\exists\alpha A(\alpha)) = \langle\max\{V_M(A(\beta)) \mid \beta \in \mathcal{C} \cup \mathcal{O}\},$ $V_M(A(a))[C^1], \ldots, V_M(A(a))[C^{n_{A(a)}}] \rhd\rangle$.

A $M \Vdash A$ (a $\mathfrak{T}$-valued **CLuN**-model M verifies A) iff $V_M(A) \in \mathfrak{T}_D$; M is a model of $\{A_1, \ldots, A_n\}$ iff $M \Vdash A_1, \ldots,$ and $M \Vdash A_n$. $B_1, \ldots, B_n \vDash^{\mathfrak{T}}_{\mathbf{CLuN}} A$ iff every $\mathfrak{T}$-valued **CLuN**-model of $\{B_1, \ldots, B_n\}$ verifies A; $\Gamma \vDash^{\mathfrak{T}}_{\mathbf{CLuN}} A$ iff $B_1, \ldots, B_n \vDash^{\mathfrak{T}}_{\mathbf{CLuN}} A$ for some $B_1, \ldots, B_n \in \Gamma$.

The only "inconvenience" with this semantics is that, for some Γ and A, $\Gamma \nvDash^{\mathfrak{T}}_{\mathbf{CLuN}} A$ while no $\mathfrak{T}$-valued **CLuN**-model of Γ falsifies A. An example is $\{\neg^{3n}p, \neg^{3n+1}p, \neg^{3n+2}p \supset q \mid n \in \mathbb{N}\} \nvDash^{\mathfrak{T}}_{\mathbf{CLuN}} q$ in which $\neg^i$ abbreviates a sequence of i occurrences of $\neg$. Every two-valued **CLuN**-model M of $\{\neg^{3n}p, \neg^{3n+1}p, \neg^{3n+2}p \supset q \mid n \in \mathbb{N}\}$ that falsifies q has, for all $n \in \mathbb{N}$, $v(\neg^{3n+1}p) = 1$ and $v(\neg^{3n+2}p) = 0$. There is a $\mathfrak{S}$-valued **CLuN**-model that corresponds to M but obviously not a $\mathfrak{T}$-valued one. Indeed, as $V_{M'}(p) = V(p)$ counts finitely many digits, say m, there is bound to be a n such that $\#_p(\neg^{3n+2}p) > m$, whence $V_{M'}(\neg^{3n+2}p) \in \mathfrak{T}_D$; but then $V_{M'}(\neg^{3n+2}p \supset q) \notin \mathfrak{T}_D$ or $V_{M'}(q) \in \mathfrak{T}_D$; so if M' is a model of $\{\neg^{3n}p, \neg^{3n+1}p, \neg^{3n+2}p \supset q \mid n \in \mathbb{N}\}$, it falsifies q. Of course the inconvenience has no effect on the semantic consequence relation in view of the special way in which it is defined.

Fact 13 The $\mathfrak{T}$-valued **CLuN**-semantics is recursive.

Fact 14 All logical symbols are truth-functions in the $\mathfrak{T}$-valued **CLuN**-semantics.

Lemma 4.4 If $M = \langle D, v\rangle$ is a two-valued **CLuN**-model, $M \Vdash A$, $M \nVdash B$, and $M' = \langle D, v'\rangle$ is obtained from M by letting $v(\neg C) = 1$ whenever $\neg C \notin \{A, B\} \cup \text{sub}(A) \cup \text{sub}(B)$, then M' is a two-valued **CLuN**-model, $M' \Vdash A$ and $M' \nVdash B$.

Proof By an obvious induction on the length of A or of B, whichever is longer. □

Lemma 4.5 If $M = \langle D, v\rangle$ is a two-valued **CLuN**-model, and, for every $A \in \mathcal{W}^a_{\mathcal{O}}$, $v(B) = 0$ for at most finitely many $B \in \text{fsup}(A)$, then there is a $\mathfrak{T}$-valued **CLuN**-model $M' = \langle D, V'\rangle$ such that, for all $A \in \mathcal{W}_{\mathcal{O}}$, $V_{M'}(A)[A] = v_M(A)$ and, where $L_A = \langle B_1, B_2, \ldots\rangle$, $V_{M'}(A)[B_i] = v(B_i)$ for all $i \in \{1, 2, \ldots\}$.

Proof Suppose that the antecedent is true. In view of Lemma 4.1, there is a $\mathfrak{S}$-valued **CLuN**-model $M'' = \langle D, V''\rangle$ with the required property. The $\mathfrak{S}$-valued **CLuN**-model M'' is transformed to the required $\mathfrak{T}$-valued **CLuN**-model $M' = \langle D, V'\rangle$ by the following steps. Consider an $A \in S$ and let $V_{M''}(A) = \mathfrak{s}$. As, for every $A \in \mathcal{W}^a_{\mathcal{O}}$, $v(B) = 0$ for at most finitely many $B \in \text{fsup}(A)$, there is bound to be a last 0 in the sequence $\mathfrak{s}$. So the result of removing trailing 1s from $\mathfrak{s}$ results in a member of $\mathfrak{T}$. The reasoning for formulas $\pi^r\alpha_1 \ldots \alpha_r$ proceeds similarly. Next one invokes an obvious induction on the length of $A \in \mathcal{W}_{\mathcal{O}}$ as in Lemma 4.1. □

Lemma 4.6 If $M' = \langle D, V\rangle$ is a $\mathfrak{T}$-valued **CLuN**-model, then there is a two-valued **CLuN**-model $M = \langle D, v\rangle$ such that, for all $A \in \mathcal{W}_{\mathcal{O}}$, $v_M(A) = V_{M'}(A)[A]$ and, where $L_A = \langle B_1, B_2, \ldots\rangle$, $v(B_i) = V_{M'}(A)[B_i]$ for all $i \in \{1, 2, \ldots\}$.

Proof Obvious. □

Lemma 4.7 If M is a two-valued **CLuN**-model, M' is a $\mathfrak{T}$-valued **CLuN**-model, and M and M' correspond in the sense of Lemmas 4.5 and 4.6, then $M \Vdash A$ iff $M' \Vdash A$ for all $A \in \mathcal{W}_{\mathcal{O}}$.

Proof A proof by cases gives one the result for $A \in \mathcal{W}^a_{\mathcal{O}}$. This provides the basis for the obvious induction on the complexity of $A \in \mathcal{W}_{\mathcal{O}}$. □

Theorem 4.8 $\Gamma \vDash^{\mathfrak{T}}_{\mathbf{CLuN}} A$ iff $\Gamma \vDash_{\mathbf{CLuN}} A$.

Proof ⇒ Suppose that $\Gamma \nvDash_{\mathbf{CLuN}} A$. Consider any $B_1, \ldots, B_n \in \Gamma$. As **CLuN** is compact, $B_1, \ldots, B_n \nvDash_{\mathbf{CLuN}} A$. Note that $B_1 \wedge \ldots \wedge B_n \nvDash_{\mathbf{CLuN}} A$. In view of the relevant Löwenheim–Skolem theorem and of Lemma 4.4, there is a countable binary **CLuN**-model $M = \langle D, v\rangle$ such that $M \Vdash B_1 \wedge \cdots \wedge B_n$, and $M \nVdash A$ and $v(C) = 0$ for at most finitely many formulas $C \notin \mathcal{W}^a_{\mathcal{O}}$. In view of Lemmas 4.5 and 4.7, it follows that there is a $\mathfrak{T}$-valued **CLuN**-model M' such that $M' \Vdash B_1 \wedge \cdots \wedge B_n$ and $M' \nVdash A$.

⇐ Suppose that $\Gamma \nvDash^{\mathfrak{T}}_{\mathbf{CLuN}} A$. The definition of $\vDash^{\mathfrak{T}}_{\mathbf{CLuN}}$ entails, for arbitrary $B_1, \ldots, B_n \in \Gamma$, that $B_1 \wedge \cdots \wedge B_n \nvDash^{\mathfrak{T}}_{\mathbf{CLuN}} A$. So a $\mathfrak{T}$-valued **CLuN**-model verifies $B_1 \wedge \cdots \wedge B_n$ and falsifies A. By Lemmas 4.6 and 4.7, there is a two-valued **CLuN**-model M' such that $M' \Vdash B_1 \wedge \cdots \wedge B_n$ and $M' \nVdash A$. □

The members of $\mathfrak{T}$ are obtained from members of $\mathfrak{S}$ by removing trailing 1s. The reader may find it more convenient to consider finite sequences of 0s and 1s that are obtained by removing trailing 0s. Another alternative is finite sequences from which trailing 1s are removed except for the first one as well as finite sequences from which trailing 0s are removed except for the first one. In both cases, the result corresponds to the one presented above and the inconvenience is exactly the same.

5 A Different Many-Valued Approach

Take another look at the two-valued **CLuNs**-semantics from Sect. 2, comparing it to the three-valued **CLuNs**-semantics from Sect. 3. That a two-valued model verifies $p \wedge \neg p$ is the result of $v(p) = 1$ and $v(\neg p) = 1$ and these are separate and independent "facts." If a two-valued model verifies $A \wedge \neg A$ for complex A, then again $M \Vdash A$ depends on one set of "facts" whereas $M \Vdash \neg A$ depends on a separate set of "facts." The idea behind the three-valued semantics is completely different. That a three-valued model verifies $p \wedge \neg p$ depends on the sole "fact" that $V(p) = I$. Moreover, if $M \Vdash A \wedge \neg A$ for complex A, then, as a little inspection readily reveals, there are always inconsistent "facts" on which both $M \Vdash A$ and $M \Vdash \neg A$ depend.

I now set out to construct a very different kind of many-valued semantics, called tuaf semantics. For a start, the assignment of the tuaf semantics will be the same as the assignment of the two-valued semantics. I shall retain the convention that valuation values are determined starting from the least complex formulas. If $v(p) = 1$, then p will obtain the valuation value t (for *true*), independent of the valuation value of $\neg p$. If $v_M(p) = t$ and $v(\neg p) = 0$, then $v_M(\neg p) = f$ (for *false*). However, if $v_M(p) = t$ and $v(\neg p) = 1$, then $v_M(\neg p) = u$ (for *glut*). Similarly, if $v_M(p) = f$ and $v(\neg p) = 0$, then $v_M(\neg p) = a$ (for *gap*).[22] The idea is that a formula receives the valuation value u, respectively a, iff it has the wrong truth-value with respect to its subformulas. This idea is clearly different from the usual one, described in Sect. 3. However, as we shall soon see, some choices have still to be made in order to turn the approach into something workable.

By way of preparation, we start with an alternative formulation of the **CL**-semantics. The semantics from Sect. 1 will be called the *clausal semantics*. Let us turn it into a *tabular semantics* by leaving the assignment function unchanged, replacing the ten clauses specifying the valuation function by the following ten tables – the last two are amalgamated.

Where $A \in S$:

$v(A)$	A
1	1
0	0

Where $\alpha_1, \ldots, \alpha_n \in \mathcal{C} \cup \mathcal{O}$ and $\pi \in \mathcal{P}^n$:

$\langle v(\alpha_1), \ldots, v(\alpha_n)\rangle, v(\pi)$	$\pi\alpha_1 \ldots \alpha_n$
$\in$	1
$\notin$	0

[22] Some values will be absent for some logics; **CLuNs**, for example, does not allow for gaps.

Where $\alpha, \beta \in \mathcal{C} \cup \mathcal{O}$:

$v(\alpha), v(\beta)$	$\alpha = \beta$
$=$	1
$\neq$	0

	$\neg$
1	0
0	1

$\wedge$	1	0
1	1	0
0	0	0

$\vee$	1	0
1	1	1
0	1	0

$\supset$	1	0
1	1	0
0	1	1

$\equiv$	1	0
1	1	0
0	0	1

$\{v_M(A(\alpha)) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$	$\forall\alpha(A(\alpha))$	$\exists\alpha(A(\alpha))$
$\{1\}$	1	1
$\{0, 1\}$	0	1
$\{0\}$	0	0

We have seen that some logics display gluts or gaps or both. For them, I shall articulate a semantics in which the valuation has the form $v_M \colon \mathcal{W}_{\mathcal{O}} \to \{t, u, a, f\}$, in which the values intuitively stand for true, glut, gap, and false – hence the name *tuaf semantics*. The designated values are t and u. So $M \Vdash A$ iff $v_M(A) \in \{t, u\}$. This settles at once the semantic consequence relation $\Gamma \vDash A$.

Some logics do not allow for gluts or for gaps. So only two or three of the values will be used in their tuaf semantics. For example, the tuaf semantics for **CL** is boringly isomorphic to **CL**'s tabular semantics: for the valuation, every 1 is replaced by t and every 0 by f – the point of the replacement will soon become clear. I spell out this semantics for future reference.[23]

Where $A \in \mathcal{S}$:

$v(A)$	A
1	t
0	f

Where $\alpha_1, \ldots, \alpha_n \in \mathcal{C} \cup \mathcal{O}$ and $\pi \in \mathcal{P}^n$:

$\langle v(\alpha_1), \ldots, v(\alpha_n)\rangle, v(\pi)$	$\pi\alpha_1 \ldots \alpha_n$
$\in$	t
$\notin$	f

Where $\alpha, \beta \in \mathcal{C} \cup \mathcal{O}$:

$v(\alpha), v(\beta)$	$\alpha = \beta$
$=$	t
$\neq$	f.

These three first tables, which concern the atomic formulas, are identical for all subsequent logics. They will not be repeated.

	$\neg$
t	f
f	t

$\wedge$	t	f
t	t	f
f	f	f

$\vee$	t	f
t	t	t
f	t	f

$\supset$	t	f
t	t	f
f	t	t

$\equiv$	t	f
t	t	f
f	f	t

$\{v_M(A(\alpha)) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$	$\forall\alpha(A(\alpha))$	$\exists\alpha(A(\alpha))$
$\{t\}$	t	t
$\{f, t\}$	f	t
$\{f\}$	f	f

[23] I use the same notation, $v_M(A)$, for the valuation function in all three kinds of semantics and I shall do so for all logics. The matter is always disambiguated by the context.

The matter gets interesting when we move to logics that tolerate gluts or gaps. Let us start with **CLuN**. Its tabular semantics is identical to that for **CL**, except that the table for negation is replaced.

A	$v(\neg A)$	$\neg A$
1	0	0
1	1	1
0	(any)	1

The table describes $v_M(\neg A)$ as a function of $v_M(A)$ and of $v(\neg A)$. The "(any)" indicates that the value of $v(\neg A)$ has no effect at this point, viz. where $v_M(A) = 0$.

Let us turn to the tuaf semantics of **CLuN**. Its assignment is as for all two-valued semantics in this chapter and the three valuation tables for atomic formulas is as for **CL**. The rest of the valuation function is determined by the following tables – some explanation follows:

A	$v(\neg A)$	$\neg A$
t	0	f
t	1	u
u	0	f
u	1	u
f	(any)	t

$\wedge$	t	u	f
t	t	t	f
u	t	t	f
f	f	f	f

$\vee$	t	u	f
t	t	t	t
u	t	t	t
f	t	t	f

$\supset$	t	u	f
t	t	t	f
u	t	t	f
f	t	t	t

$\equiv$	t	u	f
t	t	t	f
u	t	t	f
f	f	f	t

$\{v_M(A(\alpha)) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$	$\forall\alpha(A(\alpha))$	$\exists\alpha(A(\alpha))$
$\subseteq \{t, u\}$	t	t
$= \{f\}$	f	f
(other)	f	t

As the value u is *introduced* by the table for negation – if $v_M(A) \in \{t, u\}$ and $v(\neg A) = 1$, then $v_M(\neg A) = u$ – the value u has to occur in all tables in which the input entries are valuation values. The "(any)" has the same meaning as in the tabular semantics. The "(other)" obviously means that the set $\{v_M(A(\alpha)) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$ contains at least one f and at least one t or u.

Until now, the design of the tuaf semantics proceeded on somewhat insecure grounds. There apparently is a clear design behind it, but the design is not made fully explicit. So let us see where precisely the tuaf semantics assigns the value u? This question may be answered in several ways and, depending on the answer, the tuaf semantics of other logics will vary. That the question may be answered in at least two different ways is caused by the fact that **CLuN** has the following remarkable property: for any formula A, an adequate **CLuN**-semantics has models M and M' such that (i) $M \Vdash A$ and $M \Vdash \neg A$, (ii) $M' \Vdash A$ and $M' \nVdash \neg A$, and (iii) M and M' verify exactly the same subformulas of A. So for no true A, however complex, does the truth of its negation result from its proper subformulas. **CLuN** has also a different, actually more general, property. Consider a complexity function that assigns to a formula A a complexity $c(A)$ that is higher than the complexity it assigns to any proper subformula of A. For any formula A, there are two-valued **CLuN**-models M and M' such that the aforementioned (i) and (ii) obtain whereas M and M' verify exactly the same subset of $\{B \mid c(B) < c(A)\}$.

A first view on the tuaf semantics of **CLuN** may be called the *agreement view*. In the *tabular* semantics for **CL**, every table defines, for a nonatomic form A, $v_M(\mathsf{A})$ as a function of the valuation value of subformulas of A. So it is easy to check whether the valuation function of **CLuN**, or of any other logic **L** allowing for gluts or gaps, agrees with **CL** *at a specific point*. If a two-valued **L**-model M has $v_M(A) = v_M(B) = v_M(A \wedge B) = 1$, then the tuaf **L**-model M' has $v_{M'}(A \wedge B) = t$ because $v_M(A \wedge B) = 1$ agrees with all **CL**-models M that have $v_M(A) = v_M(B) = 1$. If a two-valued **L**-model M has $v_M(A) = 0$ and $v_M(A \wedge B) = 1$, then the tuaf **L**-model M' has $v_{M'}(A \wedge B) = u$ because **CL**-models have $v_M(A \wedge B) = 0$ whenever they have $v_M(A) = 0$. By a similar reasoning, if the **L**-model M has $v_M(A) = v_M(B) = 1$ and $v_M(A \wedge B) = 0$, then the tuaf **L**-model M' has $v_{M'}(A \wedge B) = a$. It is left to the reader to check that the occurrence of output entries u in the tuaf semantics of **CLuN** is in line with the agreement view.

Next there is what I shall call the *interference view*. The two output entries u in the tuaf semantics of **CLuN** depend on the assignment. For both, $v(\neg A) = 1$. If this is modified to $v(\neg A) = 0$, the value of $v_M(\neg A)$ is modified to f. So, on the interference view, a complex formula A obtains the valuation value u, respectively a, iff the two-valued valuation value depends on the assignment value $v(A)$, and not only on valuation values of subformulas of A. So this view incorporates the agreement view and moreover takes into account whether the valuation value of a *specific* complex formula is a function of the valuation values of its subformulas. Put differently, the values u and a are assigned at points where gluts or gaps *originate*.

An example clarifies this even further. In the two-valued **CLuNs**-semantics, $v_M(\neg\neg A) = v_M(A)$. Let $v_M(p) = 1 = v(\neg p)$ in a two-valued **CLuNs**-model M, whence $v_M(\neg p) = v_M(\neg\neg p) = 1$. So the corresponding tuaf model M has $v_M(\neg p) = u$ because if $v(\neg p)$ were 0, then $v_M(\neg p)$ would be 0 in the two-valued semantics. However, $v_M(\neg\neg p) = t$ on the interference view. Indeed, although $v_M(\neg p) = 1$ and $v_M(\neg\neg p) = 1$ in the two-valued semantics, the latter value does not depend on $v(\neg\neg p)$ but is a direct result of $v_M(p) = 1$.[24] Note that the agreement view leads to a different result at this point for **CLuNs**; it leads to $v_M(\neg\neg p) = u$ in the considered example.

It is again left to the reader to check that the occurrence of output entries u in the tuaf semantics of **CLuN** is in line with the interference view. Both views explain the absence of output entries u outside the negation table. They also clarify in general why $\neg A$ may have the valuation value u whereas A will never have that valuation value, unless of course in case A itself has the form $\neg B$. In the sequel of this chapter, I shall restrict attention to the interference view.

The tabular semantics as well as the tuaf semantics of **CLuNs** require the equivalence classes defined in Sect. 2. The tabular semantics is just like that for **CL**, except that the

[24] The counterfactual and causal phraseology can obviously be rephrased extensionally (in terms of all models that have certain properties).

table for negation is replaced by the following tables:

Where $A \in \mathcal{W}^a_\mathcal{O}$:

A	$\{v(\neg B) \mid B \in [\![A]\!]\}$	$\neg A$
1	$= \{0\}$	0
1	$\neq \{0\}$	1
0	(any)	1

A	$\neg\neg A$
$A \wedge \neg B$	$\neg(A \supset B)$
$\neg A \vee \neg B$	$\neg(A \wedge B)$
$\neg A \wedge \neg B$	$\neg(A \vee B)$
$(A \vee B) \wedge (\neg A \vee \neg B)$	$\neg(A \equiv B)$
$\exists\alpha\neg A(\alpha)$	$\neg\forall\alpha A(\alpha)$
$\forall\alpha\neg A(\alpha)$	$\neg\exists\alpha\neg A(\alpha)$
1	1
0	0

The lower table is obviously a summary of seven tables, each stating that the formula in the right column has the same value as the formula in the left column. Of course, the fascinating bit is the tuaf semantics. Again, the assignment is as for **CL** and so are the three tables for the atomic formulas.

Where $A \in \mathcal{W}^a_\mathcal{O}$:

A	$\{v(\neg B) \mid B \in [\![A]\!]\}$	$\neg A$
t	$= \{0\}$	f
t	$\neq \{0\}$	u
f	(any)	t

A	$\neg\neg A$
$A \wedge \neg B$	$\neg(A \supset B)$
$\neg A \vee \neg B$	$\neg(A \wedge B)$
$\neg A \wedge \neg B$	$\neg(A \vee B)$
$(A \vee B) \wedge (\neg A \vee \neg B)$	$\neg(A \equiv B)$
$\exists\alpha\neg A(\alpha)$	$\neg\forall\alpha A(\alpha)$
$\forall\alpha\neg A(\alpha)$	$\neg\exists\alpha\neg A(\alpha)$
t	t
u	t
f	f

$\wedge$	t	u	f
t	t	t	f
u	t	t	f
f	f	f	f

$\vee$	t	u	f
t	t	t	t
u	t	t	t
f	t	t	f

$\supset$	t	u	f
t	t	t	f
u	t	t	f
f	t	t	t

$\equiv$	t	u	f
t	t	t	f
u	t	t	f
f	f	f	t

$\{v_M(A(\alpha)) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$	$\forall\alpha(A(\alpha))$	$\exists\alpha(A(\alpha))$
$\subseteq \{t, u\}$	t	t
$= \{f\}$	f	f
(other)	f	t

There is only one output entry u in all these tables. Atomic formulas never receive the value u. Formulas of which the central symbol is not a negation cannot receive a u because their valuation value (does not depend on their assignment value but) is fully determined by the valuation value of less complex formulas. In the table for the negation of complex formulas, there is a u among the input entries. Please note that the only formulas in the left column that may have a valuation value u are those of the form A that moreover have the form $\neg C$. But even if this formula has the value u, the formula of the corresponding form $\neg\neg A$ needs the value t because the assignment does not interfere. Indeed, it holds within the two-valued **CLuNs**-semantics that $v_M(\neg\neg A) = 1$ if $v_M(A) = 1$, whatever $v(\neg\neg A)$.

I mentioned before that the tuaf semantics introduces values u and a where the gluts or gaps *originate*. Please check this. If $v_M(Pa) = v_M(\neg Pa) = 1$ in the clausal or tabular **CLuNs**-semantics, the tuaf semantics settles for $v_M(\neg Pa) = u$. It holds within the two-valued **CLuNs**-semantics that if $v_M(Pa) = v_M(\neg Pa) = v_M(Qb) = 1$, then $v_M(Pa \wedge Qb) = v_M(\neg Pa \vee \neg Qb) = 1$, and hence also $v_M(Pa \wedge Qb) = v_M(\neg(Pa \wedge Qb)) = 1$. The tuaf semantics settles for $v_M(\neg(Pa \wedge Qb)) = t$. This is precisely as we want it: the glut does not originate with $\neg(Pa \wedge Qb)$; it originates with $\neg Pa$.

Consider the tuaf semantics of a very weak extension of **CLuN**, viz. with $A \supset \neg\neg A$, for which I shall use the rather arbitrary name **CLuN**^{NN}. The tuaf semantics is identical to that of **CLuN**, except for the tables for negation – there are two of them. Let $\mathcal{W}^n_O$ be the set of formulas that do not have $\neg$ as their first symbol.

Where $A \in \mathcal{W}^n_O$:

A	$v(\neg A)$	$\neg A$
t	0	f
t	1	u
f	(any)	t

A	$v(\neg\neg A)$	$\neg\neg A$
t	(any)	t
u	(any)	t
f	0	f
f	1	u

There are no input entries u in the first table because only formulas of the form $\neg B$ can have the value u. The output value of the second line of the second table is t because the assignment does not interfere. If $v_M(A) = f$, then $v_M(\neg A) = t$, whence $v_M(\neg\neg A) \in \{u, f\}$.

There are two output entries u in this semantics. So negation gluts originate at two kinds of points in **CLuN**^{NN}-models, first where the negation of a nonnegated verified formula is itself verified and next where the double negation of a falsified formula is verified – the negation of the formula is then verified and, by the interference of the assignment, also its double negation.

Let us, as a nonparaconsistent illustration, consider the tuaf semantics for the logic with the beautiful name **CLuCoDaM**, in words, the logic that leaves room for conjunction gluts, for both disjunction gluts and disjunction gaps, and for implication gaps. Let us consider the version in which Replacement of Identicals is not added. So we do not need the equivalence classes from the **CLuNs**-semantics. Moreover, I skip the tabular semantics. The reader may very easily construct it in case the tuaf semantics would not be obvious at once. The assignment and the valuation tables for atomic formulas are as for **CL**.

	$\neg$
t	f
u	f
a	t
f	t

$\equiv$	t	u	a	f
t	t	t	f	f
u	t	t	f	f
a	f	f	t	t
f	f	f	t	t

$v(A \wedge B) = 1$:

$\wedge$	t	u	a	f
t	t	t	u	u
u	t	t	u	u
a	u	u	u	u
f	u	u	u	u

$v(A \vee B) = 1$:

$\vee$	t	u	a	f
t	t	t	t	t
u	t	t	t	t
a	t	t	u	u
f	t	t	u	u

$v(A \supset B) = 1$:

$\supset$	t	u	a	f
t	t	t	f	f
u	t	t	f	f
a	t	t	t	t
f	t	t	t	t

$v(A \wedge B) = 0$:

$\wedge$	t	u	a	f
t	t	t	f	f
u	t	t	f	f
a	f	f	f	f
f	f	f	f	f

$v(A \vee B) = 0$:

$\vee$	t	u	a	f
t	a	a	a	a
u	a	a	a	a
a	a	a	f	f
f	a	a	f	f

$v(A \supset B) = 0$:

$\supset$	t	u	a	f
t	a	a	f	f
u	a	a	f	f
a	a	a	a	a
f	a	a	a	a

$\{v_M(A(\alpha)) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$	$\forall\alpha(A(\alpha))$	$\exists\alpha(A(\alpha))$
$\subseteq \{t, u\}$	t	t
$\subseteq \{a, f\}$	f	f
(other)	f	t

This semantics illustrates a variety of cases. As there are no gluts or gaps with respect to negation, equivalence, and the quantifiers, the output entries are all t and f in the tables for those logical symbols. For conjunction, there are only gluts. So if $v(A \wedge B) = 0$, one obtains the normal table; if $v(A \wedge B) = 1$ every f in the normal table is replaced by a u. Implication and disjunction illustrate the other cases.

The tuaf semantics of da Costa's $\mathbf{C}_n$ logics [23] illustrates a possible complication. Let $A^{(1)}$ abbreviate $\neg(C \wedge \neg C)$ and let $A \equiv^c B$ denote that A and B are congruent in the sense of Kleene *or* that one formula results from the other by deleting vacuous quantifiers – Kleene [30, p. 153] summarizes his definition as follows: "two formulas are congruent, if they differ only in their bound variables, and corresponding bound variables are bound by corresponding quantifiers." The congruence requirement may be handled by first defining a prevaluation, which looks just like a tuaf semantics itself, and next defining a valuation from the prevaluation. The tuaf semantics of $\mathbf{C}_1$ clarifies the matter.

The assignment function is again the general one, as in the **CL**-semantics from the beginning of this section. For atomic formulas, the prevaluation $v_M : \mathcal{W}_{\mathcal{O}} \to \{t, u, f\}$ has the same tables as the tuaf valuation of **CLuN** – these tables are not repeated.

Where $A \in \mathcal{W}^a_{\mathcal{O}}$:

$v_M(A)$	$v(\neg A)$	$v_M(\neg A)$
t	0	f
t	1	u
f	(any)	t

Where $\dagger \in \{\vee, \wedge, \supset\}$ and $A \dagger B$ has not the form $C \wedge \neg C$:

$v_M(A \dagger B)$	$v_M(A^{(1)})$	$v_M(B^{(1)})$	$v(\neg(A \dagger B))$	$v_M(\neg(A \dagger B))$
t	t	t	(any)	f
t	(other)		0	f
t	(other)		1	u
f	(any)		(any)	t

Where $\mathsf{Q} \in \{\forall, \exists\}$:

$v_M(\mathsf{Q}\alpha A(\alpha))$	$\{v_M(A(\beta)^{(1)}) \mid \beta \in \mathcal{C} \cup \mathcal{O}\}$	$v(\neg \mathsf{Q}\alpha A(\alpha))$	$v_M(\neg \mathsf{Q}\alpha A(\alpha))$
t	$= \{t\}$	(any)	f
t	$\neq \{t\}$	0	f
t	$\neq \{t\}$	1	u
f	(any)	(any)	t

The other tables apply to all members of $\mathcal{W}_{\mathcal{O}}$:

$v_M(\neg A)$	$v(\neg\neg A)$	$v_M(\neg\neg A)$
t	(any)	f
u	0	f
u	1	u
f	(any)	t

$\neg A$	$A^{(1)}$
t	t
u	f
f	t

$\wedge$	t	u	f
t	t	t	f
u	t	t	f
f	f	f	f

$\vee$	t	u	f
t	t	t	t
u	t	t	t
f	t	t	f

$\supset$	t	u	f
t	t	t	f
u	t	t	f
f	t	t	t

$\equiv$	t	u	f
t	t	t	f
u	t	t	f
f	f	f	t

$\{v_M(A(\alpha)) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$	$\forall\alpha A(\alpha)$	$\exists\alpha A(\alpha)$
$\in \wp\{t, u\}$	t	t
$= \{f\}$	f	f
(other)	f	t

Let $f(A)$ be obtained by first deleting all vacuous quantifiers in A and then systematically replacing all variables in the result by the first variables of the alphabet in alphabetical order. Next, define the valuation values V_M in terms of the prevaluation values v_M by $V_M(A) = v_M(f(A))$.

Alternatively, a prevaluation v_M is called a valuation iff $v_M(A) = v_M(B)$ whenever $A \equiv^c B$.

Transforming the above semantics to a $\mathbf{C}_n$ logic (for $1 < n < \omega$) is an easy exercise left to the reader – the formulation of the tables for $\mathbf{C}_1$ and the plot described in the previous paragraph indicate the road. For $\mathbf{C}_\omega$, one replaces the tables for negation by the left and middle tables below; for $\mathbf{C}_{\overline{\omega}}$ (which is $\mathbf{C}_\omega$ extended with classical negation, $\sim$) one adds the table to the right below.

$v_M(A)$	$v(\neg A)$	$v_M(\neg A)$
t	0	f
t	1	u
f	(any)	t

$v_M(\neg A)$	$v(\neg\neg A)$	$v_M(\neg\neg A)$
t	(any)	f
u	0	f
u	1	u
f	(any)	t

A	$\sim A$
t	f
u	f
f	t

Incidentally, an indeterministic tuaf semantics is often more transparent than its deterministic counterpart. As an indeterministic tuaf semantics does not refer to the valuation, it has less clutter in the heads of the tables. So let me display the relevant tables, viz. negation tables, for $\mathbf{C}_1$.

Where $A \in \mathcal{W}^a_{\mathcal{O}}$:

A	$\neg A$
t	$[f, u]$
f	t

Where $\dagger \in \{\vee, \wedge, \supset\}$ and $A \dagger B$ has not the form $C \wedge \neg C$:

$A * B$	$A^{(1)}$	$B^{(1)}$	$\neg(A * B)$
t	t	t	f
t	(other)		$[f, u]$
f	(any)		t

Where $\mathrm{Q} \in \{\forall, \exists\}$:

$\mathrm{Q}\alpha A(\alpha)$	$\{v_M(A(\beta)^{(1)}) \mid \beta \in \mathcal{C} \cup \mathcal{O}\}$	$\neg \mathrm{Q}\alpha A(\alpha)$
t	$\{t\}$	f
t	(other)	$[f, u]$
f	(any)	t

The other tables apply to all members of $\mathcal{W}_{\mathcal{O}}$:

$\neg A$	$\neg\neg A$
t	f
u	$[f, u]$
f	t

$\neg A$	$A^{(1)}$
t	t
u	f
f	t

The expression $[f, u]$ indicates that the value may be f or u – this is an indeterministic semantics. Note that the "normal" value, the one that agrees with **CL** at this point, is f. So if the value is u, it "drops from the sky" as far as the indeterministic semantics is concerned – in the deterministic semantics, the assignment function interferes at this point. The "dropping from the sky" holds for the semantics only; a premise set may require that some values are u in its models. For other logics, a premise set may require some values to be a. The metaphor is helpful, however, because it highlights that the values u and a occur at points where an abnormality is generated.

The application of the semantics should clearly be separated from the underlying idea. Even in indeterministic versions, the occurrence of output values u should be understood in terms of the assignment's interference.

6 An Application: The Flip–Flop Danger

A *logic* assigns a set of consequences to every premise set.[25] A logic **L** is *adaptive* if it adapts itself to the specific premises to which it is applied. One way to explicate this phrase is by saying that there are rules R such that **L** does not validate R, but **L** validates some applications of R to some premise sets. Adaptive logics were developed with the aim to obtain precise formulations of defeasible reasoning forms. These reasoning forms are methodological, rather than logico-deductive.

This is not the place to present an introduction to adaptive logics – many survey papers are available and a new state-of-the-art overview is on its way [11, 12, 14]. I shall merely present an example to give the reader a feel of adaptive logics. The example will allow me to point out a problem that is solved by the new type of many-valued semantics.

Let $\Gamma_1 = \{p, q, \neg p \vee r, \neg q \vee s, \neg q\}$. Note that $\Gamma_1 \nvdash_{\mathbf{CLuN}} s$ and $\Gamma_1 \nvdash_{\mathbf{CLuN}} r$. Yet, there is a clear difference between p and q. While Γ_1 requires that q behaves inconsistently, it does not require that p behaves inconsistently. This holds intuitively and **CLuN** leads to exactly the same insight: $\Gamma_1 \vdash_{\mathbf{CLuN}} q \wedge \neg q$ whereas $\Gamma_1 \vdash_{\mathbf{CLuN}} p$ but $\Gamma_1 \nvdash_{\mathbf{CLuN}} \neg p$. The idea behind inconsistency-adaptive logics is to interpret premise sets as normal as possible, where normality is obviously connected to consistency. Interpreting Γ_1 as normally as possible entails that q is considered as inconsistent whereas p and many other formulas are considered as consistent. Doing so results in r and not in s. Indeed, $\Gamma_1 \vdash_{\mathbf{CLuN}} r \vee (p \wedge \neg p)$ and $p \wedge \neg p$ is considered to be false; $\Gamma_1 \vdash_{\mathbf{CLuN}} s \vee (q \wedge \neg q)$ but $q \wedge \neg q$ is true anyway. Precisely this result is delivered by the inconsistency-adaptive logic $\mathbf{CLuN}^m$.

A crucial component of adaptive logics is their set of abnormalities. In the case of $\mathbf{CLuN}^m$, this set is $\{\exists(A \wedge \neg A) \mid A \in \mathcal{F}\}$, the existential closure of contradictory formulas. As suggested before, the idea is that, if B is an abnormality and A is not, if $A \vee B$ is **CLuN**-derivable, and if the premises allow one on systematic and formal grounds to consider B as false,[26] then A is an adaptive consequence, in this case a $\mathbf{CLuN}^m$-consequence.

Now consider another inconsistency-adaptive logic, $\mathbf{CLuNs}^m$. One difference is that **CLuN** is replaced by **CLuNs**. This has the effect that all **CLuNs** consequences are derivable from the premises independently of the fact that one aims at a maximally normal interpretation. Thus $\Gamma_1 \vdash_{\mathbf{CLuNs}^m} \neg\neg q$, whereas $\Gamma_1 \nvdash_{\mathbf{CLuN}^m} \neg\neg q$.[27]

In defining $\mathbf{CLuNs}^m$, one might be tempted to consider the same set of abnormalities as in the case of $\mathbf{CLuN}^m$, viz. $\{\exists(A \wedge \neg A) \mid A \in \mathcal{F}\}$. If one does so, however, one looses the adaptive effect. Indeed, although $\Gamma_1 \vdash_{\mathbf{CLuNs}} r \vee (p \wedge \neg p)$ and $\Gamma_1 \nvdash_{\mathbf{CLuN}} p \wedge \neg p$, one cannot simply consider p as behaving consistently on Γ_1. This is so because $\Gamma_1 \vdash_{\mathbf{CLuN}} (p \wedge \neg p) \vee ((r \wedge q) \wedge \neg(r \wedge q))$, whereas $\Gamma_1 \nvdash_{\mathbf{CLuN}} (r \wedge q) \wedge \neg(r \wedge q)$. So $(p \wedge \neg p) \vee ((r \wedge q) \wedge \neg(r \wedge q))$ is a minimal disjunction of abnormalities that is **CLuNs**-derivable from Γ_1. One of the disjuncts is false, but we do not know which one and cannot decide on logical grounds for one or for the other *if* both disjuncts count as

[25] This weak characterization is preferable in order to avoid prejudged narrowing of the domain. It is equivalent to the characterization offered by Béziau [18].
[26] The matter is handled by an adaptive strategy – see the referred survey papers.
[27] If this sounds puzzling, please realize that $\{q, \neg q, \neg\neg q, \neg\neg\neg q\}$ is more inconsistent than $\{q, \neg q, \neg\neg\neg q\}$ and that **CLuN** does not validate $A \supset \neg\neg A$.

abnormalities. It can be shown that this reasoning can be generalized. If $\mathbf{CLuNs}^m$ is given $\{\exists(A \wedge \neg A) \mid A \in \mathcal{F}\}$ as its set of abnormalities, the consequences derivable by $\mathbf{CLuNs}^m$ from any inconsistent premise set Γ are identical to the consequences derivable by **CLuNs** from Γ. Such an adaptive logic is called a flip–flop: its consequence set is identical to the **CL**-consequence set if Γ is consistent and to the **CLuNs**-consequence set otherwise.[28]

This does not mean that a decent inconsistency-adaptive logic $\mathbf{CLuNs}^m$ cannot be defined. Such a logic is obtained, e.g., by defining $\{\exists(A \wedge \neg A) \mid A \in \mathcal{F}^a\}$ as the set of abnormalities, $\mathcal{F}^a$ being the set of atomic formulas of $\mathcal{L}$. In this case $p \wedge \neg p$ can be considered as false and r will be a $\mathbf{CLuNs}^m$-consequence of Γ_1.[29]

Abnormality is used here as a technical term. There is obviously a relation to the intuitive sense of the term. The intuitive sense may be seen as defined by **CL**.[30] Every inconsistent theory is *intuitively abnormal*: it has no **CL**-models. This justifies the choice of $\{\exists(A \wedge \neg A) \mid A \in \mathcal{F}^a\}$ as the set of abnormalities for the inconsistency-adaptive logic $\mathbf{CLuNs}^m$. Indeed, if Γ is inconsistent, then there is an $A \in \{\exists(A \wedge \neg A) \mid A \in \mathcal{F}^a\}$ such that $\Gamma \vdash_{\mathbf{CLuNs}} A$. So if Γ is normal in the intuitive sense, then it is $\mathbf{CLuNs}^m$-normal; and vice versa.

The handiest way to describe gluts and gaps is available when all classical logical symbols are present in the language. A negation gap will, for example, be described by $\check{\neg} A \mathbin{\check{\wedge}} \check{\neg}\neg A$, in which the "checked" symbols have their **CL**-meaning.[31] Similarly, a disjunction glut may be described by $\check{\neg} A \mathbin{\check{\wedge}} \check{\neg} B \mathbin{\check{\wedge}} (A \vee B)$ or, more transparently, by $\check{\neg}(A \mathbin{\check{\vee}} B) \mathbin{\check{\wedge}} (A \vee B)$.

A possible source of confusion should be clarified here. If $v_M(\neg p) = u$ in a tuaf-semantics, then $M \Vdash \check{\neg} p \wedge \check{\neg}\neg p$ – the classical negation of p is true while its standard negation is false. If $v_M(\neg p) = a$ in a tuaf-semantics, then $M \Vdash \check{\neg}\check{\neg} p \wedge \neg p$ – the classical negation of p is false while its standard negation is true.[32] Similarly for other logical symbols. So it is important to distinguish between the formula A that "displays" a glut or gap in a model M, whether on the agreement view or on the interference view, and the formula that "describes" the glut or gap and is verified by M. Let the $(A)^a$, respectively $(A)^u$, be the formula that "describes" that A "displays" a gap, respectively a glut.

An adaptive logic that is a flip–flop may be turned into a non-flip–flop by weakening its set of abnormalities. There is, however a danger to that as well. The danger may be illustrated by considering the inconsistency-adaptive logic – call it **X** – that is obtained by replacing the set of abnormalities of $\mathbf{CLuN}^m$ by the set $\{\exists(A \wedge \neg A) \mid A \in \mathcal{F}^a\}$. Consider the simple but explicit premise set $\Gamma_2 = \{p \vee ((q \wedge r) \wedge \neg(q \wedge r))\}$. The fact is that $\Gamma_2 \nvdash_{\mathbf{X}} p$ and the reason is that there are no abnormalities $A_1, \ldots, A_n \in \{\exists(A \wedge \neg A) \mid A \in \mathcal{F}^a\}$ such

[28] In some exceptional cases, one wants an adaptive logic that is a flip–flop.

[29] The formula $(r \wedge q) \wedge \neg(r \wedge q)$ is not an abnormality but it is a $\mathbf{CLuNs}^m$-consequence of Γ_1. Indeed, $q \wedge \neg q$ is a **CLuNs**-consequence (and hence a $\mathbf{CLuNs}^m$-consequence) of Γ_1, r is a $\mathbf{CLuNs}^m$-consequence of Γ_1, and $q \wedge \neg q, r \vdash_{\mathbf{CLuNs}} (r \wedge q) \wedge \neg(r \wedge q)$.

[30] Still and to the best of my knowledge, paraconsistent logics that do not allow for other gluts or gaps and were proposed to serve a sensible purpose agree with **CL** in classifying Γ as inconsistent.

[31] The "checked" symbols are metalinguistic names for certain symbols of the language $\mathcal{L}$ of logic **L**. If the $\wedge$ is a classical conjunction in **L** and $\sim$ is a classical negation in **L**, then the formula in the text stands for $\sim A \wedge \sim\neg A$.

[32] Note that $M \Vdash p \wedge \neg p$ will do just as good.

that $\Gamma_2 \vdash_{\mathbf{CLuN}} p \vee (A_1 \vee \cdots \vee A_n)$. So here lurks a different danger: that the technical sense of abnormality is too weak with respect to the intuitive sense of abnormality, whence some consistent sets, like Γ_2, are not assigned all **CL**-consequences. If the aim is to interpret theories as consistently as possible in the sense of **CL**, then the inconsistency-adaptive logic **X** is clearly too weak.

So while there is, on the one hand, the flip–flop danger, there is, on the other hand, (what may be called) the wimp danger. It is typical for adaptive logics that both extending the set of abnormalities and reducing it may lead to a weakening of the consequence set. The matter is too complex to fully discuss it here, but the use of the many-valued logics from the previous section may still be illustrated. They do not offer a single criterion, but rather two criteria. Moreover, they obviously can only be applied if the considered logics have a tuaf semantics. Finally, I cannot show that the (partial) criteria in the subsequent paragraph are correct because this requires much more technical information on adaptive logics – people familiar with adaptive logics will find the matter rather transparent.

Let **L** be the deductive logic that underlies the adaptive logic – like **CLuN** and **CLuNs** in the previous examples – and let the semantic phraseology refer to the tuaf semantics of **L**. Let the adaptive logic be $\mathbf{L}^m$ and $\Omega \subset \mathcal{W}$ its set of abnormalities.

(a) $\mathbf{L}^m$ is not a flip–flop if, for every $A \in \mathcal{W}$, (i) $(A)^a \notin \Omega$ if there is a **L**-tuaf-model M such that $v_M(A) = t$ and $M \Vdash (A)^a$ and (ii) $(A)^u \notin \Omega$ if there is a **L**-tuaf-model M such that $v_M(A) = f$ and $M \Vdash (A)^u$.[33]
(b) $\mathbf{L}^m$ is not a wimp if, for every $A \in \mathcal{W}$, (i) $(A)^a \in \Omega$ if there is a **L**-tuaf-model M such that $v_M(A) = a$ and $M \Vdash (A)^a$ and (ii) $(A)^u \in \Omega$ if there is a **L**-tuaf-model M such that $v_M(A) = u$ and $M \Vdash (A)^u$.[34]

The main antecedent of (a) is not fulfilled for any flip–flops but is fulfilled for some wimps. The main antecedent of (b) is not fulfilled for any wimps, but is fulfilled for some flip–flops. If it sounds confusing, realize that (a) and (b) delineate extremes of the sets of abnormalities, not of the adaptive consequence sets.

7 Some Reflections

The tuaf semantics introduced in Sect. 5 provides insights that are useful independent of the flip–flop problem.[35] The tuaf semantics delineates the points at which inconsistencies originate – similarly for other nonstandard features, but I shall continue to concentrate on inconsistencies. So the information provided by a tuaf semantics is very valuable for

[33] The main antecedent warrants that there are **L**-models M and M' of intuitively abnormal Γ such that $\{A \in \Omega \mid M \Vdash A\} \subset \{A \in \Omega \mid M' \Vdash A\}$. That some **L**-models of intuitively abnormal Γ are not minimally abnormal **L**-models of Γ entails that $\mathbf{L}^m$ is not a flip–flop.
[34] The main antecedent warrants that every minimally abnormal **L**-model M of an intuitively normal Γ is such that $\{A \in \Omega \mid M \Vdash A\} = \emptyset$. This entails that, for all intuitively normal Γ, $\Gamma \vdash_{\mathbf{L}^m} A$ iff $\Gamma \vdash_{\mathbf{CL}} A$. So $\mathbf{L}^m$ is not a wimp.
[35] In connection with the flip–flop problem, the result may easily be generalized to, for example, modal logics. There are indeed adaptive logics in which abnormalities have the form $\Diamond A \wedge \neg A$ or the form $\Diamond A \wedge \Diamond\neg A$. This, however, should not be elaborated here.

comparing different paraconsistent logics and even for understanding specific paraconsistent logics separately. A typical difference between **CLuNs** and **CLuN** is that the former makes all inconsistencies dependent on inconsistencies in atomic formulas, whereas **CLuN** makes inconsistencies independent of each other.[36] While it is not difficult to understand the behavior of inconsistencies in those two logics, the matter is more difficult for other paraconsistent logics, such as $\mathbf{C}_1$ or $\mathbf{CLuN}^{NN}$ – and also for other nonstandard behavior as allowed by logics like **CLuCoDaM**. In all such cases, the tuaf semantics is definitely clarifying. Some readers may question the use of these many-valued logics in view of the fact that the information they provide may also be obtained (in a more laborious way) from the two-valued semantics. This objection does not hold water. If it did, it would just as well be an objection against usual many-valued semantic systems.

The main conclusion on semantic systems in general is that one should separate technical features from philosophical ones. That a logic has a many-valued semantics is a technical feature, and so is the fact that the logical symbols are or are not truth-functions in that semantics. Such technicalities do not determine the ontological structure of domains to which the logic may sensibly be applied. The same logic may very well agree with different ontological views and each of these may suggest a different set of valuation values. A nice example is that Priest's **LP** has a three-valued semantics in which all logical symbols are truth-functions, but that the ontology underlying this semantics is clearly at odds with Priest's dialetheism [34, §19.7] – see also footnote 15.

A logic **L** need not to be given an interpretation that agrees with an **L**-semantics in which all logical symbols are truth-functions. In some cases it is hard to imagine an interpretation that would go along with such a semantics – the $\mathfrak{S}$-valued and the $\mathfrak{T}$-valued semantics of **CLuN** are ready examples. Which is the set of truth-values, or more generally of valuation values, that statements may take, is a philosophical question. A sensible person might hold that there are three truth-values, say plain truth, plain falsehood, and inconsistency, and this person might want to allow for complex inconsistencies that have only consistent components, some true, some false. This person might end up with **CLuN** as her preferred logic and might end up with a three-valued **CLuN**-semantics, in which negation is not a truth-function, as the best way to picture the world's ontology.

The aim of this chapter was to raise questions, rather than to draw conclusions. The aim of the questions was to criticize prejudices, especially prejudices on many-valued logics, on the use or need to express the valuation values within the object language, on truth-functionality, and on the connection of all this to the semantics' ontological significance.

In a sense this chapter concerns consequences of Suszko's aforementioned result. If many-valued logics have a two-valued semantics and if this semantics, unlike the many-valued one, expresses the truth-preservation underlying the consequence relation, then the many-valued semantics is bound to serve a different purpose. But obviously there are several such purposes and these will lead to different many-valued semantic characterizations of the same logic.

Much work remains to be done in connection with the two preceding paragraphs. An obvious topic of research is the generalization of the $\mathfrak{S}$-semantics and of the $\mathfrak{T}$-semantics

[36] A conjunct of an inconsistency may be an inconsistency itself, as is the case for $(p \wedge \neg p) \wedge \neg(p \wedge \neg p)$. Even then the complex inconsistency is independent of the less complex one.

to other logics than **CLuN** and the study of the properties of logics and of classes of logics revealed by this generalization. Coding the information from a worlds-semantics into a many-valued semantics seems a closely related task. A very different topic concerns the articulation of many-valued logics originating from the agreement view, as opposed to the interference view, and the study of insights offered by both types of many-valued semantics. While **CL** was considered as the absolute point of reference in the present chapter, shifting to a different point of reference may have enlightening effects. All such research will help us, logicians, to overcome traditional prejudices and to better understand the aims, properties, and uses of logic at the service of reasoning and thus of understanding and action.

References

1. Anderson, A.R., Belnap Jr., N.D., Dunn, J.M.: Entailment. The Logic of Relevance and Necessity, vol. 2. Princeton University Press, Princeton (1992)
2. Arieli, O., Avron, A., Zamansky, A.: Maximal and premaximal paraconsistency in the framework of three-valued semantics. Stud. Log. **97**, 31–60 (2011)
3. Avron, A.: On an implication connective of RM. Notre Dame J. Form. Log. **27**, 201–209 (1986)
4. Avron, A.: Natural 3-valued logics—Characterization and proof theory. J. Symb. Log. **56** 276–294 (1991)
5. Avron, A.: Non-deterministic matrices and modular semantics of rules. In: Béziau, J.-Y. (ed.) Logica Universalis, pp. 149–167. Birkhäuser, Basel, Boston, Berlin (2005)
6. Avron, A., Ben-Naim, J., Konikowska, B.: Cut-free ordinary sequent calculi for logics having generalized finite-valued semantics. Log. Univers. **1**, 41–70 (2007)
7. Avron, A., Konikowska, B.: Multi-valued calculi for logics based on non-determinism. J. Interest Group Pure Appl. Log. **10**, 365–387 (2005)
8. Batens, D.: Paraconsistent extensional propositional logics. Log. Anal. **90–91**, 195–234 (1980)
9. Batens, D.: A bridge between two-valued and many-valued semantic systems: n-tuple semantics. Proceedings of the 12th International Symposium on Multiple-Valued Logic. IEEE: Los Angeles, 318–322 (1982)
10. Batens, D.: Inconsistency-adaptive logics. In: Orlowska, E. (ed.) Logic at Work. Essays Dedicated to the Memory of Helena Rasiowa, pp. 445–472. Physica Verlag (Springer), Heidelberg, New York (1999)
11. Batens, D.: A general characterization of adaptive logics. Log. Anal. **173–175**, 45–68 (2001), appeared 2003
12. Batens, D.: A universal logic approach to adaptive logics. Log. Univers. **1**, 221–242 (2007)
13. Batens, D.: Adaptive Logics and Dynamic Proofs. Mastering the Dynamics of Reasoning, forthcoming (201x)
14. Batens, D.: Tutorial on inconsistency-adaptive logics. In Béziau, J.-Y., Chakraborty, M., Dutta, S. (eds.) New Directions in Paraconsistent Logic, in print. Springer (2015)
15. Batens, D.: Spoiled for choice? J. Log. Comput., in print. doi:10.1093/logcom/ext019 (1913)
16. Batens, D., De Clercq, K.: A rich paraconsistent extension of full positive logic. Log. Anal. **185–188**, 227–257 (2004), appeared 2005
17. Batens, D., Mortensen, C., Priest, G., Van Bendegem, J.P. (eds.) Frontiers of Paraconsistent Logic. Research Studies Press, Baldock, UK (2000)
18. Béziau, J.-Y.: What is paraconsistent logic? In: Batens et al. [17], pp. 95–111.
19. Béziau, J.-Y.: Are paraconsistent negations negations? In: Carnielli, W.A., Loffredo D'Ottaviano, I.M. (eds.) Paraconsistency. The Logical Way to the Inconsistent, pp. 465–486. Marcel Dekker, New York (2002)

20. Carnap, R.: Meaning and Necessity. University of Chicago Press, Chicago (1947)
21. Carnap, R.: Logical Foundations of Probability. University of Chicago Press, Chicago (1950)
22. Carnielli, W.A., Marcos, J., de Amo, S.: Formal inconsistency and evolutionary databases. Log. Log. Philos. **8**, 115–152 (2001), appeared 2002
23. da Costa, N.C.: On the theory of inconsistent formal systems. Notre Dame J. Form. Log. **15**, 497–510 (1974)
24. D'Ottaviano, I.M.L.: Sobre uma Teoria de Modelos Trivalente (in Portuguese). PhD thesis, State University of Campinas (Brazil) (1982)
25. D'Ottaviano, I.M.L.: The completeness and compactness of a three-valued first-order logic. Proceedings of the 5th Latin American Symposium on Mathematical Logic, pp. 77–94. Revista Colombiana de Matemáticas, 1–2 (1985)
26. D'Ottaviano, I.M.L.: The model extension theorems for J3-theories. In: Di Prisco, C.A. (ed.) Methods in Mathematical Logic: Proceedings of the 6th Latin American Symposium on Mathematical Logic, Lecture Notes in Mathematics 1130, pp. 157–173. Springer (1985)
27. D'Ottaviano, I.M.L.: Definability and quantifier elimination for J3-theories. Stud. Log. **46**(1), 37–54 (1987)
28. D'Ottaviano, I.M.L., Epstein, R.L.: A paraconsistent many-valued propositional logic: J3. Rep. Math. Log. **22**, 89–103 (1988)
29. Esser. O.: A strong model of paraconsistent logic. Notre Dame J. Form. Log. **44**, 149–156 (2003)
30. Kleene, S.C.: Introduction to Metamathematics. North-Holland, Amsterdam (1952)
31. Kraft, V.: Der Wiener Kreis. Springer, Wien (1950)
32. Lev, I.: Preferential systems for plausible non-classical reasoning. Master's thesis, Department of Computer Science, Tel-Aviv University (2000), unpublished M.A. dissertation
33. Priest, G.: The logic of paradox. J. Philos. Log. **8**, 219–241 (1979)
34. Priest, G.: In Contradiction. A Study of the Transconsistent, 2nd expanded edn. (1st edn. 1987). Oxford University Press, Oxford (2006)
35. Rosser, J.B., Turquette, J.B.: Axiom schemes for m-valued propositional calculi. J. Symb. Log. **10**, 61–82 (1945)
36. Schütte, K.: Beweistheorie. Springer, Berlin (1960)
37. Smirnova, E.D.: An approach to the justification of semantics of paraconsistent logics. In: Batens, D. et al. [17], pp. 255–262
38. Suszko, R.: The Fregean axiom and Polish mathematical logic in the 1920s. Stud. Log. **36**, 377–380 (1977)
39. Verdée, P., Batens, D.: Nice embedding in classical logic, under review. Stud. Log., in print (2015)
40. von Wright, G.H.: Truth-logics. Acta Philos. Fennica **60**, 71–91 (1996)

D. Batens (✉)
Centre for Logic and Philosophy of Science, Ghent University, Ghent, Belgium
e-mail: Diderik.Batens@UGent.be

Logics and Their Galaxies

Hilan Bensusan, Alexandre Costa-Leite, and Edelcio Gonçalves de Souza

Abstract This chapter introduces some concepts that help exploring the ontological import of universal logic. It studies the notions of an *antilogic* and *counterlogic* associated with each logic and shows some of their properties. It presents the notion of *galaxy*, as the class of possible worlds compatible with a given logic. We explore some consequences of these developments.

Keywords Universal logic · Antilogics · Counterlogics · Galaxies

Mathematics Subject Classification (2000) Primary 03B22 · Secondary 03A99

1 Towards a Universal Metaphysics

Metaphysicians make constant use of logic. They need it to make claims about what is *logically* possible, what is *logically* necessary or contingent. They need it to reason on what ought to be the case once some assumptions are made – issues of *logical* compossibility. They need it whenever they investigate *a priori* knowledge and whenever they appeal to a sufficient reason. Without an underlying logic, no thesis concerning contingency, compossibility, necessity, or sufficient reason can be substantial. The use of logic by metaphysics is evident when we consider the framework of possible worlds – taken by David Lewis [11], pp. 1–4, to be a paradise for philosophers. At least since the Lewis work, back in the 1970s, metaphysicians have been making use of possible worlds to discuss and evaluate claims and arguments concerning several topics such as causation [9, 10], dispositions [13], monism and pluralism [4], semantics [16], or the existence of necessary objects [22]. The appeal to possible worlds, however, hinges always on a distinction between possible and impossible worlds (distinction drawn by Lewis himself [11], p. 1). A *logically* impossible world can arguably be dismissed from metaphysical considerations.

Now, there are different logics. In fact, an infinite number of them. Clearly, as metaphysics is entangled with logic, different underlying logics produce different results in metaphysics. Things are evaluated differently and distinct claims are made possible when we move from one logic to another. To see this, it is enough to focus on what makes

Work supported by a grant of the National Council for Scientific and Technological Development – CNPq – Brazil (486635/2013-9).

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_10

a world impossible. What is impossible in classical logic is not necessarily so in paraconsistent or intuitionistic logics. The question arises then as to which logic to choose. In general, metaphysicians have dealt with this in an easy and oversimplified way: just use classical logic. To be sure, nevertheless, the choice of a logic is a tricky issue that, at the face of it, involves circularity: we cannot argue that one particular logic is somehow necessary without appeal to an underlying logic. Given these difficulties, it is quite common to dismiss the question altogether by saying that in the absence of any reason to select any other logic, classical logic is best retained. Such a response, nevertheless, has been challenged by recent developments in universal logic.

Universal logic looks at the plurality of logics not in order to choose one among them but rather to study the relations between them. It explores how to compare them, how to put them together and how to build new logics out of the known ones. It aims at a general investigation of logics in a way that is similar to what has been achieved with universal algebra (see [1]). To this purpose, universal logic has been using methods such as the combination of logics, techniques of translating logics, and tools in category theory. It looks at the space of all logics and how it constrains the scope of a particular logic.

The present chapter explores the idea that universal logic – and not any logic in particular – should underlie the efforts of metaphysicians. It points towards a metaphysics grounded on universal logic – a *universal metaphysics*. We begin our exploration by considering possible worlds in a framework where different logics take different worlds to be impossible. In order to do that, we develop the notion of a galaxy compatible with a given logic – a galaxy is a class of possible worlds. It is a way to formalize the notion of underlying logic of a theory. If we are right, metaphysicians should be looking at galaxies and their properties whenever they make claims involving possibility, necessity, and related concepts.

2 Logics, Antilogics, and Counterlogics

In order to consider the infinite number of logics and to make evident how different logics yield different ontological commitments, we make use of the notions of *antilogic* and *counterlogic* defined for a suitable logic. Here we will not consider any logic in particular, but rather a logic in an abstract sense.

A *logic* L is a structure $(F, \vdash_L)$ such that F is a set and $\vdash_L$ is a binary relation on $\wp(F) \times F$ without any restriction.[1] We use $\Gamma \vdash_L \varphi$ to indicate that $(\Gamma, \varphi) \in \vdash_L$ and we say that φ is a *consequence* of Γ in L. So, $\vdash_L$ is called a consequence relation of L.

The *antilogic* of a given logic L is denoted by $\bar{L}$.[2] The *counterlogic* of a given logic L is denoted by $\tilde{L}$. Antilogics and counterlogics are also logics in the sense that they are sets and consequence relations without axioms. If we define logics using consequence relations with Tarskian conditions, then antilogics and counterlogics would not be logics.

[1] Tarski proposed an operator of logical consequence with conditions on it (see [20]). We use a consequence relation without restrictions in the sense of [2].

[2] The concept of *antilogic* has been developed by Łukasiewicz in [12], and studied in [17] as *rejected propositions* or in [18] under the label of *refutation systems*. For anticlassical propositional logic, there are results in [3], [6], and [21].

The antilogic $\bar{L}$ of a given logic $L = (F, \vdash_L)$ is a pair $(F, \vdash_{\bar{L}})$ such that

$$\Gamma \vdash_{\bar{L}} \varphi \text{ if and only if it is not the case that } \Gamma \vdash_L \varphi\,.$$

It is clear that for each L there is exactly one $\bar{L}$. Moreover, considering that consequence relations are sets of pairs, a natural consequence of the above definitions is that there is nothing in common between the set corresponding to a logic and the one corresponding to its antilogic:

$$\vdash_L \cap \vdash_{\bar{L}} = \varnothing;$$
$$\vdash_L \cup \vdash_{\bar{L}} = \wp(F) \times F\,.$$

This means that the two sets cover all the well-formed formulas expressed in the vocabulary where L (and of course, $\bar{L}$) is defined. In this sense (but not necessarily in another to be considered in the next section involving the concept of galaxies), we can say that $\bar{L}$ is the complement of L.

Let $L = (F, \vdash_L)$ be a logic such that F is closed by an unary operation of negation, that is, if $\varphi \in F$, then $\neg\varphi \in F$. In this case, we say that L is a *logic with negation*. Given a logic with negation $L = (F, \vdash_L)$, the $\neg$-*counterlogic*[3] of L is the logic given by the pair $\tilde{L} = (F, \vdash_{\tilde{L}})$ such that

$$\Gamma \vdash_{\tilde{L}} \varphi \text{ if and only if } \Gamma \vdash_L \neg\varphi\,.$$

Again, it is clear that for each L, and for each negation operation, there is exactly one $\tilde{L}$.

We already have elements to prove that soundness and completeness are preserved from a logic to its antilogic and its counterlogic.

Theorem 2.1 If $L = (F, \vdash_L)$ is sound and complete with respect to a semantic structure (which is also a logic in the sense above) $S = (F, \vDash_S)$, that is, if it holds that: $\Gamma \vdash_L \varphi$ if and only if $\Gamma \vDash_S \varphi$; then, $\bar{L}$ is sound and complete with respect to $S = (F, \vDash_{\bar{S}})$, such that $\vDash_{\bar{S}}$ is given by $\Gamma \vDash_{\bar{S}} \varphi$ if and only if it is not the case that $\Gamma \vDash_S \varphi$. A similar result can be stated to counterlogics.

As for interactions of a logic with negation, its antilogic and its counterlogic, it is important to note that

$$\bar{\bar{L}} = L\,;$$
$$\tilde{\tilde{L}} = L\,;$$
$$\bar{\tilde{L}} = \tilde{\bar{L}}\,;$$
$$\Gamma \vdash_{\bar{L}} \varphi \text{ if and only if it is not the case that } \Gamma \vdash_{\tilde{L}} \neg\varphi\,;$$
$$\Gamma \vdash_{\tilde{L}} \varphi \text{ if and only if it is not the case that } \Gamma \vdash_{\bar{L}} \varphi\,.$$

[3] Henceforth, whenever nonambiguous we use simply "counterlogic".

Note also that, as the antilogic and the counterlogic are defined in terms of the consequence relation, a logic that entails nothing ($\vdash_L = \varnothing$) is its own counterlogic. Moreover, its antilogic is the trivial logic ($\vdash_L = \wp(F) \times F$).

3 Antilogic, Counterlogic, and the Square of Oppositions

The antilogic and the counterlogic of a logic can define interesting oppositions concerning relations between logics. These oppositions can be thought in terms of the framework of the square of oppositions. Such a framework (and its extensions in the so-called geometry of oppositions (see [14]) has been used to study all sorts of relations of opposition. Different logics are such that they can hold relations analogous to those of contradiction, contrariety and subalternity between them. In this section, we begin to study these relations of oppositions by seeing different logical systems as poles in a triangle or a square. By presenting a square of logics that capture some of their relations, we pave the way for a more systematic study of the oppositions between logics.

We start with some brief remarks concerning the square. We call a square *complete* if it is a square with all four oppositions: contradiction, contrariety, subcontrariety, and subalternation. A square is *standard* if it fits any family of concepts satisfying traditional oppositions. A square is *perfect* if it is complete and standard. Moreover, any square which is not complete or/and standard is called *degenerate square*.

We can now explore the oppositions between a logic and its counter and antilogic in terms of the square. The main question is to determine what are the oppositions between these logics.

Because the vertices of the square (or of the triangle formed by a logic, its counterlogic and its antilogic) are not propositions, we reconstruct the classical oppositions accordingly. We define them in terms of relations between logics – instead of truth-values.

First, let $L_1 = (F, \vdash_{L_1})$ and $L_2 = (F, \vdash_{L_2})$ be two logics. We say that L_1 is a *sublogic* of L_2 if and only if $\vdash_{L_1} \subseteq \vdash_{L_2}$. We use the following notation: if $\vdash_L \subseteq \wp(F) \times F$, then $\nvdash_L$ is the complement of $\vdash_L$ with respect to $\wp(F) \times F$.

Now we introduce some relations between logics:

(d) L_1 and L_2 are *contradictories* if and only if $\vdash_{L_1} \cap \vdash_{L_2} = \varnothing$ and $\nvdash_{L_1} \cap \nvdash_{L_2} = \varnothing$;
(c) L_1 and L_2 are *contraries* if and only if $\vdash_{L_1} \cap \vdash_{L_2} = \varnothing$ and $\nvdash_{L_1} \cap \nvdash_{L_2} \neq \varnothing$;
(sc) L_1 and L_2 are *subcontraries* if and only if $\vdash_{L_1} \cap \vdash_{L_2} \neq \varnothing$ and $\nvdash_{L_1} \cap \nvdash_{L_2} = \varnothing$;
(s) L_1 is *subaltern* to L_2 if and only if L_2 is a sublogic of L_1.

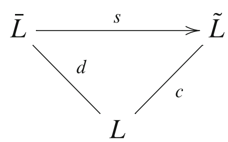
It is not straightforward to present oppositional structures for any logic. We will proceed by introducing some restrictions. First, we restrict ourselves to logics which accept elimination of double negation in an obvious sense. Additionally, let L be a logic with negation. We say that L is *well behaved* if and only if for every pair (Γ, φ), it is not the case that ($\Gamma \vdash_L \varphi$ and $\Gamma \vdash_L \neg\varphi$).

Theorem 3.1 If L is well behaved then $\tilde{L}$ is a sublogic of $\bar{L}$.

Suppose that $\Gamma \vdash_{\tilde{L}} \varphi$. By definition, $\Gamma \vdash_L \neg\varphi$. Since L is well behaved, this is not the case that $\Gamma \vdash_L \varphi$. Therefore, $\Gamma \vdash_{\bar{L}} \varphi$.

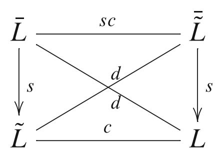
Let $L = (F, \vdash_L)$ be a logic with a negation. Consider $\varphi \in F$. We say that L is *φ-silent* if and only if we have neither $\varnothing \vdash_L \varphi$ nor $\varnothing \nvdash_L \varphi$. L is called *silent* if and only if there exists $\varphi \in F$ such that L is φ-silent. Note that if L is silent then its corresponding $\bar{L}$ is not well behaved.

We can start out by remarking that L and $\bar{L}$ are contradictories. Also, if L is well behaved and with double negation elimination rule, L and $\tilde{L}$ are contraries. Note that $\tilde{L}$ is then also well behaved. Since L is well behaved, $\tilde{L}$ is a sublogic of $\bar{L}$ and therefore the former is subaltern to the latter. We can then have a triangle of oppositions for a well-behaved L:



We can now consider a fourth vertex. Consider $\bar{\tilde{L}}$ which we have seen is equivalent to $\tilde{\bar{L}}$. Clearly, for any L that is well behaved, silent, and with the double negation elimination rule, so is $\tilde{L}$. Then we can see that if L is well behaved, $\tilde{\tilde{L}}$ is a sublogic of $\bar{\tilde{L}}$, by the theorem above. As $\tilde{\tilde{L}} = L$, then L is subaltern to $\bar{\tilde{L}}$ (and therefore to $\tilde{\bar{L}}$). Furthermore, if L_1 and L_2 are contraries, then $\bar{L}_1$ and $\bar{L}_2$ are subcontraries.

We can then present a square of oppositions for a well behaved, silent with double negation L:



This square of logics displays oppositions between any well behaved, silent, and accepting double negation logic on the one hand and three other logics constructed from it on the other. It is a complete square, but not a standard one and therefore it is a degenerate square. To be sure, it could be possible to find a perfect square of logics by extending differently the triangle above. However, the degenerate square is enough to show how any given well behaved and silent logic is accompanied by four other ones in relations to opposition. It is a square of logics showing how different (and yet related) logics are in relation to opposition.

4 Galaxies

The relations between any logic and its corresponding antilogic and counterlogic provides a convenient set-up to investigate the nature of logical necessity and therefore what counts as a possible world outside classical logic. It is clear that to each logic there is an associated class of possible worlds. To think of different logics in terms of possible worlds enable us to consider not only (accessibility) relations between different worlds made possible by a single logic but also relations between worlds made possible by different logics. The connection between a logic and its associated class of possible worlds began to be made explicit (and fruitful) by Kripke's semantics.

Indeed, in modal logic, it is customary to understand a logic in terms of the possible worlds associated with it (see [5]). In Kripke's semantics, different modal logics (primarily the normal ones **K**, **T**, **S4**, and **S5**) are each associated with the different accessibility relations between worlds and therefore to a class of possible worlds. What discriminates between those classes is the type of accessibility relations between worlds that they feature. In general, modal logics makes it clear that classes of possible worlds are associated to consequence relations. The more abstract idea is that a consequence relation can be understood in terms of an associated class of possible worlds. To these classes, we turn now.

Let $\mathcal{L}$ be a class of all logics. Consider, also, a class $\mathcal{W}$ which elements will be called *possible worlds*. Let $\Vdash$ be a binary relation in $\mathcal{W} \times \mathcal{L}$. We call this *compatibility relation*. When $w \Vdash L$ we say that the possible world w is *compatible* with the logic L.

Let G be a function (of classes) defined as

$$\begin{array}{rccl} G: & \mathcal{L} & \rightarrow & \wp(\mathcal{W}), \\ & L & \mapsto & G(L) := \{w \in \mathcal{W} : w \Vdash L\}. \end{array}$$

Therefore, given a logic L, $G(L)$ is the set of possible worlds compatible with L. $G(L)$ is called the *galaxy* of L. Note that for each logic L associated to a single consequence relation there is one and only one $G(L)$. A galaxy can be seen as the ontological counterpart of a logic.

The following condition on the compatibility relation has to be satisfied:

Sublogic Condition (SC): Consider that L_1 is a sublogic of L_2, if $w \Vdash L_1$, then $w \Vdash L_2$.

Further, we can also introduce an extra condition relating a logic and its antilogic:

Strong Constraint Condition (SCC): For every $w \in \mathcal{W}$, $w \Vdash L$ if and only if it is not the case that $w \Vdash \bar{L}$.

Under SCC, we have

$$G(L) \cap G(\bar{L}) = \varnothing;$$
$$G(L) \cup G(\bar{L}) = \mathcal{W}.$$

SCC, though, could be too strong for it fails to capture the standard way to talk about possible worlds. To see this, consider the following toy example. Let $L = (F, \vdash_L)$ be a logic, such that $F = \{\alpha, \beta\}$. Suppose further that $\varnothing \vdash_L \alpha$ and it is not the case that $\varnothing \vdash_L \beta$. Therefore, we have $\varnothing \vdash_{\bar{L}} \beta$ and it is not the case that $\varnothing \vdash_{\bar{L}} \alpha$. Let $w_1 \in \mathcal{W}$ and $w_2 \in \mathcal{W}$ be possible worlds such that α and β are the case in w_1, and α and β are not the case in w_2. It would be reasonable to consider that for all $w \in \mathcal{W}$, $w \in G(L)$ if and only if for any $\varphi \in F$, $\varnothing \vdash_L \varphi$ implies that φ is the case in w. Thus, in our toy example, $w_1 \in G(L)$ and $w_2 \notin G(L)$ while $w_1 \in G(\bar{L})$ and $w_2 \notin G(\bar{L})$. In other words, $G(L) \cap G(\bar{L}) \neq \varnothing$, and $G(L) \cup G(\bar{L}) \neq \mathcal{W}$.

It can be therefore natural to weaken the constraint on the compatibility relation. An alternative condition would then be

Weak Constraint Condition (WCC): Given a logic L and its antilogic $\bar{L}$, for every $w \in W$, it is not the case that ($w \Vdash L$ if and only if $w \Vdash \bar{L}$).

It is interesting to consider how the galaxy of a logic and of its counterlogic relate. It could be that

$$G(L) \cup G(\tilde{L}) \neq \mathcal{W}\,.$$

Neither a logic and its respective counterlogic can cover all formulas where the logic is defined nor their respective galaxies can cover all worlds. Also, it could be that

$$G(L) \cap G(\tilde{L}) \neq \varnothing\,.$$

When this is the case, this is a remarkable intersection class for it is composed of only inconsistent worlds. All logics that admit inconsistencies are such that their galaxies are in this intersection.

So, let a class of all inconsistent logics be defined as such

$$\varDelta = \{L \in \mathcal{L}\colon \ \text{there is a } \varphi \text{ such that } \ \varnothing \vdash_L \varphi \ \text{ and } \ \varnothing \vdash_L \neg\varphi\}\,.$$

If so we can say that for any L

$$G(L) \cap G(\tilde{L}) \subseteq G(\varDelta) := \bigcup\{G(L)\colon L \in \varDelta\}\,.$$

In other words, the intersection between the galaxy of a logic and of its counterlogic is a subclass of the galaxy of all inconsistent logics. The intersection is always inconsistent. We call a world w such that $w \in \varDelta$ a *dialethea*, or an inconsistent world.[4] It follows that

$$\text{for each } \ w \in G(\varDelta) \ \text{ there is an } L \text{ such that } \ w \in G(L) \cap G(\tilde{L})\,.$$

These prospective remarks can be viewed as the starting point of an *algebra of galaxies* that is neither reducible nor isomorphic to the algebra of logics (or the algebra of consequence relations (see [19])). Such an algebra is a tool to study how different logics intersect each other. Also, operations on galaxies reflect operations on logics – such as fusion (in the sense of combination of modal logics (see [8])). So, we have

$$G(L_1 \oplus L_2) = G(L_1) \cup G(L_2)\,.$$

The algebra is a useful tool to look at the ontological issues raised by universal logic.

[4] The term is used here in a way that is reminiscent but not equal to that used by Priest (see [15]). He takes a dialethea to be a part of reality which harbours inconsistencies, and true contradictions. We take it to be simply an inconsistent world. It is also worth noticing that if L is such that it is not the case that $\varnothing \vdash_L \varphi \wedge \neg\varphi$ then $\varnothing \vdash_{\bar{L}} \varphi \wedge \neg\varphi$. Therefore $\bar{L}$ could contain dialetheas (true contradictions) also in Priest's sense.

5 Prospects for a Metaphysics of Galaxies

We can then say that a world is possible only with respect to a galaxy. This has an impact on the notion of *logical truth* – or necessity simpliciter as it is classically conceived given that (logical) truths are relative to galaxies. When we say that something is the case in all possible worlds, we ought to specify a galaxy in which these possible worlds are. We can, however, say some general things through the framework of galaxies. A relevant general result is that for any φ, if φ is contingent in a galaxy, there is a galaxy in which it is necessary.

A notion directly affected by understanding a possible world as relative to a galaxy is that of necessity. Either reduced to a single type (broadly constructed as metaphysical necessity or narrowly conceived as logical necessity) or as a multiplicity (where natural or normative necessity cannot be understood in terms of logical or metaphysical necessity,[5] necessity is understood in terms of classes of possible worlds. Now, it seems that it follows from consideration on galaxies that necessity is relative to a place in the space of galaxies. There is no such thing as absolute necessity. How would a metaphysics without a notion of absolute necessity look like?

This is the main metaphysical open problem raised by *galaxy theory*, i.e., the study of galaxies. It bifurcates into two different issues: the nature of modal claims under the framework of galaxies and the ontological status of galaxies. As for the first issue, modal claims in general appear to have to be indexed by galaxies. These claims become relative to a class of worlds. It is interesting to remark that the relative character of necessity entails also the relativity of contingency to galaxies. At first sight, a metaphysics that takes into consideration different galaxies of possible worlds is close to a Humean approach in that it exorcises necessity; and in fact it is more extreme than typical Humean approaches as it dismisses *logical* necessity as well. It would be a metaphysics of absolute contingency. However, nothing happens contingently in all galaxies. There is no such thing as (logical) contingency simpliciter. If there is nothing but things that are contingent on something else, maybe other modal notions such as dependency and compossibility should rather be taken as basic. Those notions are relational ones and make explicit that modal ties are not absolute and have to be placed in the space of galaxies. Given a galaxy (or a collection of galaxies), we can study necessary (or accidental) connections within it. As for the overall picture where we ask questions about, say, the contingency of it all, they would have to be fully reformulated.

As for the second issue, we can make some sketchy remarks about the reality of galaxies. David Lewis has championed a version of modal realism [11] that maintains that all possible worlds exist – the actual one being no more than the one which happens to be this. There is no special property that renders any world actual, it is a world among others and it is actual for those who are in it. Actuality is indexical while existence is shared by all worlds (within a chosen galaxy – the classical one). From the point of view of galaxies, for a world to exist, it is enough to be in the right galaxy. Now, given the plurality of galaxies, one could formulate a stronger version of modal realism that would include all

[5] Kit Fine [7], for example, considers three irreducible varieties of necessity: metaphysical necessity, natural necessity and normative necessity.

worlds across all galaxies. One could be a realist concerning worlds possible in all logics and therefore extend realism to galaxies. If galaxies are part of the furniture of the universe, then the plurality of logics – the concern of universal logic – becomes a plurality not only *about* reality but *in* reality. In Lewis' realism about possible worlds, each world is actual for its citizens. Analogously, a galactic (modal) realism would take the actual world as no more than our current address. But then further questions concerning the plurality of logics in reality could be raised. The issue can be framed in terms of where does this world belong within the space of galaxies.

The plurality of galaxies is the opening door to a metaphysics informed by universal logic. This plurality does not entail the relativity of all metaphysical truths but rather guides the metaphysical gaze towards this plurality. Kit Fine [7] has studied three options with respect to the tension between a notion of reality that is neutral, absolute, and coherent and the plurality of perspectives that arises from McTaggart's problems with the reality of tense. Those three alternatives could have counterparts in the context of the plurality of galaxies – and its connection to reality. The first would be to choose one among the many alternative galaxies and take it to be real. This would take reality not to be neutral and would embrace a galactical antirealism contrasting with the modal realism formulated above. The second option would be to consider that each galaxy has its own reality – reality would then be no longer absolute but relative to each galaxy. The third option would be to consider a broader reality – Fine labels it *über-reality* – that encompasses all galaxies, even though that reality would by no means exhibit coherence. Fine himself favours the equivalent of this third option and explores some of its consequences. Here again, the third option is more clearly in line with what we mean by a universal metaphysics oriented by a galaxy theory. A metaphysics guided by universal logics attempts to investigate this über-reality formed by the multiple galaxies and the way they relate to each other.

The focus on galaxies enables us to consider metaphysical issues beyond a parochial point of view on necessity and contingency. It goes along with universal logic for it springs out from the idea that each logic has a class of possible worlds associated to it. Galaxy theory is the ontological counterpart of universal logic. The latter is the study of all logics while the former is the study of their respective galaxies. It is also the study of all logics through its associated class of possible worlds. Furthermore, as we see it, it is not simply a subsidiary of the universal logic, but rather feeds its development and can guide its endeavours.

Acknowledgment Thanks to Arnold Koslow and Graham Priest for discussions concerning ideas of this paper.

References

1. Beziau, J.-Y.: Universal logic. In: Childers, T., Majer, O. (eds.) Logica'94 – Proceedings of the 8th International Symposium, Prague, pp. 73–93 (1994)
2. Beziau, J.-Y.: From consequence operator to universal logic: a survey of general abstract logic. In: Log. Univers.: Towards a General Theory of Logics. Birkhäuser (2005)

3. 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 Publication, London (2013)
4. Bohn, E.: Must there be a top level? Philos. Q. **59**(235), 193–201 (2009)
5. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press, Cambridge (2001)
6. CAICEDO, X.: A formal system for the non-theorems of the propositional calculus. Notre Dame J. Form. Log. **19**, 147–151 (1978)
7. Fine, K.: Tense and reality. In: Modality and Tense – Philosophical Papers. Clarendon Press, Oxford (2005)
8. Gabbay, D, Kurucz, A., Wolter, F., Zakharyaschev, M.: Many-dimensional Modal Logics: Theory and Applications. Studies in Logic and the Foundations of Mathematics. Elsevier, Amsterdam (2003)
9. Hitchcock, C.: Prevention, preemption, and the principle of sufficient reason. Philos. Rev. **116**(4), 495–532 (2007)
10. Lewis, D.: Causation. J. Philos. **70**, 556–567 (1973)
11. Lewis, D.: On the Plurality of Words. Blackwell, Oxford (1986)
12. Łukasiewicz, J.: Aristotle's Syllogistic from the Standpoint of Modern Formal Logic. Clarendon Press, Oxford (1951)
13. Martin, C.B.: Dispositions and conditionals. Philos. Q. **44**, 1–8 (1994)
14. Moretti, A.: Geometry for modalities? Yes: through n-opposition theory. In: Aspects of Universal Logic, vol. 17, p. 102–145. CdRS Université de Neuchâtel, Neuchâtel (2004)
15. PRIEST, G.: In Contradiction. Clarendon Press, Oxford (2006)
16. Stalnaker, R.: Mere Possibilities. Princeton University Press, Princeton (2012)
17. Słupecki, J., Bryll, G., Wybraniec-Skardowska, U.: Theory of Rejected Propositions I. Stud. Log. **29**, 75–123 (1971)
18. Skura, T.: A Refutation Theory. Log. Univers. **3**, 293–302 (2009)
19. Stone, M.: The theory of representation for Boolean algebras. Trans. Am. Math. Soc. **40**, 37–111 (1936)
20. Tarski, A.: Remarks on fundamental concepts of the methodology of mathematics. In: Universal Logic: An Anthology, translated by R. Purdy and J. Zygmunt. Birkhäuser (2010), Original: 1929
21. Varzi, A.: Complementary logics for classical propositional languages. Kriterion **4**, 20–24 (1992)
22. Williamson, T.: Modal Logics as Metaphysics. Oxford University Press, Oxford (2013)

H. Bensusan
A. Costa-Leite (✉)
Campus Universitário Darcy Ribeiro, Instituto de Ciências Humanas, Departamento de Filosofia, ICC Ala Norte, Universidade de Brasília, Brasilia, Brazil
e-mail: costaleite@unb.br

E. Gonçalves de Souza
Departamento de Filosofia, Universidade de São Paulo, São Paulo, Brazil

Can Identity Be Relativized?

Otávio Bueno

Abstract What is the scope of identity? The intuitive response is that it applies everywhere, since every object is identical to itself. In this sense, identity is a general concept. In this chapter, I argue for the generality of identity by critically examining some attempts at relativizing it. I argue that each alleged instance of relativization ultimately presupposes identity, and as a result, identity still stands – in all of its un-relativized glory.

Keywords Identity, Relativization, Quantification, Nonclassical, Logics

Mathematics Subject Classification (2000) Primary 03B05 · Secondary 03B60

1 Introduction

Identity seems to be a general concept, one that can be, and regularly is, applied to any domain, from ordinary concrete objects through theoretical constructs to abstract entities and structures. Is the *same* concept of identity used in all of these cases? The answer seems to be affirmative: for what could different concepts of identity possibly be? In order to determine the difference (or sameness) in such concepts, we would need to invoke the concept of identity – the very concept whose application conditions and extension we are trying to determine. In this sense, identity seems to be an absolute notion rather than a relative one. Identity's generality and absolute character seems to go hand in hand.

In contrast, there have been a series of significant challenges to the effect that the concept of identity is general (see, e.g., [4]). In this chapter, I examine some of the central arguments in support of the relativization of identity, and I try to resist them. The alternative, suggested here, is one in which identity is the most general concept, and there is a single identity concept that needs to be taken as primitive. (I call this a *full, un-relativized concept of identity*.) However, I stress, no substantive metaphysics is needed in order to make sense of this concept.[1]

[1] Considerations in support of the unsubstantial character of identity can be found in [1].

To Jean-Yves Béziau

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_11

2 Relativizing Identity?

Several arguments can be, and have been, invoked to relativize identity, that is, to suggest that identity is a relative concept, which depends on contextual considerations. I'll consider, in particular, three such arguments: (a) the definability of identity; (b) the dependence of identity on the underlying logic, and (c) identity's dependence on particular frameworks. Each of the arguments will be examined in turn.

(a) The definability of identity The first argument is that identity can be defined: either explicitly (depending on the underlying logic) or implicitly (via suitable postulates). And given that definitions depend on the underlying language and the available logical resources, by changing the language, the logical resources or both, one would thereby also change identity, which is, thus, relativized.

In the case of an explicit definition, if the underlying logic is second-order, the well-known argument goes, identity can be defined via Leibniz's two principles of identity: the identity of indiscernibles and the indiscernibility of identicals [9]. Taken together, the two principles state that

$$\text{(L)} \quad x = y \leftrightarrow \forall P(Px \leftrightarrow Py)\,,$$

where "P" ranges over properties. Given that identity is not invoked in the right biconditional, identity is then defined in terms of second-order quantification over properties, and the resulting definition, so goes the argument, is not circular. But this definition, of course, crucially requires second-order quantification. If such quantification were not available, one would be unable to define identity explicitly.

Alternatively, one could adopt a broader understanding of definition, in which identity is implicitly defined *via* suitable postulates ([4], p. 328, note 2). Among these postulates, the reflexivity and substitutivity of identity are obvious candidates (see [6], p. 95):

$$\text{(R)} \quad \forall x\ x = x \quad \text{(reflexivity)}\,.$$
$$\text{(S)} \quad \text{If } x = y\,, \text{ then } Bxx \rightarrow Bxy \quad \text{(substitutivity)}\,.$$

Reflexivity expresses the fact that each object is self-identical – a basic feature of identity that is taken to hold universally over all objects (however broadly they are taken to be). Substitutivity allows one to infer that the same objects have the same properties. In this case, although identity is not explicitly defined by these two conditions, (R) and (S), it is implicitly formulated as that which satisfies these postulates. Thus, given a broad enough understanding of definition, identity can still be (implicitly) defined.

However, it is now a matter of the underlying logic how such an implicit definition is ultimately implemented. Different strategies can be adopted to define identity implicitly, depending, once again, on the resources of the underlying language and logic. In this sense, even though, when defined, identity is taken to hold universally, the concept is now relativized to the particular logical and linguistic devices available to implement the implicit definition. As a result, identity is not taken as an absolute notion: it is relativized.

(b) The relativization of identity to logic One of the reasons that have been advanced for the relativity of identity is that identity depends on the underlying logic, not only in terms of the resources available to express, or to define, identity (as we have just seen), but also in terms of the properties that identity is supposed to have (as will be discussed now).

Depending on the logic that one considers, different concepts of identity emerge, in the sense that different principles of identity are associated with the logic under consideration (see [4], pp. 329–330). Here are a couple of illustrations:

(i) In first-order logic, identity can be formulated – although not fully defined – in a slightly more direct way than the one discussed above, in terms of the following condition:

$$\text{(F)} \quad x = y \leftrightarrow (Px \leftrightarrow Py)\,,$$

in which x and y are arbitrary individual variables, and F is an arbitrary formula. Of course, given the limitations in expressive power of first-order logic, (F) does not express the condition that identical objects have the same *properties*. Strictly speaking, (F) only states that identical objects will satisfy the same *formulas*. But clearly this is not enough to guarantee the identity of the objects in question. After all, suppose that there are uncountably many properties.[2] In this case, we will not be able to express each one of them with a formula in first-order logic (since there are at most countably many such formulas). As a result, even if the right-hand bi-conditional in (F) were satisfied, this would not guarantee that the objects in question were indeed identical.

(ii) This problem, of course, will not emerge in second-order logic, given that one can then quantify over properties, and express directly Leibniz's principles (L), discussed above. These principles clearly provide a better way of formulating identity. But they also have significant limitations. Suppose that on the right-hand bi-conditional in (L), one quantifies over *all* properties of an object, and consider a concrete object, that is, an object located in spacetime. In two different instants in time, the object will no longer be the same, given the distinct properties it has. In fact, on this formulation, identity is extremely hard to preserve for spatiotemporal objects.

I am assuming here that it makes sense to talk of different instants in time in the context of Leibniz's principles (L). It may be argued that classical logic, whether in a first-order or in a higher order setting, is timeless, and so relativization to time is not something this logic can, strictly speaking, capture. On this view, any time relativization naturally requires some change in logic. This means that in order not to violate classical logic, one is

[2] The assumption is not unreasonable. Consider a physical magnitude, such as gravitational potential in field theories, and suppose that each value of this magnitude is a particular property. In this case, as one moves away from the source of the field, the field intensity decreases continuously. As a result, there would be uncountably many properties (see [10]). A far more controversial illustration would be this: Consider the properties *identical to the real number 0*, *identical to the real number 1*, *identical to the real number 2*, and so on. Note that each property of this kind is a property only of the corresponding number: *identical to 0* is a property of 0, and 0 alone; *identical to 1* is a property of 1, and 1 alone, etc. Since there are uncountably many real numbers, there are uncountably many properties. Identity already figures here, as it should, as an extremely general (relational) property; in fact, it is arguably the most general one. It may be complained that identity is not a property, but a relation of an object to itself and to no other object (see, for instance, [11]). But this clearly presupposes identity [2]: the reference to "no *other* object" requires that there is no object *distinct* (that is, *not identical* to) the one under consideration. And it is unclear what else identity could be.

not quantifying over *all* properties, but only over all *non-temporal* properties. As a result, the full force of the principles of identity is not properly captured, since what is required for objects to be the same is that they share *all* properties, rather than all properties of a *certain kind*. If, however, one restricts the properties that are quantified over – e.g., to essential properties, non-temporal properties, non-spatiotemporal properties – one would have, at best, a limited, restricted formulation of identity rather than identity *simpliciter* (for further discussion of this issue, see [3]).

(c) The relativization of identity to particular frameworks Identity is not only relative to the logic that is adopted in a given context, but it is also relative to the conceptual resources of the framework that is employed. Two illustrations can be offered of this phenomenon (see [4], pp. 329–330).

First, identity can be relativized to the particular kind of objects one considers, thus yielding a form of *objectual relativization*. In classical set theory, such as Zermelo–Fraenkel set theory with the axiom of choice (ZFC), to mention a well-known case, identity for sets is formulated in terms of the extensionality axiom:

$$\text{(E)} \quad x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y)\,.$$

This axiom expresses the fact that the identity of a set is determined only by its members: sets are the same just in case the same members constitute them. This is not a general characterization of identity, nor is it meant to be. But it provides a particular instance of identity to a given domain of objects: it is tailor-made to the relevant framework – not surprisingly, it applies particularly well to it. In this way, we have a particular kind of framework relativization of identity.

Second, identity can also be relativized to a given condition (formulated in terms of a collection of properties). Consider what Krause and Béziau ([4], p. 329) call the "conceptual relativization" of the principles of identity: x and y are identical relative to a collection of properties C just in case, for every property in that collection, x has that property as long as y has it, and vice versa. More formally, we have

$$\text{(C)} \quad x =_C y \leftrightarrow \exists C \forall P(P \in C \rightarrow (Px \leftrightarrow Py))\,.$$

This provides an additional illustration of the relativity of identity to a given framework; in this case, a conceptual relativization of the principles of identity to a certain condition C.

3 Resisting the Challenges

Can these challenges to the relativization of identity be resisted? I think they can, and I will examine each of them in turn.

(a) The undefinability of identity The alleged fact that identity can be defined provides no grounds to support the relativity of identity. First, it is not clear that identity can, in fact, be defined – either explicitly or implicitly. Any allegedly explicit definition of identity – for instance, in second-order logic – ultimately presupposes identity [1, 5], and [8]. The intelligibility of Leibniz's principles of identity presupposes that the *same* variables are used in each side of the two bi-conditionals:

$$\text{(L)} \quad x = y \leftrightarrow \forall P(Px \leftrightarrow Py)\,.$$

If the variables x and y, on the left side of the main bi-conditional, are not the *same* as the corresponding variables on the right, the resulting statement will not express the intended meaning: objects x and y are not the same if some object z, distinct from them, has the properties that are quantified over. Moreover, if the variable P is not the *same* on both sides of the right bi-conditional, the resulting statement will also fail to be adequate, since two objects are not the same if they have different properties.

It may be objected that, in this case, identity is only required in the metalanguage in order to express Leibniz's principles. Identity is nowhere to be found, as a primitive concept, in the object language. In particular, it does not figure in the right side of (L)'s main bi-conditional in terms of which the concept is defined. Thus, no circularity is, in fact, involved.

In response, the fact that identity is needed in the metalanguage clearly indicates that the concept is ultimately presupposed (see also [1]). For it is the intelligibility of the *object-language expression* that is at stake. If the *same* variables are not used on both sides of the bi-conditionals in Leibniz's principles, the statement fails to express the intended concept. As a result, Leibniz's principles can properly express the target concept only by presupposing the very notion that the principles are supposed to define, namely, identity. Although perhaps it is only in the metalanguage that identity is explicitly invoked, the fact remains that identity is presupposed in the object language. And this is enough to challenge the claim that identity has been successfully defined.

Moreover, even if the object language is sufficiently impoverished so that it is unable to express fully the concept of identity – such as a fragment of classical propositional logic – the notion of identity is still presupposed. Consider classical logic's conjunction-elimination rule that allows one to infer A from the conjunction $(A \wedge B)$. One needs to use the *same* variable A in the premise and in the conclusion of the rule in order to properly express it; otherwise, an invalid rule would be formulated: clearly C does not follow logically from $(A \wedge B)$ if C is different from both (and is not itself a logical truth). But this means that identity is presupposed even in logics that cannot express the notion (for further discussion, see [1]).

Second, even attempts at defining identity implicitly, *via* postulates, presuppose identity. Consider the reflexivity and the substitutivity of identity, (R) and (S), mentioned above. Both postulates clearly presuppose that *the same* variables are used. In the case of reflexivity – namely, $\forall x\ x = x$ – the same variable x needs to be in place on both sides of the identity sign. Similarly, in the case of the substitutivity of identity – that is, if $x = y$, then $Bxx \rightarrow Bxy$ – the same variables x and y also need to be invoked. In both cases, failure to do so would imply an inadequate formulation of the relevant postulates. But their proper formulation clearly *presupposes* the very concept that they attempt, albeit implicitly, to define.

(b) The nonrelativity of identity to logic It was argued above that identity is relative to the logic under consideration. Depending on the logic, the argument goes, identity is expressed differently – if it can be expressed at all. As a result, identity has different scopes: it is more restricted in some logics and less so in others.

Consider the formulation of the principle of identity in first-order logic:

$$\text{(F)} \quad x = y \leftrightarrow (Px \leftrightarrow Py)\,.$$

Is this a relativization of identity to a logic in which one is unable to quantify over properties? I don't think it is. First, note that identity is presupposed in the formulation of (F): the *same* variables x and y need to be quantified over on both sides of the main bi-conditional. Thus, rather than providing a first-order, restricted formulation of identity, (F) ultimately relies on the full, un-relativized concept in order to be properly expressed. Second, since one cannot quantify over properties in a first-order context, identity cannot be fully characterized. But the fact that it has not been thus characterized does not entail that identity has been somehow relativized. After all, as noted, the full concept is invoked to ensure the adequacy of (F), by guaranteeing that the *same* variables are used throughout. Therefore, identity is not restricted or relativized in (F): the un-relativized concept is employed, and full identity is ultimately invoked.

Now, in (F) an arbitrary propositional variable is free. Of course, given the quantificational resources of first-order logic, this variable cannot be quantified over. But what does it range over? There are two obvious answers: predicates (perhaps formulas) or properties. The former are linguistic items, the latter are features of the world. But none of these options works in this context. If the variable ranges over predicates, the content of (F) will fail to capture identity fully, since identity is supposed to be a relevant trait of objects rather than something limited to linguistic expressions. Furthermore, as is well known, there are countably many predicates in a language, but there are (at least) uncountably many properties. So, there is a mismatch between what is minimally required to formulate properly identity (that is, quantification over uncountably many properties) and the available expressive resources of first-order languages (that is, countably many predicates).

The same point goes through, *mutatis mutandis*, if the predicate variable in (F) is taken to range over formulas (see [4], p. 329), since there are at most countably many of them. Note that formulas don't seem to be the right items in this case. Consider the one-place formula "x is thought of by Frege." Presumably the identity of an object does not depend in general on the relations it can enter into. Thus, whether x is identical to y or not should not depend on whether Frege thought of x or y. But it would if the predicate variable in (F) ranged over formulas.

In contrast, let us suppose that in (F) the propositional variable ranges over properties. This option at least puts the proposal on the right track. However, even this would not do. After all, presumably (F) can only express identity if what it expresses is true. Since the propositional variable P in (F) is not quantified over, (F) is an open formula that lacks truth value. We can, of course, consider individual instances of (F) for particular values of P. But these instances do not have the quantificational force to express identity. In fact, as a definition of identity, any particular instance of (F) would be false: objects are not the same just in case they both satisfy a *single* property.

We thus have a dilemma: Either (F) has generality or it does not. If it does have generality, and thus at least in principle could range over a number of different properties, it lacks truth value; after all, in this case, (F) is an open formula. If (F) does not have generality, and a particular property is instantiated in each case, (F) will have a truth value, but each instance of (F), as noted above, will then be false, and thus inadequate as a formulation of identity. In either case, (F) fails to express identity properly. As a result, since identity has not been adequately expressed, it has not been relativized either. Any relativization of identity should be recognized, first, as identity. However, this is not the case, given (F)'s falsity in one horn and its lack of truth value in the other.

These objections would not emerge in the context of second-order logic, since quantification over properties is allowed. We thus have Leibniz's principles, discussed above:

$$\text{(L)} \quad x = y \leftrightarrow \forall P(Px \leftrightarrow Py)\,.$$

But we still do not have a relativization of identity. First, as noted above, identity is not defined by (L), since it is presupposed in the formulation of (L) itself. The expression, thus, fails to define properly the concept.

Furthermore, even if identity were somehow characterized by (L), it would not thereby be relativized. Although what can be expressed in a given formal language does indeed depend on the resources of that language, this fact does not entail that identity itself has been relativized to the language in question. Given the limitations in expressive power of any given language, a particular aspect of identity may have been formulated, but that is different from fully capturing a relativized concept of identity. Talk of a relativized concept suggests that identity – full identity – has been been formulated, and it turns out to be somehow relativized. But given that full identity is needed in order to make sense of any such relativization, it is unclear that any significant relativization has in fact been reached. A purely linguistic relativization, which simply acknowledges the fact that in a given language certain things cannot be expressed, is not philosophically significant. Such a linguistic relativization may well be right, but it is largely irrelevant. What would be relevant is a different concept of identity – a genuine concept of identity – that has in fact been relativized. But it is unclear that we can have that since expressions of that concept ultimately presuppose full identity, for the reasons discussed in this chapter, and thus the identity concepts at stake are not really rivals.

These points also apply to (L). As noted, its formulation does presuppose identity. Moreover, as also highlighted above, suppose that the properties that are quantified over in (L) are only of a certain kind: they may involve essential properties, temporal properties, spatiotemporal properties, etc. Since quantification is restricted to the relevant properties, even in the context of (L) one obtains at best a limited formulation of identity, which focuses on the particular properties that are quantified over. However, since (L) still presupposes the full concept of identity, no philosophically significant relativized notion emerges in the end.

(c) The nonrelativization of identity to particular frameworks Two forms of relativization of identity to a framework were considered above: the first involved the relativization to a certain kind of objects, the second focused on conceptual relativization (expressed in terms of the characterization of a domain *via* suitable properties). Neither form of relativization goes through, though – or, at least, none has the implication that identity has indeed been relativized.

Consider, first, objectual relativization. The extensionality axiom in set theory provides identity conditions for sets, that is, it specifies the conditions under which a set differs from other sets and the conditions under which it does not. This does not entail, however, that identity itself has been relativized. If sets are the same as long as they have the same members, the notion of identity that is invoked remains unaltered. In fact, had it been altered, the extensionality axiom would not have the content it is supposed to have, namely, that the identity of a set depends only on its members, and so any difference between two sets has to be traced back to a difference between the relevant members.

To see why this is the case, consider the notion of *difference* that is invoked. Given extensionality, two sets x and y can differ only if there is an object o that is a member of x but not of y, or vice versa. Thus, if o is a member of x, it needs to be *distinct* from each member of y; alternatively, if o is a member of y, it has to be distinct from each member of x. To be distinct, of course, just is not to be identical. And if this notion of identity is somehow relativized – in this case, to a particular set – the extensionality axiom will fail to guarantee that the identity of sets depends only on their members. After all, whether the object o is distinct from each member of y (or of x, depending on the case) would depend on whether o is itself a member of *some other* set, given the presumed relativization to a set. Thus, whether sets are identical or not would then depend not only on their members, but also on whether their members belong to some other set, which provides the presumed relativization of identity. But this is clearly against the extensionality axiom.

Moreover, in this case, we would also face a regress. After all, the set with respect to which identity is relativized should have well-defined identity conditions (so that it is determined whether o is itself a member of that set), but this presupposes the extensionality axiom. Since this axiom is taken, by hypothesis, as invoking a relativized notion of identity, whether an object is, or is not, a member of a given set ultimately depends on whether this object is a member of *some other set* with respect to which identity is relativized. This set, in turn, also needs to have well-defined identity conditions, which also presupposes extensionality, and a regress emerges. To block such a regress what is needed is a *nonrelativized* concept of identity – that is, one that is *not* dependent on any set. For in this case, two sets differ as long as there is an object which belongs to one of them and is *distinct* (that is, not identical) to any member of the other. (Note that identity is presupposed in order for an object not to be a member of a given set, since the object needs to be *distinct* from each member of that set.) Given that the difference between this object and the relevant members of the set does not depend on membership to any set, there's no need to invoke extensionality, and the regress is blocked.

The fact that such a regress does not emerge in the first place when classical set theory is developed already indicates that a nonrelativized concept of identity is ultimately in place. For these reasons, the very content of extensionality requires the nonrelativized concept of identity. But this means that this axiom does not relativize identity after all: it *presupposes* the usual, nonrelativized identity concept – on pain of not capturing the axiom's content.

Second, conceptual relativization is also not a relativization of the concept of identity. As we saw, x and y are identical relative to a collection of properties C if, and only if, for every property in that collection, x has that property if, and only if, y has it as well. This was then formulated in terms of the following equivalence:

$$\text{(C)} \quad x =_C y \leftrightarrow \exists C \forall P (P \in C \rightarrow (Px \leftrightarrow Py)) .$$

Identity is not relativized in this case. Expressed in (C) is not the concept of identity, but that of similarity (that is, identity with respect to some properties). And it's crucial that the underlying concept of identity, which is presupposed in (C), is itself not relativized. Otherwise, (C) will fail even to express similarity between the objects in question. Note, to begin with, that to ensure that the objects x and y are considered throughout, it is crucial that *the same* variables are invoke on both sides of the main bi-conditional in (C). Furthermore, since x and y are not required to share the same properties in general, but only those of a restricted class C, clearly these objects need not be identical. But their similarity – identity with respect to C – presupposes (nonrelativized) identity. After all, *all* those properties that are in the restricted class C are involved in establishing the relevant similarity between x and y, that is, any property *distinct* from those in C need not be shared by these objects. Otherwise, instead of similarity, we would have identity *via* Leibniz's principles, since the objects x and y would then have the same properties. Once again, identity is not relativized here: it is presupposed.

If identity is not relativized, does this mean that it is an absolutely unrestricted concept?[3] I don't think it does. We have here a pragmatic point. Similarly to any other concept, the use of identity presupposes a context in which it is applied. Relative to that context – which may be determined by a framework, by particular conditions, or by a given logic – identity seems to be relativized. This provides the appearance that the concept has indeed undergone some sort of relativization. In practice, however, appearances are deceptive, and the nonrelativized notion of identity – the single notion that is presupposed in each use – is invoked. In each particular instance, the same concept is employed: the one just referred to in this very sentence!

But, as with any quantificational use, it is always relative to a context that a particular application is articulated. In the very statement of the thesis that identity is a general concept – that is, one that is applied *everywhere* – a particular context is at stake, namely, the attempt at determining the scope of identity. And in different contexts, different considerations are put forward. As we saw, there are attempts at restricting identity to certain frameworks, or particular conceptual conditions, or given logics. Any individual application context of identity provides the relevant range of contrasts for the concept. The important point, however, is that, despite the different contexts, the same concept of identity is used, or presupposed, throughout. It is unclear that there is another concept in any case. After all, as argued above, in each alleged relativization of identity, the full, unrelativized concept of identity is ultimately presupposed.

This supports the conclusion that identity is an extremely general concept. However, this does not entail that identity is an absolutely unrestricted concept. Just as quantification is always restricted – a particular domain for the quantifier needs to be specified in order for the quantifier to be properly determined and to avoid potential incoherence – identity is similarly constrained. Despite its generality, identity is applied in particular contexts, and

[3] For discussions of absolute generality, see the papers in [7].

in different contexts different constraints emerge. These are constraints about the context (frameworks, conditions, logics), which restrict particular uses of the concept – despite the fact that, in each case, the concept is the same.

4 Conclusion

I examined three considerations to the effect that identity can be relativized: (a) Identity can be defined, and since definitions depend on the underlying logical resources, different concepts of identity emerge. (b) Identity is also relative to logic; hence, changes in logic entail changes in identity, which is, thus, relativized. (c) Finally, identity is relative to a framework, and as a result, another source of relativity of identity is in place. I argued that there are difficulties for each of these claims, and given the fact that full, un-relativized identity is presupposed throughout, it is unclear that identity has indeed been relativized. Identity is too basic a notion for that.

Acknowledgment My thanks go to Jean-Yves Béziau, Newton da Costa, Steven French, Décio Krause, and Alan Weir for extremely helpful discussions. It is a pleasure to offer this paper to Jean-Yves on the occasion of his 50th birthday. It still feels as though it was only yesterday that he arrived for the first time in Brazil to change so much of the logical landscape in the country – and elsewhere. It's surprising that two decades have already passed since then!

References

1. Bueno, O.: Why Identity is Fundamental. Am. Philos. Q. **51**, 325–332 (2014)
2. Deutsch, H.: Relative identity. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Winter 2008 Edition). (2008). http://plato.stanford.edu/archives/win2008/entries/identity-relative
3. French, S. Krause, D.: Identity in Physics: A Historical, Philosophical, and Formal Analysis. Clarendon Press, Oxford (2006)
4. Krause, D., Béziau, J.-Y.: Relativizations of the Principle of Identity. Log. J. Interest Group Pure Appl. Log. **5**, 327–338 (1997)
5. McGinn, C.: Logical Properties. Clarendon Press, Oxford (2000)
6. Mendelson, E.: Introduction to Mathematical Logic, 4th edn. Chapman and Hall, London (1997)
7. Rayo, A., Uzquiano, G. (eds.): Absolute Generality. Clarendon Press, Oxford (2006)
8. Savellos, E.: On Defining Identity. Notre Dame J. Form. Log. **31**, 476–484 (1990)
9. Shapiro, S.: Foundations without Foundationalism. Oxford University Press, Oxford (1991)
10. Swoyer, C., Orilia, F.: Properties. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Fall 2014 Edition). (2014). http://plato.stanford.edu/archives/fall2014/entries/properties
11. Zalabardo, J.: Introduction to the Theory of Logic. Westview Press, Boulder (2000)

O. Bueno (✉)
Department of Philosophy, University of Miami, Coral Gables, USA
e-mail: otaviobueno@me.com

From (Paraconsistent) Topos Logic to Universal (Topos) Logic

Luis Estrada-González

Abstract In this chapter, I describe how complement toposes, with their paraconsistent internal logic, lead to a more abstract theory of topos logic. Béziau's work in Universal Logic – including his ideas on logical structures, axiomatic emptiness and on logical many-valuedness – is central in this shift and therefore it is with great pleasure that I wrote this chapter for the present commemorative volume.

Keywords Standard topos · Complement-topos · Bare topos · (Bare) internal logic

Mathematics Subject Classification (2000) 03A05, 03B53, 03G30, 18B25

> The point of these observations is not the reduction of the familiar to the unfamiliar (...) but the extension of the familiar to cover many more cases.
> —Saunders Mac Lane, *Categories for the Working Mathematician*.

1 Introduction

Probably most readers of this *Festschrift* are familiar with Béziau's personal journey from paraconsistent logic to Universal Logic, as accounted for example in [7]. In this chapter, I want to explore a similar conceptual shift in the case of topos logic. I will show that the notion of complement topos, with its paraconsistent internal logic, plays a significant role in finding the truly universal, structural features of topos logic, since, in spite of what many category theorists think, until nowadays there are plenty of material, non-invariant or non-structural elements in topos logic.

The plan of the chapter, is as follows. In the next section, I will expound the basics of what I call *standard topos theory* and the view of the internal logic, or topos logic, arising from it. In Sect. 3, I will present complement toposes, the features of their internal logic and the exact elements of standard topos logic that they help to exhibit as non-structural, namely certain particular Skolemizations in the equational structure of a topos. The very existence of complement toposes runs against a theorem that is usually read as stating

To Jean-Yves Béziau for his 50th birthday, and also to Christian Edward Mortensen for his 70th birthday.

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_12

that the internal logic of a topos is in general intuitionistic. In Sect. 4, I examine the preconditions for proving that theorem, and that make its usual reading posible, and show that many of them are also non-structural elements from standard topos logic. This allows us to give, in Sect. 5, a more structural, invariant, purely equational formulation of topos logic, which closely resembles Béziau's notion of logical structure in his Universal Logic.

The reader is assumed to know classical logic and to understand first-order languages and naïve set-theoretic notation. Those are the prerequisites. A fluent reading presupposes the knowledge of some category theory, order theory and algebra.[1] There is a convention to keep on mind: I use the adjective 'categorial' exclusively used as shorthand for 'category theoretic', but note that this convention has not been applied to quotations, where 'categorical' is commonly used.

2 Basics of Standard Topos Logic

> When the main contradictions of a thing have been found,
> the scientific procedure is to summarize them in slogans
> which one then constantly uses as an ideological weapon
> for the further development and transformation of the thing.
> —William Francis Lawvere, 'Quantifiers and sheaves'.

2.1 Introduction: Toposes as Universes of Sets

A category can be thought of as a universe of *objects* and their transformations or connections, called *morphisms*, subject to some very general conditions. An example of a category is **Set**, whose objects are sets and its morphisms are functions between sets. In **Set** there is a special kind of objects, namely objects with two elements. As objects with two elements, all these objects are isomorphic to each and each of them has all and only those mathematical properties (as expressible in categorial terms) as any other, so the sign '$\mathbf{2_{Set}}$' can be used to denote any of them and speak as if there were only one of them. We will say that an object with the property of having exactly two elements is unique up to isomorphism. $\mathbf{2_{Set}}$ act as truth values object in **Set** in the sense that suitable compositions with codomain $\mathbf{2_{Set}}$ serve to expresses that certain sets are part of others. Hence, the two elements of $\mathbf{2_{Set}}$ are conveniently called $true_{\mathbf{Set}}$ and $false_{\mathbf{Set}}$.

Other logical notions besides truth values, such as zero- and highero rder connectives, can be defined in **Set**. It can be proved that the right logic to study the objects in **Set**, its *internal logic*, is that induced by the algebra formed by $\mathbf{2_{Set}}$ and the connectives, which turns out to be classical. This logic is called *internal* for two main reasons. First, because it is formulated exclusively in terms of the objects and morphisms of the category in question. Second, because it is the right logic to reason about the category in question since it is determined by the definitions of its objects and its morphisms, in a way that

[1] A good starting point are Chapters 1, 2 and 4 of [26].

using a different logic for that purpose would alter the defining properties of those objects and morphisms and thus it would not be a logic for the intended objects and morphisms; in other words, it cannot be a canon imposed "externally" to reason about the category.

As in usual axiomatic membership-based set theories like ZF(C), most of mathematics can be interpreted and carried out in **Set**. However, a set theory developed from a category-theoretic point of view is not based on the notion of membership, but rather on those of function and composition (of functions).

There are other **Set**-like categories, called *elementary toposes* or simply *toposes*. In a topos $\mathcal{E}$ there are objects which play the role of $\mathbf{2_{Set}}$ in the particular case of **Set**, i.e. they serve to express that certain objects are part of others via suitable compositions of morphisms. An object that plays such a role in a topos is also unique up to isomorphism and any of them can be denoted by '$\Omega_{\mathcal{E}}$' and speaks as if there were only one of them. Logical notions such as truth values and zero- and higher order connectives can also be defined in a topos. However, in general, $\Omega_{\mathcal{E}}$ has more than two elements and, since $\Omega_{\mathcal{E}}$ has all the same universal properties as $\mathbf{2_{Set}}$ and the latter can be considered a truth values object, so can the former. In addition, the logic appropriate for dealing with the objects and morphisms in a topos, its internal logic, is in general intuitionistic, not classical. This is precisely a logic arising from objects and morphisms themselves, not from our devices to reason about them. Like **Set**, toposes also allow for the interpretation of set theoretical notions and hence of significant parts of mathematics, but the reconstruction of mathematics carried out in a topos corresponds to mathematics as done in an intuitionistic set theory. If toposes can be considered universes of sets and, given that at least parts of mathematics can be reconstructed in a set theory, toposes also allow for the reconstruction of those parts of mathematics, then the universal laws of mathematics are those valid across all universes of sets, namely, the laws of intuitionistic logic.

2.2 Properties and a Comprehension Axiom

For our convenience, think of an object O of a topos as a type, collection of things, or generalized set – the O's.[2] Thus an object O is the objects of o's, in the same way that a product is the object of pairs $\langle x, y\rangle$ such that x is in X and y is in Y. The basic means of getting logic in a topos will be by a generalized notion of comprehension of subobjects by 'properties'. There are two things one needs to know about such properties:

Properties are local: A property is always a property of o's of some O; thus, every property has a fixed domain of significance.
Properties are variable propositions: If φ is a property with a domain of significance O, and a is a constant element of type O, then $\varphi(a)$ is a proposition.[3]

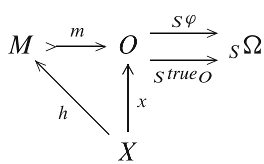
So in a topos a property with a domain of significance O will be called a *propositional function on O*. Every morphism must have a codomain, so a topos will include an ob-

[2] This elucidation of toposes in logical terms follows closely [1].
[3] As Awodey has noted, this is Russell's notion of propositional function, for example in *The Principles of Mathematics* § 22 or *Principia Mathematica*, pp. 14 and 161.

ject Ω of propositions or (algebraic) truth values. Its elements (if any) $p : \mathbf{1} \longrightarrow \Omega$ are propositions, and its generalized elements $\varphi : X \longrightarrow \Omega$ are variable propositions, hence propositional functions. If the proposition p factors as $p = \varphi(a) : \mathbf{1} \longrightarrow O \longrightarrow \Omega$, then p results from evaluating the propositional function φ for the element a of O. Like for every object of a category, the elements of Ω form a partial order, i.e. for any propositions p, q and r,

$$p \leq p\,;$$
$$\text{If } p \leq q \text{ and } q \leq p, \text{ then } p = q\,;$$
$$\text{If } p \leq q \text{ and } q \leq r, \text{ then } p \leq r\,.$$

The core assumption in standard topos theory – at least for the part concerning the theory of the internal logic of toposes, or *topos logic* – is that there is a proposition $true : \mathbf{1} \longrightarrow \Omega$ satisfying a certain comprehension principle. I will use a subscript S to denote all the morphisms, objects and constructions that depend on this assumption and I will explain later in more detail and more precisely the notation. Thus, such a proposition ${}_S true : \mathbf{1} \longrightarrow \Omega$ is said to satisfy the following **(Standard) Comprehension axiom**. For each ${}_S\varphi : O \longrightarrow {}_S\Omega$ there is an equalizer of ${}_S\varphi$ and ${}_S true_O$, and each monic $m : M \rightarrowtail O$ is such an equalizer for a unique ${}_S\varphi$. In diagrams, ${}_S true$ is such that for every ${}_S\varphi$ and every object T and morphism $o : T \longrightarrow O$, if $m \circ_S \varphi = m \circ_S true_O$ and $x \circ_S \varphi = x \circ_S true_O$, then there is a unique $h : X \longrightarrow M$ that makes the diagram below commutative:



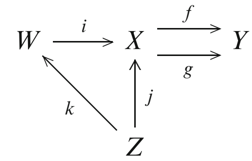
The propositional function ${}_S\varphi$ is also called ‘the (standard) characteristic (or classifying) morphism of m’, denoted by ${}_S\varphi_m$ for more convenience. A subobject classifier is unique up to isomorphism and so is ${}_S\varphi_m$. Now a topos can be defined more precisely: A category ${}_S\mathcal{E}$ with equalizers, (binary) products, coequalizers, coproducts, exponentials, and a (standard) subobject classifier is called *elementary (standard) topos*.[4]

Then, for any object O in a topos, the composite ${}_S true \circ !_O : O \longrightarrow \mathbf{1} \longrightarrow {}_S\Omega$ denotes a constant, ${}_S true$-valued propositional function on O, abbreviated as ${}_S true_O$. Propositional functions specify subobjects as follows. Given a propositional function $\varphi : O \longrightarrow {}_S\Omega$, one gets the part of the o’s of which ${}_S\varphi$ is true, if any, as an equalizer $m : M \rightarrowtail O$ of ${}_S\varphi$ and ${}_S true_O$. This subobject will be named accordingly the *extension* of the propositional function ${}_S\varphi$.

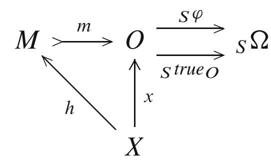
The connection of this *Comprehension axiom* with more traditional logical notions is much less mysterious than it might appear at first sight. Consider the diagram in the

[4] Note by the way that, unlike many authors, I prefer the equalizers presentation of logic, not the pullbacks one.

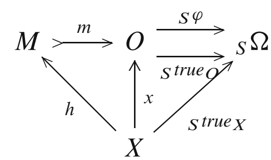
definition of an equalizer:[5]



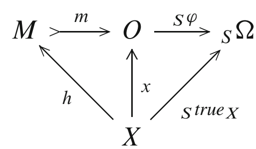
As a particular case for the *Comprehension axiom* one has



The only morphism from X to ${}_S\Omega$ that makes the diagram above commutative is ${}_S true_X$:

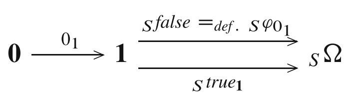


Thus, the following diagram is obtained:



Note that, according to the definition of an equalizer, h must be the only morphism that, among other properties, $x = m \circ h$. But this suffices to satisfy the categorial definition of $x \in m$. Hence, what the *Comprehension axiom* states is that ${}_S\varphi(x) = {}_S true_X$ (because of the right commutative triangle) if and only if $x \in m$ (because of the left commutative triangle).

Given the notion of a subobject classifier, one can also define ${}_S false : \mathbf{1} \longrightarrow {}_S\Omega$ as the character of 0_1, the only morphism from an initial object to a terminal one:



[5] Let $f : X \longrightarrow Y$ and $g : X \longrightarrow Y$ be morphisms in a category $\mathbf{C}$. An *equalizer* in $\mathbf{C}$ for f and g is given by an object W and a morphism $i : W \longrightarrow X$ in $\mathbf{C}$ with the following two properties: (1) $f \circ i = g \circ i$ and (2) for any morphism $h : Z \longrightarrow X$ in $\mathbf{C}$, if $f \circ h = g \circ h$, then there is exactly one morphism in $\mathbf{C}$ $k : Z \longrightarrow W$ such that $h = i \circ k$.

Example 2.1 Let **Set** be the (standard) category of (abstract constant) sets as objects and functions as morphisms. ${}_S\Omega_{\mathbf{Set}}$ has only two elements with the order ${}_Sfalse_{\mathbf{Set}} < {}_Strue_{\mathbf{Set}}$. Hence, in this category ${}_S\Omega_{\mathbf{Set}} = \mathbf{2}_{S\,\mathbf{Set}}$. Thus, for every element t of O, $t : \mathbf{1} \longrightarrow O$, $t \in O$ if and only if ${}_S\varphi \circ t = {}_Strue_{\mathbf{Set}}$, and $t \notin O$ if and only if ${}_S\varphi_m \circ t = {}_Sfalse_{\mathbf{Set}}$, since ${}_Sfalse_{\mathbf{Set}}$ is the only morphism distinct from ${}_Strue_{\mathbf{Set}}$. According to the aforementioned convention, I will use '${}_S\mathbf{Set}$' to denote that $\Omega_{\mathbf{Set}}$ is ${}_S\Omega_{\mathbf{Set}}$. A similar convention will be used for the categories below.

Example 2.2 ${}_S\mathbf{Set}^{\rightarrow}$ is the standard category of functions. A terminal object in this category, $\mathbf{1}_{S\,\mathbf{Set}^{\rightarrow}}$, is the identity function from $\mathbf{1}_{S\,\mathbf{Set}}$ to $\mathbf{1}_{S\,\mathbf{Set}}$.

Consider two objects of $\mathbf{Set}^{\rightarrow}$, $f : A \longrightarrow B$ and $g : C \longrightarrow D$. If f is a subobject of g, then $A \subseteq C$, $B \subseteq D$ and f is the restriction of g, that is, $f(x) = g(x)$ for $x \in A$. To the question 'Is a given element x of C also an element of B?' there are only two possible answers: Either it is or it is not, so the codomain of a function playing the role of a subobject classifier can be ${}_S\Omega_{S\,\mathbf{Set}}$. But before giving that definite answer, one must compute whether x is in A or not. One then has three options:

(i) Either $x \in A$, so the final answer to the original question is 'Yes', because $g(x) \in B$; or
(ii) $x \notin A$, but the final answer to the original question will be 'Yes', because $g(x) \in B$ after all; or
(iii) $x \notin A$, but the final answer will be 'No' because $x \notin B$ too.

Then, the domain of a function playing the role of a subobject classifier will be any three-element set to represent these three options. Let me use '1', '$\frac{1}{2}$' and '0' to denote each of those elements, respectively. So ${}_S\Omega_{\mathbf{Set}^{\rightarrow}}$ looks like this:

$$t : \mathbf{3}_{S\,\mathbf{Set}} \longrightarrow {}_S\,\Omega_{\mathbf{Set}}$$

with $t(0) = {}_Sfalse_{\mathbf{Set}}$ and $t(\frac{1}{2}) = t(1) = {}_S\,true_{\mathbf{Set}}$.

Thus, a subobject classifier in this category is ${}_Strue_{\mathbf{Set}^{\rightarrow}} : \mathbf{1}_{\mathbf{Set}^{\rightarrow}} \longrightarrow {}_S\Omega_{\mathbf{Set}^{\rightarrow}}$, i.e. a pair of morphisms $\langle t'_{\mathbf{Set}}, true_{S\,\mathbf{Set}}\rangle$ from $id_{\mathbf{1}_{\mathbf{Set}}} : \mathbf{1}_{\mathbf{Set}} \longrightarrow \mathbf{1}_{\mathbf{Set}}$ to ${}_S\Omega_{\mathbf{Set}^{\rightarrow}}$. There are only two truth values in this category. The calculation is straightforward and can be left to the reader (Hint: There seems to be an additional value; let us denote it as ${}_S\alpha_{\mathbf{Set}^{\rightarrow}} = \langle t'_{\mathbf{Set}}, true_{S\,\mathbf{Set}}\rangle$. Note that although $t' \neq t''$, ${}_S\alpha_{\mathbf{Set}^{\rightarrow}} = {}_S\,true_{\mathbf{Set}^{\rightarrow}}$).

Example 2.3 ${}_SS^{\downarrow\downarrow}$ is the category of (standard irreflexive directed multi-) *graphs* and graph structure preserving maps.[6] An object of ${}_SS^{\downarrow\downarrow}$ is any pair of sets equipped with a parallel pair of maps $A \overset{s}{\underset{t}{\rightrightarrows}} V$ where A is called the set *arrows* and V is the set of *dots* (or *nodes* or *vertices*). If a is an element of A (an arrow), then $s(a)$ is called the *source* of a, and $t(a)$ is called the *target* of a.

[6] Nice introductions to this category can be found in [38] and [22].

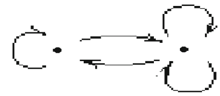


Fig. 1 Truth values object of ${}_S\mathbf{S}^{\downarrow\downarrow}$

Morphisms of ${}_S\mathbf{S}^{\downarrow\downarrow}$ are also defined so as to respect the graph structure. That is, a morphism $f : (A \underset{t}{\overset{s}{\rightrightarrows}} V) \longrightarrow (E \underset{t'}{\overset{s'}{\rightrightarrows}} P)$ in ${}_S\mathbf{S}^{\downarrow\downarrow}$ is defined to be any *pair* of morphisms of **Set** $f_a : A \longrightarrow V$, $f_v : E \longrightarrow P$ for which both equations

$$f_v \circ s = s' \circ f_a$$
$$f_v \circ t = t' \circ f_a$$

are valid in ${}_S\mathbf{Set}$. It is said that f preserves the structure of the graphs if it preserves the source and target relations.

A terminal object in this category, $\mathbf{1}_{{}_S S^{\downarrow\downarrow}}$, is any arrow such that its source and target coincide.

This topos provides a simple yet good example of a truth values object with more than two elements. ${}_S\Omega_{S^{\downarrow\downarrow}}$ has the form of a graph like that in Fig. 1. There are exactly three morphisms $\mathbf{1}_{{}_S S^{\downarrow\downarrow}} \longrightarrow {}_S\Omega_{S^{\downarrow\downarrow}}$ in this category, which means that ${}_S\Omega_{S^{\downarrow\downarrow}}$ has three truth values with the order ${}_S\mathit{false}_{S^{\downarrow\downarrow}} < {}_S\binom{s}{t}_{S^{\downarrow\downarrow}} < {}_S\mathit{true}_{S^{\downarrow\downarrow}}$.

2.3 The Standard Connectives

A morphism $k : ({}_S\Omega \times \cdots \times {}_S\Omega)^{{}_S\Omega^{\cdot^{\cdot^{\cdot^{{}_S\Omega^X}}}}} \longrightarrow {}_S\Omega$ (with ${}_S\Omega \times \cdots \times_S \Omega$ n times and ${}_S\Omega^{\cdot^{\cdot^{\cdot^{{}_S\Omega}}}}$ t times, $n, t \geq 0$), abbreviated as $k : {}_S\Omega^{nm} \longrightarrow {}_S\Omega$, will be said to be an *(standard) n-ary connective* of order m (where $m = t - 1$). Propositions, i.e. morphisms $\mathbf{1} \longrightarrow {}_S\Omega$ can thus be considered 0-ary connectives when $n = t = 0$.

Usual connectives are defined as certain equalizers which imply the following truth conditions:

Negation: $\neg p = {}_S\mathit{true}$ if and only if $p = {}_S\mathit{false}$; otherwise $\neg p = {}_S\mathit{false}$
Conjunction: $(p \wedge q) = \inf(p, q)$
Disjunction: $(p \vee q) = \sup(p, q)$
Conditional: $(p \Rightarrow q) = {}_S\mathit{true}$ if and only if $p \leq q$, otherwise $(p \Rightarrow q) = q$
Universal quantifier: $\forall_X \varphi(x) = \inf(\varphi(x))$
Particular quantifier: $\exists_X \varphi(x) = \sup(\varphi(x))$.

2.4 The Internal Logic of a Standard Topos

There is a theorem establishing necessary and sufficient conditions for a proposition ${}_S p$ being the same morphism as ${}_S true$ in a given standard topos ${}_S\mathcal{E}$. Let '$\models_I$' indicate that logical consequence gives the results as in intuitionistic logic. Then the following theorem holds.

Theorem 2.4 For every proposition ${}_S p$, $\models_{S\mathcal{E}}\, {}_S p$ for every topos ${}_S\mathcal{E}$ if and only if $\models_I {}_S p$.

i.e. ${}_S\Omega$ is a Heyting algebra.[7]

Summarizing, the standard categorial analysis of logic implies the following:

(IL1) Propositions form a partial order, i.e. for every propositions p, q and r:
(IL1a) $p \le p$
(IL1b) If $p \le q$ and $q \le p$, then $p = q$
(IL1c) If $p \le q$ and $q \le r$, then $p \le r$
(IL2) There is a truth value called ${}_S true$ with the following property:

$$\text{For every proposition } \ p, p \le_S true.$$

(IL3) One can define a truth value called ${}_S false$ that has the following property:

$${}_S false \le_S true$$

(IL4) From (IL2) and (IL3) one can obtain

$$\text{For every proposition } \ p, {}_S false \le p$$

(IL5) Connectives obey the following truth conditions:

$$\neg p = {}_S true \ \text{ if and only if } \ p = {}_S false, \ \text{ otherwise } \ \neg p = {}_S false$$
$$(p \wedge q) = \inf(p, q)$$
$$(p \vee q) = \sup(p, q)$$
$$(p \supset q) = {}_S true \ \text{ if and only if } \ p \le q, \ \text{ otherwise } \ (p \supset q) = q$$
$$\forall_X \varphi(x) = \inf(\varphi(x))$$
$$\exists_X \varphi(x) = \sup(\varphi(x)).$$

(IL6) The categorial analysis of logic does not imply, but rather assume, the traditional, 'Tarskian', notion of logical consequence.
Let '$p \models_{S\mathcal{E}} q$' denote that q is a logical consequence of p in a standard topos $\mathcal{E}$, i.e. that whenever p is the same morphism as ${}_S true$ in ${}_S\mathcal{E}$, so is q. Equivalently, if q is not the same morphism as ${}_S true$, p neither is. $\models_{S\mathcal{E}}\ p$ means that p is the same morphism as ${}_S true$ in ${}_S\mathcal{E}$.
(IL7) From (IL1)–(IL6), the internal logic of a standard topos is in general intuitionistic.

[7] I have made a little abuse of notation, for I used '${}_S p$' in both $\models_{S\mathcal{E}}$ and $\models_I$. In rigour, ${}_S p$ is a morphism which corresponds to a formula $({}_S p)^*$ in a possibly different language, but there is no harm if one identifies them. A proof can be found in [15, see §8.3 for the soundness part and §10.6 for the completeness part].

Example 2.5 The internal logic of ${}_S\mathbf{Set}$ is classical. For example, in ${}_S\mathbf{Set}$, every proposition p is the same as one and only one of ${}_S true_{\mathbf{Set}}$ and ${}_S false_{\mathbf{Set}}$. $\neg \circ {}_S true_{\mathbf{Set}} = {}_S false_{\mathbf{Set}}$ and $\neg \circ {}_S false_{\mathbf{Set}} = {}_S true_{\mathbf{Set}}$. Hence, for any p, $\neg\neg p = p$. Also, for any p $(p \vee \neg p) = \vee \circ \langle p, \neg p\rangle = \sup(p, \neg p) = {}_S true_{\mathbf{Set}}$.

Example 2.6 Even though it is many-valued, the internal logic of ${}_S\mathbf{Set}^2$ is classical: ${}_S\Omega\mathbf{Set}^2$ is a Boolean algebra with four elements, which in turn is the Cartesian product of a two-element Boolean algebra with universe $\{{}_S true_{\Omega_{\mathbf{Set}}}, false_{\Omega_{\mathbf{Set}}}\}$ with itself (i.e. operations act coordinatewise). For example, negation gives

$$\neg {}_S true_{\mathbf{Set}^2} = \langle \neg true_{S\,\mathbf{Set}}, \neg true_{S\,\mathbf{Set}}\rangle = \langle false_{S\,\mathbf{Set}}, false_{S\,\mathbf{Set}}\rangle = {}_S true_{\mathbf{Set}^2}$$
$$\neg {}_S \alpha_{\mathbf{Set}^2} = \langle \neg true_{S\,\mathbf{Set}}, \neg false_{S\,\mathbf{Set}}\rangle = \langle false_{S\,\mathbf{Set}}, true_{S\,\mathbf{Set}}\rangle = {}_S \beta_{\mathbf{Set}^2}\ .$$

The cases of ${}_S\alpha_{\mathbf{Set}^2}$ and ${}_S false_{\mathbf{Set}^2}$ are left to the reader. It is easy verify that for every p in ${}_S\mathbf{Set}^2$, $\neg\neg p = p$ and that $(p \vee \neg p) = {}_S true_{\mathbf{Set}^2}$.

Example 2.7 As I have mentioned, ${}_S\Omega_{S\downarrow\downarrow}$ has three truth values with the order ${}_S false_{S\downarrow\downarrow} < {}_S\binom{s}{t}_{S\downarrow\downarrow} < {}_S true_{S\downarrow\downarrow}$. Negation gives the following identities of morphisms:

$$\neg {}_S true_{S\downarrow\downarrow} = {}_S false_{S\downarrow\downarrow}\,, \quad \neg {}_S\tbinom{s}{t}_{S\downarrow\downarrow} = false_{S\downarrow\downarrow}\,, \quad \neg {}_S false_{S\downarrow\downarrow} = {}_S true_{S\downarrow\downarrow}\ .$$

Since $(p \supset q) = {}_S true$ if and only if $(p \wedge q) = p$, in general $(\neg\neg p \supset p) \neq {}_S true$ in $S^{\downarrow\downarrow}$ because even though $(\neg\neg p \supset p) = {}_S true_{S\downarrow\downarrow}$ either when $p = {}_S true_{S\downarrow\downarrow}$ or when $p = {}_S false_{S\downarrow\downarrow}$, $(\neg\neg p \wedge p) \neq \neg\neg p$ when $p = {}_S\binom{s}{t}_{S\downarrow\downarrow}$. Given that $(\neg\neg p \supset p) \neq {}_S true_{S\downarrow\downarrow}$ but there is no formula Φ such that $\Phi = true$ in classical logic and $\Phi = false$ in intuitionistic logic, $(\neg\neg p \supset p) = {}_S\binom{s}{t}_{S\downarrow\downarrow}$ when $p = {}_S\binom{s}{t}_{S\downarrow\downarrow}$. Moreover, $p \vee \neg p$ fails to be the same morphism as ${}_S true_{S\downarrow\downarrow}$ since $(p \vee q) = {}_S true$ if and only if either $p = {}_S true$ or $q = {}_S true$. If $p = {}_S\binom{s}{t}_{S\downarrow\downarrow}$, $\neg p = {}_S false_{S\downarrow\downarrow}$, so neither $p = {}_S true_{S\downarrow\downarrow}$ nor $\neg p = {}_S true_{S\downarrow\downarrow}$ and hence $(p \vee \neg p) \neq {}_S true_{S\downarrow\downarrow}$.

2.5 *Standard Topos Logic in a Nutshell*

The beautiful picture of logic in a topos described above can be summarized in the following slogans:[8]

(S1) $\Omega_{\mathcal{E}}$ is (or at least can be seen as) a truth-values object. (Common categorial wisdom; see, for example, [15, 21–23].)

(S2) The internal logic of a topos is in general many-valued. (Common categorial wisdom; but see [3–5, 15, 21, 22, 26].)

[8] I use the word ‘slogan’ here pretty much in the sense of van Inwagen: ‘a vague phrase of ordinary English whose use is by no means dictated by the mathematically formulated speculations it is supposed to summarize’ [36, p. 163], ‘but that looks as if it was’, I would add.

(S3) The internal logic of a topos is in general (with a few provisos) intuitionistic. (This also is common categorial wisdom, just to name but two important texts where this is asserted; see [15] and [23].)
(S4) Intuitionistic logic is the objective logic of variable sets. (A powerful metaphor widely accepted; see [18, 19, 27].)
(S5) The universal, invariant laws of mathematics are those of intuitionistic logic (cf. again [3–5].)

With the exception of (S5),[9] which is a claim specifically due to Bell, these slogans are theses so widely endorsed by topos theorists as accurate readings of some definitions, results and constructions in topos theory that it is hardly worth documenting, but I have done it just to show that they appear in several major texts written by leading category theorists. In the remaining of the chapter, I will show that these slogans are heavily loaded, philosophically speaking, and that they are not immediate consequences of the purely mathematical features of toposes, and that there is a more purely structural characterization of toposes. The notion of complement topos plays a crucial role in finding such a more structural characterization of toposes.

3 Complement-Toposes and the Non-Structural Components of Topos Logic

(...) philosophy continues to suffer from a certain 'prejudice towards truth'. (...) But why should truth be privileged over flasehood? Why should acceptance be privileged over rejection?
—João Marcos, *Ineffable inconsistencies*.

3.1 A Categorial Approach to Inconsistency: Bi-Heyting Toposes

A *bi-Heyting algebra* is a distributive lattice which is both a Heyting algebra and a Brouwerian algebra (the dual of a Heyting algebra, also called 'co-Heyting algebra'). Clearly, a Boolean algebra is a bi-Heyting algebra. Let $c()$ be the operation of Boolean complement. Define then $a \rightarrow b = c(a) \bigvee b$ and $a - b = a \bigwedge c(b)$. In this case $-a = \urcorner a = c(a)$.

A *bi-Heyting topos* is a standard topos for which the algebra of subobjects of any object is a Brouwerian algebra. Since the algebra of subobjects of any object in a standard topos is a Heyting algebra, a bi-Heyting topos can be defined as a standard topos for which the algebra of subobjects of any object is bi-Heyting.

There might be objects in a standard topos whose algebra of subobjects is a co-Heyting algebra. In [31], following the work in [20], some examples in the category ${}_SS^{\downarrow\downarrow}$ are given. This is the closest one will get paraconsistency in standard toposes, though. The

[9] And maybe also of (S4), due mostly to the appearance of Hegelian terminology ('objective'), very frequent in Lawvere but not in other topos-theorists. Omitting that, one can add [2] and [15] as supporters of this slogan.

internal logic of a bi-Heyting topos is never dual to an intuitionistic or superintuitionistic logic. Remember that the internal logic of a topos is determined by the algebra of ${}_S\Omega$ and the connectives, not by the algebra of its subobjects, and it is a co-Heyting algebra only if it is a Boolean algebra. This is assured by the following theorems.

Theorem 3.1 Let $\delta : {}_S\Omega \longrightarrow {}_S\Omega$ a morphism such that $\delta \leq id_{S\Omega}$ and $\delta \circ {}_S true = {}_S true$. Then $\delta = id_{S\Omega}$.

(This is corollary 1.12 in [30] or proposition 4.1 in [31], where a proof is given.)

Theorem 3.2 In any topos ${}_S\mathcal{E}$ the following conditions are equivalent:

(a) ${}_S\mathcal{E}$ is Boolean.
(b) $\neg \circ \neg = id_{S\Omega}$.

(This is proved as theorem 7.3.1 in [15].)

Theorem 3.3 If ${}_S\mathcal{E}$ is Boolean, then its internal logic is classical.

(This is proved as theorem 7.4.1 in [15].)

However, these results rely heavily on the *standard* character of a topos, i.e. on a particular description of its categorial structure. In what follows, I will show that the same categorial structure can be described in an alternative, coherent way, such that the internal logic of a topos can also be described as dual intuitionistic or paraconsistent.

3.2 Introducing Complement-Toposes

Mortensen's argument for developing an inconsistency-tolerant approach to category theory is that every topological space gives a topos (the category of pre-sheaves on the space), mathematically

> (...) specifying a topological space by its closed sets is as natural as specifying it by its open sets. So it would seem odd that topos theory should be associated with open sets rather than closed sets. Yet this is what would be the case if open set logic were the natural propositional logic of toposes. At any rate, there should be a simple 'topological' transformation of the theory of toposes, which stands to closed sets and their logic [i.e. inconsistency-tolerant], as topos theory does to open sets and intuitionism. [28, p. 102]

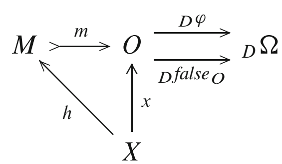
If the duality between intuitionistic logic and CSL is as deep as topological, then a representation of CSL as the internal logic of a topos should be equally natural. So Mortensen's remark amounts to this: The same categorial structure described as supporting intuitionistic logic should also be describable as supporting inconsistency-tolerance. Note that the crucial motivation is the topological motivation, and does not turn on paraconsistent ideology (even though Mortensen subscribes to the latter). In what follows, I expound Mortensen and Lavers's dualization of logical connectives in a topos.[10]

[10] It is important to set their individual contributions. Of the ten diagrams in [28, Ch. 11], Mortensen drew the first one and the final five, while Lavers drew the remaining four. The diagram for the dual-conditional

Think of the objects of complement toposes as the objects of standard toposes in section 2 and retain the definition of propositional functions. It will be assumed that there is a proposition $false : \mathbf{1} \longrightarrow \Omega$. This assumption will obligate certain names for other morphisms. '${}_D\Omega$' will denote this initial assumption about the name of a certain morphism with codomain Ω '*false*' in this case) and from here on, ${}_D f$ will denote that there is a monomorphism from ${}_D\Omega$ to the codomain of f and ${}_D\mathcal{E}$ that the morphisms with codomain of the object of propositions of $\mathcal{E}$ receive their names according to this initial assumption.

Then, for any object X in a complement topos, the composite ${}_D false \circ !_X : X \longrightarrow \mathbf{1} \longrightarrow {}_D\Omega$ denotes a constant, ${}_D false$-valued propositional function on X, abbreviated as ${}_D false_X$. Propositional functions will specify subobjects as follows. Given a propositional function ${}_D\varphi : X \longrightarrow {}_D\Omega$, one gets the part of the x's of which ${}_D\varphi$ is false, if any, as an equalizer $m : M \rightarrowtail X$ of ${}_D\varphi$ and ${}_D false_X$. This subobject will be named the *anti-extension* of the propositional function ${}_D\varphi$. A morphism ${}_D false : \mathbf{1} \longrightarrow {}_D\Omega$, called the *dual classifier*, has the following property:

Anti-comprehension axiom For each ${}_D\varphi : O \longrightarrow {}_D\Omega$ there is an equalizer of ${}_D\varphi$ and ${}_D false_O$, and each monic $m : M \rightarrowtail O$ is such an equalizer for a unique ${}_D\varphi$. In diagrams, ${}_D false$ is such that for every ${}_D\varphi$ and every object T and morphism $o : T \longrightarrow O$, if $m \circ_D \varphi = m \circ_D true_O$ and $x \circ_D \varphi = x \circ_D true_O$, then there is a unique $h : X \longrightarrow M$ that makes the diagram below commutative:



never was explicitly drawn, but it was discussed in [28, p. 109]. The full story, as told by Mortensen in personal communication is as follows. Mortensen gave a talk at the Australian National University (Canberra) in late 1986, on paraconsistent topos logic, arguing the topological motivation for closed set logic. He defined a complement topos, drew the first three diagrams from *Inconsistent Mathematics*, chapter 11, that is including the complement versions of ${}_S true$ and paraconsistent negation, and criticized Goodman's views on the conditional. But it was not seen clearly at that stage how the logic would turn out. Peter Lavers was present (also Richard Routley, Robert K. Meyer, Michael A. McRobbie, Chris Brink and others). For a couple of days in Canberra, Mortensen and Lavers tried without success to thrash it out. Mortensen returned home to Adelaide and two weeks later Lavers' letter arrived in Adelaide, in which he stressed that inverting the order is the key insight to understanding the problem, drew the diagrams for conjunction and disjunction, and pointed out that subtraction is the right topological dual for the conditional. Mortensen then responded with the four diagrams for the $S5$ conditional, and one for quantification (last five diagrams in *Inconsistent Mathematics*, chapter 11). A few months later (1987), Mortensen wrote the first paper, with Lavers as co-author, and sent it to Saunders Mac Lane and William Lawvere (also Routley, Meyer, Priest). Mac Lane replied but Lawvere did not. A later version of that paper became the eleventh chapter of *Inconsistent Mathematics*.

Let me call ${}_D\mathit{false} : \mathbf{1} \longrightarrow {}_D\Omega$ a *dual classifier*. ${}_D\varphi$ is then the *dual characteristic morphism* of m.[11] A *complement topos* ${}_D\mathcal{E}$ is then a Cartesian closed category with a dual classifier. Thus, what the *Anti-comprehension axiom* says is that ${}_D\varphi(x) = {}_D\mathit{false}_X$ if and only if $x \in m$. This implies the reconceptualization of certain components of toposes mentioned earlier: Given a propositional function ${}_D\varphi$, the part of the O's for which it is false is obtained as an equalizer of ${}_D\varphi$ and ${}_D\mathit{false}_O$ (a constant propositional in O with false as value). Such an equalizer is a subobject $m : M \rightarrowtail O$, the *anti-extension* of the propositional function ${}_D\varphi$. Now, ${}_D\Omega$ is in general a Brouwerian algebra, because as the standard connectives form a Heyting algebra, their duals form a Brouwerian algebra, which can be easily verified. This means that the internal logic of a complement topos is not intuitionistic, but a certain kind of dual intuitionistic, paraconsistent logic (examples of the dualized internal logic of concrete toposes can be found in [11]). The truth conditions implied by the definitions of the connectives would be

Negation: $\sim p = {}_D\mathit{false}$ if and only if $p = {}_D\mathit{true}$; otherwise $\sim p = {}_D\mathit{true}$
Disjunction: $(p \curlyvee q) = \sup(p, q)$
Conjunction: $(p \curlywedge q) = \inf(p, q)$
Subtraction: $(q - p) = {}_S\mathit{false}$ if and only if $q_D \leq p$; otherwise $(q - p) = q$
Particular quantifier: $E_X\varphi(x) = \sup(\varphi(x))$
Universal quantifier: $A_X\varphi(x) = \inf(\varphi(x))$

Given that ${}_D\mathcal{E}$ is a category with exponentials, one has ${}_D\Omega^X$, ${}_D\Omega^{{}_D\Omega^X}$, etc. for any X in ${}_D\mathcal{E}$, which may be regarded as representing collections of properties, properties of properties, etc. defined over X, so one can also have higher order dual propositions.

If ${}_S\mathcal{E}$ is a standard topos and ${}_D\mathcal{E}$ is the category obtained by assuming not the name ${}_S\mathit{true}$, but ${}_D\mathit{false}$ for a given morphism with codomain Ω and making the corresponding suitable choice of names for connectives, then ${}_D\mathcal{E}$ and ${}_S\mathcal{E}$ are categorially indistinguishable since terminal and initial objects, pullbacks, pushouts, and exponentials are notions and constructions prior to the characterization of classifiers and connectives. Moreover, Mortensen proved the following

Theorem 3.4 (Duality Theorem) Let S be a statement about ${}_D\mathcal{E}$ obtained by the above relabelling method from a statement S' about ${}_S\mathcal{E}$. Then S' is true of ${}_S\mathcal{E}$ if and only if S is true of ${}_D\mathcal{E}$.

A proof can be found in [28, p. 106].

Clearly, Heyting algebras and Brouwerian algebras, on one hand, and the logics they give rise to, on the other, are dual. Nonetheless, toposes ${}_S\mathcal{E}$ and ${}_D\mathcal{E}$ are not dual in the traditional categorial sense, so this other kind of duality has to be studied. A categorial characterization of the 'duality' between standard toposes and complement toposes would be most welcome, but for now I will describe in more detail the internal logic of complement toposes.[12]

[11] Mortensen and Lavers use the names *complement-classifier* and *complement topos*, which are now the names set in the literature (cf. [11, 28, 29, 37]). I have decided not to use the name 'dual topos' because the adjective 'dual' applied to categories has another well-entrenched meaning in category theory.
[12] I have attempted such a categorial description of this kind of duality in [10].

The *internal logic* of a complement topos ${}_D\mathcal{E}$ is the algebra induced by the object of propositions or algebraic truth values, ${}_D\Omega$, and the connectives ($\sim$, $\curlywedge$, etc.). Consequence is defined as usual: Let ${}_Dp \models_{{}_D\mathcal{E}} {}_Dq$ denote that whenever the morphism ${}_Dp$ is the same morphism as ${}_D true$ in ${}_D\mathcal{E}$, so is ${}_Dq$ ($\models_{{}_D\mathcal{E}} {}_Dp$ means that ${}_Dp$ is the same morphism as ${}_D true$ in ${}_D\mathcal{E}$).

There is a theorem establishing necessary and sufficient conditions for a proposition ${}_Dp$ being the same morphism as ${}_D true$ in a given complement topos ${}_D\mathcal{E}$. Let $\models_{CSL}$ be the extension of the consequence relation of closed set logic. Then the following theorem holds.

Theorem 3.5 For every topos ${}_D\mathcal{E}$ and proposition ${}_Dp$, $\models_{{}_D\mathcal{E}} {}_Dp$ if and only if $\models_{CSL} {}_Dp$.

i.e. ${}_D\Omega$ is a Brouwerian algebra (by Theorem 2.4 and the Duality Theorem 3.4 above).[13]

Summarizing, the complement-categorial analysis of logic implies the following:[14]

(DIL1) Propositions form a partial order, i.e. for every propositions p, q and r:
(DIL1a) $p \geq p$
(DIL1b) If $p \geq q$ and $q \geq p$, then $p = q$
(DIL1b) If $p \geq q$ and $q \geq r$, then $p \geq r$
(DIL2) There is a truth value called ${}_D false$ with the following property:

$$\text{For every proposition } p, p \geq_D false$$

(DIL3) One can define a truth value called ${}_D true$ that has the following property:

$${}_D true \geq_D false.$$

(DIL4) From (DIL2) and (DIL3) one can obtain

$$\text{For every proposition } p, {}_D true \geq p.$$

(DIL5) Connectives obey the following truth conditions:

$$\begin{aligned}
&\sim p = {}_D false \text{ if and only if } p = {}_D true, \text{ otherwise } \sim p = {}_D true\\
&(p \curlyvee q) = \sup(p, q)\\
&(p \curlywedge q) = \inf(p, q)\\
&(q - p) = {}_S false \text{ if and only if } q_D \leq p, \text{ otherwise } (q - p) = q\\
&E_X\varphi(x) = \sup(\varphi(x))\\
&A_X\varphi(x) = \inf(\varphi(x)).
\end{aligned}$$

---

[13] Again, I have made a little abuse of notation, for I used '${}_Dp$' in both $\models_{{}_D\mathcal{E}}$ and $\models_I$. In rigour, ${}_Dp$ is a morphism which corresponds to a formula $({}_Dp)^*$ in a possibly different language.

[14] By abuse of notation but to simplify reading, I will not indicate that the order here is dual to that in standard toposes, unless there is risk of confusion.

(DIL6) The categorial analysis of logic in complement toposes assumes the Tarskian notion of logical consequence too.
Let '$p \models_{D\mathcal{E}} q$' denote that q is a logical consequence of p in a complement topos ${}_D\mathcal{E}$, i.e. that whenever p is the same morphism as ${}_D true$ in ${}_D\mathcal{E}$, so is q. Equivalently, if q is not the same morphism as ${}_D true$, p neither is. $\models_{S\mathcal{E}} p$ means that p is the same morphism as ${}_D true$ in ${}_D\mathcal{E}$.

(DIL7) From (DIL1)–(DIL6), in the internal logic of a complement topos hold at least the laws of dual intuitionistic logic, as shown in Theorem 3.5.[15]

Example 3.6 Since classical logic is its own dual (just as a Boolean algebra is its own dual), the internal logic of e.g. **Set** is not modified by renaming and thus complement-**Set** (${}_D$**Set**) is the same as **Set**.[16]

Example 3.7 Complement-$S^{\downarrow\downarrow}$ or ${}_D S^{\downarrow\downarrow}$ has, mutatis mutandis, the same three truth values with its original order,[17] but negation gives now the following identities of morphisms:

$$\sim_D false_{S^{\downarrow\downarrow}} =_D true_{S^{\downarrow\downarrow}}, \ \sim_D \tbinom{s}{t}_{S^{\downarrow\downarrow}} =_D true_{S^{\downarrow\downarrow}}, \ \sim_D true_{S^{\downarrow\downarrow}} =_D false_{S^{\downarrow\downarrow}} .$$

In $S^{\downarrow\downarrow}$ one has $(p \vee \neg p) \neq {}_S true_{S^{\downarrow\downarrow}}$, and in the alternative labelling one obtains $(p \curlywedge \sim p) \neq {}_D false_{S^{\downarrow\downarrow}}$. Remember that in a complement topos $(p \curlywedge q) = {}_D false$ if and only if either $p = {}_D false$ or $q = {}_D false$. If $p = {}_D \tbinom{s}{t}_{S^{\downarrow\downarrow}}$ then $\sim p = {}_D true_{S^{\downarrow\downarrow}}$, so neither $\sim p = {}_D false_{S^{\downarrow\downarrow}}$ nor $p \neq {}_D false_{S^{\downarrow\downarrow}}$ and hence $(p \curlywedge \sim p) \neq {}_D false_{S^{\downarrow\downarrow}}$. Besides, in a Heyting algebra (like the algebra ${}_S\Omega$), in general, it is not the case that $q \leq (p \vee \neg p)$, which in the alternative labelling corresponds to the fact that in a Brouwerian algebra (like ${}_D\Omega$), in general, it is not the case that $(p \curlywedge \sim p) \leq q$. So, the internal logic of complement-$S^{\downarrow\downarrow}$ is not classical (nor intuitionistic!), but inconsistency-tolerant. Moreover, in complement-$S^{\downarrow\downarrow}$ both $p \curlyvee \sim p$ and $\sim (p \curlywedge \sim p)$ are the same morphism as ${}_D true_{S^{\downarrow\downarrow}}$, unlike their standard counterparts. In standard $S^{\downarrow\downarrow}$, $(p \wedge \neg p) = {}_S false_{S^{\downarrow\downarrow}}$, which in the alternative labelling gives $(p \curlyvee \sim p) = {}_D true_{S^{\downarrow\downarrow}}$. In standard $S^{\downarrow\downarrow}$, $\neg(p \vee \neg p) = {}_S false_{S^{\downarrow\downarrow}}$ (for in intuitionistic logic the negation of a classical theorem is always false), and the alternative labelling gives $\sim (p \curlywedge \sim p) = {}_D true_{S^{\downarrow\downarrow}}$.

[15] Inconsistency-tolerant categorial structures are studied further in [28, chapter 12, written by William James] and in [16].

[16] Thus, as Vasyukov ([37] p. 292) points out: '(...) in **Set** we always have paraconsistency because of the presence of both types of subobject classifiers (...)' just as we always have in it (at least) intuitionistic logic. The presence of paraconsistency within classical logic is not news. See, for example [8], where some paraconsistent negations in S5 and classical first-order logic are defined.

[17] It is easy to verify that after making all the necessary changes, i.e. changing ${}_S true_{S^{\downarrow\downarrow}}$ for ${}_D false_{S^{\downarrow\downarrow}}$, etc. the names are ordered in the same way as they are in ${}_S S^{\downarrow\downarrow}$.

3.3 Consequences for the Standard Story of Toposes

Being ${}_D\Omega$ and ${}_S\Omega$ isomorphic, standard toposes and complement toposes are both, well, toposes, because they do not differ in categorial structure. A further consequence of this isomorphism is that complement toposes are not parasitic on standard toposes: One could start with complement toposes and then obtain standard toposes by renaming. This means that, even if the categorial structure invites to be named in certain ways, it does not force it. All this helps to solve the following perplexity: If ${}_S\mathcal{E}$ and ${}_D\mathcal{E}$ should be indistinguishable because they are categorially indistinguishable, how can one in fact distinguish between them? According to the legend, Mac Lane, in response to Mortensen and Laver's paper mentioned in footnote 10, said that complement toposes are just standard toposes, that they are indistinguishable because they have the same categorial structure. However, they seem distinguishable; after all, they seem to have different internal logics, intuitionistic the ones and paraconsistent the others. The appropriate answer, I think, is this: To date, there is more than categorial structure in the study of toposes, to wit, special, intuitive names for some of the morphisms, invited, but not necessitated, by the categorial structure. It is worth emphasizing that complement toposes do not claim to be categorially different from standard toposes nor to say that dual connectives acquire further categorial properties *qua* morphisms after the renaming, but rather stress the fact that the same categorial stuff can be described in at least two different ways. Neither of the names is imposed by the categorial structure of toposes itself so, in its current form, there is more than just categorial structure in the study of toposes. As I have discussed elsewhere, given the isomorphism between ${}_D\Omega$ and ${}_S\Omega$, one cannot argue against complement toposes, as has been done for example in [9], using theorems $T_1, \ldots, T_n$ which involve ${}_S\Omega$ and connectives ${}_Sk$, when the right theorems for complement toposes are ${}_DT_1, \ldots, {}_DT_n$, which involve ${}_D\Omega$ and ${}_Dk$.

Thus, the notion of complement topos goes against slogans (S3)–(S5), but I think it is possible to advance further. The main morals of Mortensen and Lavers' study of complement toposes seem to be the following:

Moral 1. There is a 'bare' or 'abstract' categorial structure of toposes that can filled in at least two ways (the standard way and the way suggested by Mortensen and Lavers). Said otherwise, there are underlying universal properties in topos logic dissembled by certain intuitive conceptualizations of the categorial structure of toposes, yet not necessitated by this.

Moral 2. The theorem stating the intuitionistic character of the internal logic should be read rather as follows: Under certain conditions $c_1, \ldots, c_n$, most of them extra-categorial, (S3) is the case.

Moral 3. The universal, invariant laws of mathematics are not those of intuitionistic logic. They seem to be so only when $c_1, \ldots, c_n$ are adopted.

In the next section, I explore in more details these implications of complement toposes.

4 The Substance Behind the Categorial Orthodoxy

Are there any names which witness of themselves that they are not given arbitrarily,
but have a natural fitness?
—Plato, *Cratylus*.

4.1 Two Skolemizations

Note that the assumption in standard topos theory that there is a proposition $true : \mathbf{1} \longrightarrow \Omega$ is actually a twofold assumption. On the one hand, there is a categorial, 'formal' or 'structural' assumption, merely concerning the existence of a certain morphism

Existence of a truth value. There is a morphism $\nu : \mathbf{1} \longrightarrow \Omega$ On the other hand, there is a 'material' or more substantive concerning a very loaded conceptualization of such a morphism.

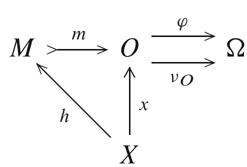
Name of a truth value. It is better thought of as '*true* : $\mathbf{1} \longrightarrow \Omega$' provided a plausible conceptualization of the properties it satisfies.

Complement-toposes share the first assumption, but give a different conceptualization of certain parts of the categorial structure of toposes in a way that it satisfies rather

Name of a truth value.* The morphism $\nu : \mathbf{1} \longrightarrow \Omega$ is better thought of as '*false* : $\mathbf{1} \longrightarrow \Omega$' provided a plausible conceptualization of the properties it satisfies. Thus, the particular Skolemization behind standard topos logic is not a purely categorial component of topos logic, and hence there should be a more abstract form of topos logic, independent from both standard and complement toposes.

A morphism $\nu : \mathbf{1} \longrightarrow \Omega$, called a *(bare) subobject classifier*, has the following property:

(Bare) Comprehension axiom For each $\varphi : O \longrightarrow \Omega$, there is an equalizer of φ and ν_O, and each monic $m : M \rightarrowtail O$ is such an equalizer for a unique φ. In diagrams, ν is such that for every φ and every object T and morphism $o : T \longrightarrow O$, if $m \circ \varphi = m \circ \nu_O$ and $x \circ \varphi = x \circ \nu_O$, then there is a unique $h : X \longrightarrow M$ that makes the diagram below commutative:



Note that, according to the definition of equalizer, h must be the only morphism that, among other things, $x = m \circ h$. But this satisfies the definition of $x \in m$. Thus, what the **(Bare) Comprehension axiom** says is that $\varphi(x) = \nu_X$ if and only if $x \in m$. This clearly invites the reading of ν as *true*, if m is thought of as the *extension* of the property φ, as is natural to think, but it also encompasses the reading in which complement toposes are based. In formal terms, both the standard and complement conceptualization are particular Skolemizations of the (equational) formula describing the (bare) subobject classifier.

The condition *Name of the truth value*, the 'standard Skolemization' of ν, obligates certain names for other categorial ingredients, which I have denoted using the subscript S, but let me state more formally how it works. Let '$\ulcorner\urcorner$' denote an instantiation device, such that '$\ulcorner x \urcorner$' denotes a constant which is the replacement of x and thus '${}_S\ulcorner x \urcorner$' denotes the standard instantiation of x. Thus,

$$({}_S true)_S \ulcorner \nu : \mathbf{1} \longrightarrow \Omega \urcorner = {}_S true : \mathbf{1} \longrightarrow {}_S \Omega \, .$$

According to this, '${}_S\Omega$' denotes that

(${}_S\Omega$) for every $f : X \longrightarrow \Omega$ in a given topos $\mathcal{E}$, $\ulcorner f : X \longrightarrow \Omega \urcorner$ is standard according to the initial Skolemization for $\nu : \mathbf{1} \longrightarrow \Omega$.
'${}_S\mathcal{E}$' denotes something similar to '${}_S\Omega$', but emphasizing the ambient topos '$\mathcal{E}$':
(${}_S\mathcal{E}$) in a given topos $\mathcal{E}$, for every $f : X \longrightarrow \Omega$, $\ulcorner f : X \longrightarrow \Omega \urcorner$ is standard according to the initial Skolemization for $\nu : \mathbf{1} \longrightarrow \Omega$ and
'${}_S f$' denotes quite the same as the two symbols above but emphasizing the morphism f:
(${}_S f$) for the morphism $f : X \longrightarrow \Omega$ in a given topos $\mathcal{E}$, $\ulcorner f : X \longrightarrow \Omega \urcorner$ is standard according to the initial Skolemization for $\nu : \mathbf{1} \longrightarrow \Omega$.

Similar conventions rule the use of the subscript D for complement toposes.

I must confess I do not know how to rinse the phrase '$\ulcorner f : X \longrightarrow \Omega \urcorner$ is standard according to the initial Skolemization for $\nu : \mathbf{1} \longrightarrow \Omega$' otherwise than by saying that the $\ulcorner f : X \longrightarrow \Omega \urcorner$'s correspond with some prior knowledge or conception of logical notions which is coherent with the initial choice of name for $\nu : \mathbf{1} \longrightarrow \Omega$. Consider the (partial) truth condition $p \# q = \nu$ if and only if $p = \nu$ and $q = \nu$: If one has chosen the name '*true*' for ν then the best name for # is 'conjunction', not 'disjunction' or some other.

Even if from a mathematical point of view all this might be regarded as uninteresting, preferring one reading above the other may have (and has had) important philosophical consequences. This is not a mere play with labels and, even though the underlying dualities between Heyting algebras and Brouwerian algebras are well known, the choice of names affects what we are considering as the internal logic of a topos. The choice of labels has profound philosophical implications. For example, the well-known Theorem 2.4 in topos theory is thought of as claiming what slogan 3 says:

(S3) The internal logic of a topos is intuitionistic (with a few provisos). (As I have said, this is common categorial wisdom.)
Given that most of ordinary mathematics can be reconstructed within a topos just as in traditional axiomatic set theories, the aforementioned result is sometimes read as stating slogan 5:

(S5) The universal, invariant laws of mathematics are those of intuitionistic logic (cf. [3–5])

There is also another philosophical claim connecting the internal logic of standard toposes and ‘the objective form of variation’:

(S4) The objective logic of variable sets is intuitionistic (cf. [18, 19, 27])

Nonetheless, complement toposes show that Theorem 2.4 should not be paraphrased laxly as (S3). Complement-toposes suggest that Theorem 2.4 should be read rather as follows: Under certain conditions $c_1, \ldots, c_n$ (and one of the c_i’s are the names chosen for morphisms with codomain Ω), including the proof of the theorem, (S3) is the case. The universal, invariant laws of mathematics do not seem to be those of intuitionistic logic, as stated in (S5). They seem to be so only when very specific $c_1, \ldots, c_n$ are adopted. Nor variation as embodied in standard toposes exhausts all kinds of variation, as claimed in (S4), for sets can also vary *backwards* (as opposed to *forwards* variation as in usual Kripke models). All this deserves more careful discussion; I just wanted to point out that choosing labels and names is not conceptually trivial and it does not lack mathematical interest.

4.2 The Tarskian Assumption and Logical Many-Valuedness

Slogans (S1) and (S2) are also in a similar, difficult position. Remember that the logical consequence in a topos is assumed to be traditional, Tarskian consequence:

q is a *(Tarskian) logical consequence* of premises Γ, in symbols $\Gamma \models^T q$, if *true* is preserved from premises to the conclusion and is not a consequence if the premises are the same morphism as *true* but the conclusion is not. A *thm* is a consequence of an empty set of premises, i.e. if it is a morphism which is the same morphism as *true*. A non-theorem is a morphism which is different from *true*. But the *two* values *true* and *not true* (or *untrue*, etc.) are the only values required to define the (Tarskian) consequence.

The internal logic of complement toposes is Tarskian, too. Even though the subobject classifier and the connectives are described in a different, dual way, the notion of consequence in the internal logic of complement toposes is the same as that of (ordinary or standard) toposes. Therefore, the subsequent discussion for the rest of this section can be cashed in terms of toposes *simpliciter*, ignoring whether they are standard or not unless otherwise indicated.

I said that slogans (S1) and (S2) are in a difficult position because there is a theorem by Roman Suszko which states that every Tarskian logic, i.e. every logic whose consequence relation is reflexive, transitive and monotonic, has a bivalent semantics. A philosophical intuition behind Suszko’s result is the distinction between *algebraic truth values* and *logical truth values*. Logical values are those values used to define a valid semantic consequence: If every premise is *true*, then so is (at least one of) the conclusion(s). In a contrapositive form, the other logical value can also be used to explain the valid semantic consequence: If the (every) conclusion is *not true*, then so is at least one of the premises. Thus only the two logical truth values *true* and *not true* or, more generally, *designated* and *antidesignated,* are needed in the definition of consequence.

Reductive results similar in spirit to Suszko's were presented independently by other logicians, for example, Newton da Costa (see, e.g., [17]), Dana Scott (cf. [33, 34]) and Richard Routley and Robert K. Meyer [32]. Moreover, there is a family of akin results of different strengths under the label 'Suszko's reduction'. *Suszko's reduction*, strictly speaking, required from the logic not only to be reflexive, transitive and monotonic, but also to be *structural*. *Suszko-da Costa's reduction* dropped the structurality requisite. *Suszko–Béziau's reduction* only requires reflexivity from the logic (cf. [35]).

Suszko declared that many-valuedness is 'a magnificent conceptual deceit' and claimed that '(...) there are but two logical values, true and false (...)'. This claim is now called *Suszko's thesis* and can be stated more dramatically as 'All logics are bivalent' or 'Many-valued logics do not exist at all'. Reductive results, especially the strongest form (Suszko-Béziau's reduction), seem to be overwhelming evidence in favour'of Suszko's thesis because virtually all logics regarded as such are in the scope of these theorems.

A possible way to resist Suszko's thesis is by extending the scope of logics to cover non-Tarskian logics, especially to non-reflexive ones to avoid Suszko–Béziau's reduction, and this reply is what I will discuss more extensively after discussing reductions in a categorial setting.

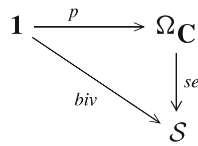
These reductive results can be given categorial content. The internal logic of a topos is said to be *algebraically n-valued* if there are n distinct morphisms $\mathbf{1} \longrightarrow \Omega$ in the given topos. As reductive results have shown, an algebraically n-valued Tarskian logic, in general, is not also *logically* n-valued. Accordingly, the internal logic of a topos is said to be *logically m-valued* if its notion of consequence implies that there are m distinct logical values. The internal logic of a topos, whether standard or complement, is a Tarskian logic, and this means that it is in the scope of Suszko's theorem. Such internal logic is defined, as usual, by distinguishing between those propositions that are the same morphism as *true* and the other ones, no matter what or how many algebraic truth values there are, i.e. what is playing the leading logical role is the bipartition *true* and *not true*, independently of the number of algebraic values (elements of Ω). So, provided that the notion of logical consequence is that usually assumed, and it is to deliver intuitionistic logic in the standard case and a dual of this in the complement case, the internal logic of a topos is logically bivalent.

The logical m-valuedness of the internal logic of a topos can be itself *internalized* just in case it can be replicated appropriately in terms of morphisms and compositions of the topos itself. I study here first the case of internalizing m-valuedness when $m = 2$ and suggest a more general definition in the next section.

Definition 4.1 In a non-degenerate category **C** with (respectively, dual) subobject classifier, a *Suszkian logical truth values object*, or *Suszkian object* for short, is an object $\mathcal{S}$ such that there are exactly two morphisms $\mathbf{1} \underset{\delta^-}{\overset{\delta^+}{\rightrightarrows}} \mathcal{S}$ and a morphism sep : Ω : $\mathbf{C} \longrightarrow \mathcal{S}$ such that sep is the unique morphism which satisfies the following properties:

(Sus1) $\text{sep} \circ p = \delta^+$ if $p = true_{\mathbf{C}}$, and
(Sus2) $\text{sep} \circ p = \delta^-$ if $p \neq true_{\mathbf{C}}$

The morphisms δ^+ and δ^- can be collectively denoted by biv and are called a *Suszkian bivaluation*. Thus, the diagram below commutes according to the above definition of biv and the conditions (S1) and (S2):



From the very definition of a Suszkian object, for every proposition φ, either sep $\circ\, \varphi = \delta^+$ or sep $\circ\, \varphi = \delta^-$. Now, for every theorem Φ, sep $\circ\, \Phi = \delta^+$, and for every non-theorem Ψ, sep $\circ\, \Psi = \delta^-$. Consider the three truth values in ${}_S S^{\downarrow\downarrow}$. Then sep $\circ_S$ $true_{S\downarrow\downarrow} = \delta^+$ and sep $\circ_S$ $false_{S\downarrow\downarrow}$ = sep $\circ_S$ $\binom{s}{t}_{S\downarrow\downarrow} = \delta^-$. Hence, for example, sep $\circ\, (\neg\neg p \Rightarrow p) = \delta^-$, for $(\neg\neg p \Rightarrow p) =_S \binom{s}{t}_{S\downarrow\downarrow}$ when $p =_S \binom{s}{t}_{S\downarrow\downarrow}$. Something similar happens with $(p \vee \neg p)$.

As a consequence of the definition, there is no morphism $\psi : \mathbf{1} \longrightarrow \mathcal{S}$ such that $\psi \in \delta^+$ and $\psi \in \delta^-$. However, this does not mean that $\cap \circ (\delta^+, \delta^-) = \varnothing$. This implies that, in general, $\mathcal{S}$ is not isomorphic to $\mathbf{2}$ in spite of having exactly two morphisms $\mathbf{1} \longrightarrow \mathcal{S}$ (nonetheless, it is easily proved that a Suszkian object and Ω are isomorphic for example in **Set**).

Unlike a subobject classifier, a Suszkian object does not necessarily classify subobjects and it does not necessarily count them, either, for it collapses every other proposition different from ${}_S true$ into δ^-. A Suszkian object provides a bivaluation biv = sep $\circ\, p$ for Ω, i.e. it says whether a proposition is logically true or not, full stop. This justifies the suggested definition of a Suszkian object, but the difficult part is proving the claim that every (non-degenerate) topos has a Suszkian object as defined here. Maybe not all toposes have a Suszkian object as defined here but only those with certain discreteness conditions. However, I only wanted to show that it seems possible to internalize the notions involved in the reductive results.

There are notions of logical consequence which are not Tarskian and that could introduce interesting complications in the theory of the internal logic of toposes.[18] Consider first Frankowski's P-consequence ('P' for 'Plausible'; cf. [13, 14]):

P-consequence. q is a logical P-consequence from premises Γ, in symbols $\Gamma \models^P q$, if and only if any case in which each premise in Γ is designated is also a case in which φ is not antidesignated. Or equivalently, there is no case in which each premise in Γ is designated, but in which q fails to be not antidesignated.

Thus logical many-valuedness in a topos could be obtained at a different level, by taking it into account from the very characterization of logical consequence. However, this would result in a change in the description of the internal logic, for it would be no

[18] Someone could argue that these are not notions of logical consequence at all, since logical consequence has to satisfy the Tarskian conditions. I guess (and hope) that readers of this *Festschrift* do not have this kind of doubts. For a defense of the logicality of non-Tarskian relations of logical consequence, see for example [12].

longer intuitionistic. The Tarskian properties are indissolubly tied to the canonical characterizations of consequence, but P-consequence is non-Tarskian: it is not transitive. Let me exemplify how radical the change would be in the internal logic if P-consequence is adopted instead of the Tarskian one.

In general, P-consequence does affect the collection of theorems. Since theorems are those propositions which are consequences of an empty set of premises, theorems according to P-consequence are those propositions that are not antidesignated. So theorems of the internal logic are not the same as those when Tarskian consequence is assumed even if $true_{\mathcal{I}}$ is taken as the only designated value. For example, let us assume as above that ${}_S true_{S^{\downarrow\downarrow}}$ is the only designated value in ${}_S\mathbf{S}^{\downarrow\downarrow}$ and that ${}_S false_{S^{\downarrow\downarrow}}$ is the only antidesignated value. $p \vee \neg p$ would be a theorem in ${}_S\mathbf{S}^{\downarrow\downarrow}$ because there is no case in which it is antidesignated.

P-consequence affects also the validity of inferences. Remember that unlike Tarskian consequence, P-consequence is not transitive. Suppose that $p = {}_S true_{S^{\downarrow\downarrow}}$, $q = {}_S\binom{s}{t}_{S^{\downarrow\downarrow}}$ and $r = {}_S false_{S^{\downarrow\downarrow}}$. Thus $p \models^P_{{}_S S^{\downarrow\downarrow}} q$ and $q \models^P_{{}_S S^{\downarrow\downarrow}} r$, but $p \not\models^P_{{}_S S^{\downarrow\downarrow}} r$, because P-consequence requires that if premises are designated, conclusions must be not antidesignated, which is not the case in this example.

However, in being reflexive P-consequence is in the scope of Béziau's reduction to bivalence, it does not assure logical many-valuedness. This is not the case with Malinowski's Q-consequence 'Q' for 'Quasi'; cf. [24, 25]), though:

Q-consequence. q is a logical Q-consequence from premises Γ, in symbols $\Gamma \models^Q q$, if and only if any case in which each premise in Γ is not antidesignated is also a case in which q is designated. Or equivalently, there is no case in which each premise in Γ is not antidesignated, but in which q fails to be designated.

The changes in the internal logic would be as follows. Theorems are those propositions which are consequences of an empty set of premises, so theorems are propositions that are always designated. This is just the usual notion of theoremhood, but whether Q-consequence affects the collection of theorems depends on what are the designated values, because one has to choose by hand, as it were, what are the designated, antidesignated and neither designated nor antidesignated values. If ${}_{\mathcal{I}}true$ is the only designated value as usual, the theorems of the internal logic are the same whether the Tarskian or Q-consequence is assumed.

Nonetheless, Q-consequence does affect the validity of inferences even if ${}_{\mathcal{I}}true$ is the only designated value. Unlike the Tarskian consequence, Q-consequence is not reflexive. For example, let us assume that ${}_S true_{S^{\downarrow\downarrow}}$ is the only designated value in ${}_S\mathbf{S}^{\downarrow\downarrow}$ and that ${}_S false_{S^{\downarrow\downarrow}}$ is the only antidesignated value. Suppose that $p = {}_S\binom{s}{t}_{S^{\downarrow\downarrow}}$. Then $p \not\models^Q_{{}_S S^{\downarrow\downarrow}} p$, because Q-consequence requires that if premises are not antidesignated, conclusions must be designated, which is not the case in this example.

The above are not the only possible changes. Consider the case when designated and antidesignated values form mutually exclusive and collectively exhaustive values; for simplicity take the Tarskian logical consequence $\models^T$, which states that if premises are designated, then the conclusions are also designated; equivalently, under the preceding assumption on values, if conclusions are antidesignated, premises are also antidesignated. Elaborating on [39], let me isolate and separate these forms of logical consequence:

D^+-consequence. q is a logical D^+-consequence from premises Γ, in symbols $\Gamma \models^{D^+} q$, if and only if any case in which each premise in Γ is designated is also a case in which q is designated. This is called *forwards preservation (of D^+).*

D^--consequence. q is a logical D^--consequence from premises Γ, in symbols $\Gamma \models^{D^-} q$, if and only if any case in which q is antidesignated is also a case in which some premise in Γ is antidesignated. This is called *forwards preservation (of D^-).*

When the arrangement of values is such that $D^+ \cup D^- \neq A$ (with A the collection of all algebraic values considered) or $D^+ \cap D^- \neq \varnothing$, D^+-consequence and D^--consequence do not coincide. Let us consider the category ${}_S\mathbf{S}^{\downarrow\downarrow}$ and let us assume that ${}_S true_{S^{\downarrow\downarrow}}$ is the only designated value and that ${}_S false_{S^{\downarrow\downarrow}}$ is the only antidesignated value. Then one has $p \wedge (p \Rightarrow q) \models^{D^+} q$ but $p \wedge (p \Rightarrow q) \not\models^{D^-} q$.

Again, *mutatis mutandis*, examples similar to the above can be given to show how each notion of consequence modifies the internal logic of a complement topos. [39] is a good source of inspiration for other notions of logical consequence. Abstraction on the notions of logical consequence could go further up to a definition of a logical structure analogous to that of an algebraic structure given in Universal Algebra such that other notions of consequence and particular logics appear as specifications of that structure: That is Béziau's project of Universal Logic; see [6] for an introduction. However, I will stop generalization here because of limitations of space and because it has been enough to show that the issue of the many-valuedness of topos logic is not as neat as thought in the categorial orthodoxy.

A problem at this point is to know whether the non-Tarskian notions of consequence can be internalized in a topos, but I will be back to that in the next section.

5 Bare Topos Logic

The undetermined is the structure of everything.
—Anaximander (in Jean-Yves Béziau's paraphrase)

5.1 Bare Toposes, Bare Order and Bare Connectives

One can forget for a moment all what one knows about toposes and tell the story from the beginning, in a way much similar to Awodey's [1] but with no particular, intuitive name for some of the main characters. I have started such description in the previous section, in describing the *(Bare) Comprehension axiom*. In this section, I will continue to make explicit the 'bare' or 'unlabelled' categorial structure of toposes, as well as equally unlabelled definitions of some logical connectives, and show how standard and complement toposes and their connectives are instances of those bare definitions. This allows an abstract definition of the internal logic of a topos from which it is evident that one gets intuitionistic logic if certain conditions are assumed, dual-intuitionistic or paraconsistency if another conditions are met, and what not if variation is taken to its last consequences. The results of this chapter summarize and end the justification of my criticisms on slogans (S1)–(S5). This is not so odd as it might seem at first sight. For example, in group theory there are different *notations* for the binary operation, which is in itself neither additive nor multiplicative. There is no need to call it, for example, 'addiplicative' since it can be referred to as *the* binary operation of the group and then specify different notations for it. This maneuver is not available here due to the presence of many interacting n-ary operations. So the bared structure of topos has to be thought as the abstract group and the specially named presentations as notations for that group.

It is well-known that the subobjects of a given object form a partial order. In particular, the elements of Ω form a partial order, which means that propositions form a partial order, i.e. for every propositions p, q and r:

- $p\mathfrak{D}p$
- If $p\mathfrak{D}q$ and $q\mathfrak{D}r$, then $p\mathfrak{D}r$
- If $p\mathfrak{D}q$ and $q\mathfrak{D}p$, then $p = q$

Note that the *direction* of $\mathfrak{D}$ depends on the name given to ν. The definition of ν implies that, for every proposition p, $p\mathfrak{D}\nu$. If ν is read as *true*, as in standard toposes, $\mathfrak{D}$ is better interpreted as $\leq$, but if ν is read as *false*, as in complement toposes, $\mathfrak{D}$ is better interpreted as $\geq$.

If $\mathfrak{D}$ is interpreted as a deducibility relation, $\vdash$, the properties above say that deducibility is reflexive, transitive and that interdeducible propositions are equivalent, no matter what the direction of the relation is.

Given a subset S of a partial order P, $insup$ denotes an element such that it is the infimum of S and, if the order in P is reversed, it turns out to be the supremum of S, or vice versa. $supin$ is the defined as the dual of $insup$, i.e. an element such that it is the supremum of S and, if the order in P is reversed, it turns out to be the infimum of S, or vice versa.

We can define another proposition, μ, as the bare classifying morphism of 0_1 (the only morphism from an initial object to a terminal one):

$$\mathbf{0} \xrightarrow{0_1} \mathbf{1} \underset{{}^{\mu}\mathbf{1}}{\overset{\mu =_{def.} \varphi_{0_1}}{\rightrightarrows}} {}_S\Omega$$

Let us call ν and μ 'special bare values'.

A morphism $k : (\Omega\times\cdots\times\Omega)^{\Omega^{\cdot^{\cdot^{\cdot^{\Omega^X}}}}} \longrightarrow \Omega$ (with $\Omega\times\cdots\times\Omega$ n times and $\Omega^{\cdot^{\cdot^{\cdot^{\Omega}}}}$ t times, $n, t \geq 0$), abbreviated as $k : \Omega^{n^m} \longrightarrow \Omega$, will be said to be an *abstract n-ary connective* of order m, where $m = 0$ if and only if $\Omega^{\cdot^{\cdot^{\cdot^{\Omega^X}}}} \cong \mathbf{1}$ and $m = (t+1)$ otherwise. Abstract propositions, i.e. morphisms $\mathbf{1} \longrightarrow \Omega$ can thus be considered 0-ary connectives (and can be of any order) with $(\Omega\times\cdots\times\Omega) \cong \mathbf{1}$.

As I have said, '${}_S\Omega$' denotes that n-ary abstract connectives receive their standard well-known names and definitions; '${}_D\Omega$' denotes that n-ary abstract connectives receive their names and definitions as in complement toposes. For convenience, sometimes I will use '${}_S p$' (respectively, '${}_D p$') as a shorthand for $p : \mathbf{1} \longrightarrow {}_S\Omega$ (respectively, $p : \mathbf{1} \longrightarrow {}_D\Omega$) and I will use a similar shorthand for propositional functions. If n-ary bare connectives of a bare topos $\mathfrak{T}$ receive all some intuitive names and definitions, then we are going to call $\mathfrak{T}$ a *concrete topos*, denoted by '$\mathcal{E}$'. '${}_S\mathcal{E}$' and '${}_D\mathcal{E}$' denote arbitrary standard and complement toposes, respectively. Respecting historical priority and if the context prevents any confusion with the abstract notions, I will occasionally omit the subscript in '${}_S\Omega$', '${}_S\mathcal{E}$', '${}_S p$' and '${}_S\varphi$'. Whenever the context allows it, I will also omit the subscript indicating the complement names. In the case of connectives, the differences between the names of the elements of Ω will be indicated by a difference in the connective symbol. For example, the difference between $k : {}_S\Omega^{nm} \longrightarrow {}_S\Omega$ and $k : {}_D\Omega^{nm} \longrightarrow {}_D\Omega$ will be indicated by writing $\sigma_a : \Omega^{nm} \longrightarrow \Omega$, $\sigma_b : \Omega^{nm} \longrightarrow \Omega$, respectively, where σ_a and σ_b are two different symbols. Thus, $\sigma_a(p_1, \ldots, p_n)$ (respectively, $\sigma_b(p_1, \ldots, p_n)$) will stand for $\mathbf{1} \xrightarrow{\langle p_1,\ldots,p_n\rangle} {}_S\Omega^{nm} \xrightarrow{\sigma_a} \Omega$ ($\mathbf{1} \xrightarrow{\langle p_1,\ldots,p_n\rangle} {}_D\Omega^{nm} \xrightarrow{\sigma_b} \Omega$). For simplicity, I will avoid the composition notation in the case of unary and binary connectives (the only ones I will deal with here). That is, instead of writing $k^1 \circ p$ or $k^2 \circ (p, q)$, where k is a connective where the superscript indicates the arity, I will write kp and pkq, respectively.

So ν is called ${}_S$*true* in standard toposes and ${}_D$*false* in complement toposes. Clearly, μ is called ${}_S$*false* in standard toposes and ${}_D$*true* in complement toposes. The assumed name for the special value ν in standard toposes is ${}_S true : \mathbf{1} \longrightarrow {}_S\Omega$ and $false : \mathbf{1} \longrightarrow {}_D\Omega$ in complement toposes. Let us consider now just three unary and three binary bare connectives.

Let be $\nu : \mathbf{1} \longrightarrow \Omega$. Then $\ominus : \Omega \longrightarrow \Omega$ is the bare characteristic morphism of μ:

$$\mathbf{1} \xrightarrow{\mu} {}_S\Omega \underset{\nu_\Omega}{\overset{\ominus =_{def.} \varphi_\mu}{\rightrightarrows}} {}_S\Omega$$

This implies the following truth condition for $\ominus$:

$$\ominus p = \nu \text{ if and only if } p = \mu; \text{ otherwise } \ominus p = \mu\,.$$

It says that $\ominus p$ has one of the special values if and only if p has the other one; otherwise $\ominus p$ has that other special value. For standard toposes it gives

$$\neg p = {}_S true \text{ if and only if } {}_S p = {}_S false, \text{ otherwise } \neg p = {}_S false$$

and for complement toposes

$$\sim p = {}_D false \text{ if and only if } {}_D p = {}_D true, \text{ otherwise } \sim p = {}_D true\,.$$

$\ltimes : \Omega \times \Omega \longrightarrow \Omega$ is defined as the bare characteristic morphism of $\langle \nu, \nu \rangle : \mathbf{1} \longrightarrow \Omega \times \Omega$:

$$\mathbf{1} \xrightarrow{\langle \nu,\nu\rangle} \Omega \times \Omega \underset{\nu_{\Omega\times\Omega}}{\overset{\ltimes =_{def.} \varphi_{\langle \nu,\, \nu\rangle}}{\rightrightarrows}} \Omega$$

This implies that $p \ltimes q = \text{insup}(p, q)$. For standard toposes it gives the diagram

$$\mathbf{1} \xrightarrow{\langle {}_S true,\ {}_S true\rangle} {}_S\Omega \times_S \Omega \underset{{}_S true_{{}_S\Omega\times {}_S\Omega}}{\overset{\wedge =_{def.}\ {}_S\varphi_{\langle {}_S true,\ {}_S true\rangle}}{\rightrightarrows}} {}_S\Omega$$

which implies the following truth condition:

$$(p \wedge q) = \inf({}_S p,\ {}_S q)$$

For complement toposes one has instead the following diagram:

$$\mathbf{1} \xrightarrow{\langle {}_D false,\ {}_D false\rangle} {}_D\Omega \times_D \Omega \underset{{}_D false_{{}_D\Omega\times {}_D\Omega}}{\overset{\curlyvee =_{def.}\ {}_D\varphi_{\langle {}_D false,\ {}_D false\rangle}}{\rightrightarrows}} {}_D\Omega$$

which implies the following truth condition:

$$(p \curlyvee q) = \sup({}_D p,\ {}_D q).$$

$\rtimes : \Omega \times \Omega \longrightarrow \Omega$ is defined as the characteristic morphism of the image of $[\langle \nu\ , \text{id}_\Omega\rangle, \langle \text{id}_\Omega, \nu\rangle] : \Omega + \Omega \longrightarrow \Omega \times \Omega$:

$$\Omega + \Omega \xrightarrow{Im[\langle \nu,\ id_\Omega\rangle, \langle id_\Omega,\ \nu\rangle]} \Omega \times \Omega \underset{\nu_{\Omega\times\Omega}}{\overset{\rtimes =_{def.}\ \varphi_{Im[\langle \nu,\ id_\Omega\rangle,\ \langle id_\Omega,\ \nu\rangle]}}{\rightrightarrows}} \Omega$$

This implies that $p \rtimes q = \text{supin}(p, q)$. For standard toposes, it gives the diagram

$${}_S\Omega +_S \Omega \xrightarrow{Im\left[\left\langle {}_S true, id_{{}_S\Omega}\right\rangle, \left\langle id_{{}_S\Omega}, {}_S true\right\rangle\right]} {}_S\Omega \times_S \Omega \underset{{}_S true_{{}_S\Omega\times {}_S\Omega}}{\overset{\vee =_{def.}\ {}_S\varphi_{Im\left[\left\langle {}_S true, id_{{}_S\Omega}\right\rangle, \left\langle id_{{}_S\Omega}, {}_S true\right\rangle\right]}}{\rightrightarrows}} {}_S\Omega$$

which implies the following truth condition:

$$(p \vee q) = \sup({}_S p,\ {}_S q).$$

For complement toposes the truth condition is

$$(p \curlywedge q) = \inf({}_D p,\ {}_D q)$$

which is expressed in the following diagram:

$${}_D\Omega +_D \Omega \xrightarrow{Im\left[\left\langle {}_D false, id_{{}_D\Omega}\right\rangle, \left\langle id_{{}_D\Omega}, {}_D false\right\rangle\right]} {}_D\Omega \times_D \Omega \underset{{}_D false_{{}_D\Omega\times {}_D\Omega}}{\overset{\curlywedge =_{def.}\ {}_D\varphi_{Im\left[\left\langle {}_D false, id_{{}_D\Omega}\right\rangle, \left\langle id_{{}_D\Omega}, {}_D false\right\rangle\right]}}{\rightrightarrows}} {}_D\Omega$$

$\rightleftharpoons : \Omega \times \Omega \longrightarrow \Omega$ is defined as the characteristic morphism of $\epsilon : \preceq \longrightarrow \Omega \times \Omega$ (the equalizer of $\ltimes : \Omega \times \Omega \longrightarrow \Omega$ and the first projection p):

$$\mathfrak{D} \overset{e}{\rightarrowtail} \Omega \times \Omega \underset{\nu_{\Omega \times \Omega}}{\overset{\rightleftharpoons \, =_{def.} \, \varphi_e}{\rightrightarrows}} \Omega$$

In standard toposes, one has the morphism $\Rightarrow : {}_S\Omega \times {}_S\Omega \longrightarrow {}_S\Omega$, defined as the characteristic morphism of $e : \leq \longrightarrow {}_S\Omega \times {}_S\Omega$, the equalizer of $\wedge : {}_S\Omega \times {}_S\Omega \longrightarrow {}_S\Omega$ and the first projection p:

$${}_S\leq \overset{e}{\rightarrowtail} {}_S\Omega \times_S \Omega \underset{{}_S true_{{}_S\Omega \times {}_S\Omega}}{\overset{\Rightarrow \, =_{def.} \, {}_S\varphi_e}{\rightrightarrows}} {}_S\Omega$$

whose corresponding truth condition is the following one:

$(p \Rightarrow q) = {}_S true$ if and only if ${}_S p \leq {}_S q$ (equivalently, if and only if $(p \wedge q) = p$); otherwise $(p \Rightarrow q) = q$

In complement toposes, *dual-implication*, *subtraction* or *pseudo-difference* $- : {}_D\Omega \times {}_D\Omega \longrightarrow {}_D\Omega$ is defined as the characteristic morphism of $\bar{e} : \geq \longrightarrow {}_D\Omega \times {}_D\Omega$, where $\bar{e}$ is the equalizer of $\curlyvee : {}_D\Omega \times {}_D\Omega \longrightarrow {}_D\Omega$ and the first projection p:

$${}_D\geq \overset{\bar{e}}{\rightarrowtail} {}_D\Omega \times_D \Omega \underset{{}_D false_{{}_D\Omega \times {}_D\Omega}}{\overset{- \, =_{def.} \, {}_D\varphi_{\bar{e}}}{\rightrightarrows}} {}_D\Omega$$

which implies the following truth condition:

$(p - q) = {}_D false$ if and only if ${}_D p \geq {}_D q$ (equivalently, if and only if $(p \curlyvee q) = p$); otherwise $(p - q) = q$).

Finally, let us consider two first-order unary connectives. $\binom{\mathfrak{A}}{\mathfrak{S}}_X : \Omega^X \longrightarrow \Omega$ is defined as the characteristic morphism of $\lambda_x.\nu_X$, the 'name' (exponential transposition) of $\nu_X \circ pr_X : \mathbf{1} \times X \longrightarrow X \longrightarrow \Omega$:

$$\mathbf{1} \xrightarrow{\lambda_x.\nu_X} {}_S\Omega^X \underset{\nu_{\Omega^X}}{\overset{\binom{\mathfrak{A}}{\mathfrak{S}}_X \, =_{def.} \, \varphi_{\lambda_x.\nu_X}}{\rightrightarrows}} \Omega$$

This implies the following truth condition:

$$\binom{\mathfrak{A}}{\mathfrak{S}}_X \varphi(x) = \text{insup}(\varphi(x)) \,.$$

It says that $\binom{\mathfrak{A}}{\mathfrak{S}}_X \varphi(x)$ has the special value ν if and only if $\varphi(x)$ has that special value for all instances of x (in the domain X). A more illustrative way to put it is:

$$\binom{\mathfrak{A}}{\mathfrak{S}}_X \varphi(x) = \nu \ \text{ if and only if } \ \varphi(x) = \nu, \ \text{ for all } \ x \,.$$

For standard toposes it gives the following diagram:

$$\mathbf{1} \xrightarrow{\lambda_x.{}_S true_X} {}_S\Omega^X \underset{{}_S true_{{}_S\Omega^X}}{\overset{\forall_X \, =_{def.} \, {}_S\varphi_{\lambda_x.{}_S true X}}{\rightrightarrows}} {}_S\Omega$$

according to which $\forall_X \varphi(x) = {}_S true$ if and only if ${}_S\varphi(x) = {}_S true$, for all x. The exact truth condition implied by the above definition is

$$\forall_X \varphi(x) = \inf(\varphi(x)) .$$

For complement toposes one has instead the following truth condition:

$$E_X \varphi(x) = {}_D false \text{ if and only if } {}_D\varphi(x) = {}_D false, \text{ for all } x,$$

or more exactly

$$E_X \varphi(x) = \sup({}_D\varphi(x))$$

embodied in the following diagram:

$$\mathbf{1} \xrightarrow{\lambda x \cdot {}_D false_X} {}_D\Omega^X \underset{{}_D false_{{}_D\Omega^X}}{\overset{E_X =_{def.} {}_D\varphi_{\lambda x \cdot {}_D false X}}{\rightrightarrows}} {}_D\Omega$$

$(^{\mathfrak{S}}_{\mathfrak{A}})_X : \Omega^X \longrightarrow \Omega$ is defined as the characteristic morphism of the composite $p_X \circ \in_X (\in)$: $\in \rightarrowtail \Omega^X \times X \longrightarrow \Omega^X$ (where p_x is the first projection and $\in_X$ is the subobject of $\Omega^X \times X$ whose character is the evaluation morphism $e_X : \Omega^X \times X \longrightarrow \Omega$):

$$p_X \circ \in_X (\in) \xrightarrow{Im(p_X \circ \in_X)} \Omega^X \underset{\nu_{\Omega^X}}{\overset{(^{\mathfrak{S}}_{\mathfrak{A}})_X =_{def.} \varphi_{Im(p_X \circ \in_X)}}{\rightrightarrows}} \Omega$$

This implies the following truth condition:

$$(^{\mathfrak{S}}_{\mathfrak{A}})_X \varphi(x) = \nu \text{ if and only if } \varphi(x) = \nu, \text{ for some } x$$

or, more exactly,

$$(^{\mathfrak{S}}_{\mathfrak{A}})_X \varphi(x) = insup(\varphi(x)).$$

It says that $(^{\mathfrak{A}}_{\mathfrak{S}})_X \varphi(x)$ has the special value ν if and only if $\varphi(x)$ has that special value for some instances of x (in the domain X).

For standard toposes one has then the following diagram:

$${}_S p_X \circ \in_X (\in) \xrightarrow{Im({}_S p_X \circ {}_S\in_X)} {}_S\Omega^X \underset{{}_S true_{{}_S\Omega^X}}{\overset{\exists_X =_{def.} {}_S\varphi_{Im({}_S p_X \circ {}_S\in_X)}}{\rightrightarrows}} {}_S\Omega$$

and the following truth condition:

$$\exists_X \varphi(x) = {}_S true \text{ if and only if } {}_S\varphi(x) = {}_S true, \text{ for some } x .$$

or more exactly,

$$\exists_X \varphi(x) = \sup(\varphi(x)),$$

whereas for complement toposes one has the following truth conditions:

$A_X\varphi(x) = \inf(\varphi(x))$, that is, $A_X\varphi(x) = {}_Dfalse$ if and only if ${}_D\varphi(x) = {}_Dfalse$, for some x

and the corresponding diagram:

$$ {}_Dp_X \circ \in_X (\in) \xrightarrow{Im({}_Dp_X \circ {}_D\in_X)} {}_D\Omega^X \underset{{}_Dfalse_{{}_D\Omega^X}}{\overset{A_X =_{def.} {}_D\varphi_{Im({}_Dp_X \circ {}_D\in_X)}}{\rightrightarrows}} {}_D\Omega $$

5.2 Logical m-Valuedness

In the previous section, I showed how to internalize certain Suszkian ideas. I also argued that the notion of logical consequence in toposes is externally assumed by the theorists rather than internally imposed by their categorial structure. However, I left as an open problem the issue of describing and proving the existence of objects suitable to internalize those other notions of consequence. Without trying to settle that question here, I will probe an idea at least for consequences based on a form of forwards preservation.

I say that a topos is algebraically n-valued if there are n morphisms from $\mathbf{1}$ to Ω. A topos is said to be *logically m-valued* if the assumed notion of consequence, $\Vdash$, implies that there are m logical values. Logically m-valuedness is *internalized* if

(1) there is an object $\mathfrak{V}$ such that it is the codomain of exactly m morphisms with domain $\mathbf{1}$ such that to each logical value implied by $\Vdash$ corresponds one and only one morphism from $\mathbf{1}$ to $\mathfrak{V}$; and
(2) there is a unique morphism $\mathfrak{sep} : \Omega \longrightarrow \mathfrak{V}$ such that $\mathfrak{sep}$ satisfies the following properties:
   (2.1) For every $\delta_i : \mathbf{1} \longrightarrow \mathfrak{V}$ there is a $p : \mathbf{1} \longrightarrow \Omega$ such that $\mathfrak{sep} \circ p = \delta_i$.
   (2.2) If $p \Vdash q$ implies that p and q have certain $\Vdash$-logical values v_i and v_j, respectively, then if $\mathfrak{sep} \circ p = \delta_i$, $\mathfrak{sep} \circ q = \delta_j$ (where δ_i corresponds to v_i and δ_j corresponds to v_j).

The morphisms $\delta_1, \ldots \delta_m$ can be collectively denoted by m-val and are called *a logical m-valuation (based on $\Vdash$)*. Thus, the diagram below commutes according to the definition of m-val just given and conditions (1) and (2):

$$ \begin{array}{ccc} \mathbf{1} & \xrightarrow{p} & \Omega \\ & \searrow^{m\text{-val}} & \downarrow \mathfrak{sep} \\ & & \mathfrak{V} \end{array} $$

The morphism δ_i such that $\mathfrak{sep} \circ true_{\mathfrak{T}} = \delta_i$ will be called 'morphism of designated values' and will be denoted by 'δ^+'; the morphism δ_j such that $\mathfrak{sep} \circ false_{\mathfrak{T}} = \delta_j$ will be called 'morphism of antidesignated values' and will be denoted by 'δ^{+}'. A similar procedure has to be followed to individuate each additional logical value, if any.

Again, an open problem is to determine which kind of toposes are logically m-valued in the sense defined above and to provide a definition that could encompass all toposes.

5.3 Bare Internal Logic

A *(bare) internal logic* $\mathfrak{L}_{\mathfrak{T}}$ of a topos $\mathfrak{T}$ is a tuple $\langle o_{\Omega_{\mathfrak{T}}1}, \ldots, o_{\Omega_{\mathfrak{T}}n}; \mathfrak{N}; \Vdash; {}_{\mathfrak{T}}\delta_1, \ldots, {}_{\mathfrak{T}}\delta_m \rangle$, where the $o_{\Omega_{\mathfrak{T}}i}$ are operations of $\Omega_{\mathfrak{T}}$ (Ω in $\mathfrak{T}$), $\mathfrak{N}$ is a collection of labels or names for those operations, $\Vdash$ is a consequence relation and the ${}_{\mathfrak{T}}\delta_j$ are morphisms of logical values. As we have seen, the common place stating that the internal logic of a topos is intuitionistic is grounded on a theorem which presupposes that

(${}_S\mathfrak{N}$) the names for the morphisms with codomain Ω are the standard ones,
(${}_{\text{top}S}\delta^{+}$) ${}_S true$ is the only designated value, and
($\Vdash^{T}$) the underlying notion of consequence is *Tarskian*.

The same bare categorial structure of toposes may support several internal logics, depending on what particular $\mathfrak{N}$, $\Vdash$ and ${}_{\mathfrak{T}}\delta_j$ are considered. For example, the internal logic of a complement topos satisfies ($\Vdash^{T}$) but not (${}_S\mathfrak{N}$) (hence, it does not exactly satisfy (${}_{\text{top}}\,{}_{\mathfrak{T}}\delta^{+}$), either). Instead, a complement topos satisfies

(${}_D\mathfrak{N}$) the names of the morphisms with codomain Ω are described by $\varrho({}_S p)$,
(${}_{\text{top}D}\delta^{+}$) ${}_D true$ is the only designated value.

6 Conclusions

> Thus we seem to have partially demonstrated that even in foundations not Substance but invariant Form is the carrier of the relevant mathematical information.
> —William Francis Lawvere, *An elementary theory of the category of sets*.

I have expounded the basics of the standard theory of topos logic, which leads to set the following slogans:

(S1) Ω is (or at least can be thought of as) a truth values object.
(S2) In general, the internal logic of a topos is many-valued.
(S3) In general, the internal logic of a topos is (with some provisos) intuitionistic.
(S4) Intuitionistic logic is the objective logic of variable sets.
(S5) The universal, invariant laws of mathematics are intuitionistic.

However, complement toposes give rise to doubts about the standard description of the internal logic of a topos. Specifically, they directly go against (S3), (S4) and (S5), and allow us to raise doubts about the other slogans. Thus, complement toposes, with their paraconsistent internal logic, suggest that current topos theory gives us just part of the concept of topos, that some common theorems on topos logic tell just part of the relevant story and that in a further, more abstract development the slogans above have just limited application and that logicality lies beyond any particularity of a logic.

Moreover, I gave a categorial version of Suszko's thoughts on many-valuedness. According to Suszko's reduction, every Tarskian logic (a logic whose consequence relation is reflexive, transitive and monotonic) has a bivalent semantics, which implies the rejection of (S1) and (S2). I did this working mainly with standard toposes, but the result can be easily transferred to complement toposes. I showed that (S2) can be maintained (with 'can be many-valued' instead of 'is many-valued') changing the underlying notion of logical consequence, but this does not save the other slogans and, indeed, can be used as a further case against them.

I described how do a topos and an internal logic look like once all the non-categorial assumptions that gave rise to (S1)–(S5) are removed. Unsurprisingly, it turns out that a topos, and especially its internal logic, is a truly Protean categorial creature which can accommodate the most diverse descriptions and support an enormous variety of logics besides that mentioned in the slogans. For reasons of space I could not discuss other important parts of topos theory, like Kripke and sheaf semantics or Lawvere–Tierney topologies, but they are subject to dualization too, and also to the step into abstraction.

If I was to use the Hegelian–Lawverean terminology, I would say that this is what toposes are, and this is what their internal logic is. The labels and special names are 'subjective', 'substance' befouling the objective and invariant forms. Standard toposes do not provide us the full story about variable sets and, again, the objective logic of variable sets would be that surviving all the variations in the parameters $\mathfrak{N}$ (the names for certain parts of the equational structure), $\Vdash$ (the underlying notion of logical consequence), ${}_{\mathfrak{T}}\delta_j$ (the adopted logical values). But even if my attempted step into abstraction were misguided and my attack on the presuppositions fell short and they were really unshakeable, I think I have succeeded in highlighting them and, thus, in contributing to make clearer the foundations of topos logic. The best of the study of the interactions between toposes, philosophy and logic is yet to come.

Acknowledgement I want to thank for the support from the CONACyT project CCB 2011 166502 "Aspectos filosóficos de la modalidad". Diagrams were drawn using Paul Taylor's diagrams package v. 3.94.

References

1. Awodey, S.: Structure in mathematics and logic: A categorical perspective. Philos. Math. **4**(3), 209–237 (1996)
2. Awodey, S.: Continuity and logical completeness: An application of sheaf theory and topoi. In: van Benthem, M.R.J., Heinzmann, G., Visser, H. (eds.) The Age of Alternative Logics. Assesing Philosophy of Logic and Mathematics Today, pp. 139–149. Springer, Dordrecht (2006)
3. Bell, J.L.: Lectures on the Foundations of Mathematics. Available at publish.uwo.ca/∼jbell/foundations%20of%20mathematics.pdf
4. Bell, J.L.: From absolute to local mathematics. Synthese **69**, 409–426 (1986)
5. Bell, J. L.: Toposes and Local Set Theories: An Introduction. Clarendon Press, Oxford (1988)

6. 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)
7. Béziau, J.-Y.: From paraconsistent logic to universal logic. Sorites **12**, 5–32 (2001)
8. Béziau, J.-Y.: S5 is a paraconsistent logic and so is first-order classical logic. Log. Investig. **9**, 301–309 (2002)
9. de Queiroz, G.d.S.: On the Duality between Intuitionism and Paraconsistency, in Portuguese. PhD thesis, Universidade Estadual de Campinas (1998)
10. Estrada-González, L.: Quaternality in a topos-theoretical setting, unpublished typescript (2011)
11. Estrada-González, L.: Complement-topoi and dual intuitionistic logic. Australasian J. Log. **9** 26–44 (2010)
12. Estrada-González, L.: Fifty (more or less) shades of logical consequence. In: Arazim, P., Dančák, M. (eds.) LOGICA Yearbook 2014. College Publications, London (2015)
13. Frankowski, S.: Formalization of a plausible inference. Bull. Sect. Log. **33**, 41–52 (2004)
14. Frankowski, S.: P-consequence versus q-consequence operations. Bull. Sect. Log. **33**, 197–207 (2004)
15. Goldblatt, R.: Topoi: The Categorial Analysis of Logic. Studies in Logic and the Foundations of Mathematics, vol. 98, revised edition. North Holland Publishing, Amsterdam (1984)
16. James, W., Mortensen, C.: Categories, sheaves, and paraconsistent logic. Unpublished typescript (1997)
17. Kotas, J., da Costa, N.C.A.: Some problems on logical matrices and valorizations. In: Arruda, A., da Costa, N.C.A., Sette, A.M. (eds.) Proceedings of the Third Brazilian Conference on Mathematical Logic, pp. 131–145. Sociedade Brasileira de Lógica, São Paulo (1980)
18. Lawvere, F.W.: Continuously variable sets: Algebraic geometry = geometric logic. In: Rose, H.E., Shepherdson, J.C. (eds.) Logic Colloquium '73 (Bristol, 1973), pp. 135–156. North Holland, Amsterdam (1975)
19. Lawvere, F.W.: Variable quantities and variable structures in topoi. In: Heller, A., Tierney, M. (eds.) Algebra, Topology, and Category Theory. A Collection of Papers in Honor of Samuel Eilenberg, pp. 101–131. Academic Press, New York (1976)
20. Lawvere, F.W.: Intrinsic co-Heyting boundaries and the Leibniz rule in certain toposes. In: Carboni, A., Pedicchio, M.C., Rosolini, G. (eds.) Proceedings of the 1990 Meeting on Category Theory held in Como, Italy, Lecture Notes in Mathematics, vol. 1488, pp. 279–281. Springer, Berlin (1991)
21. Lawvere, F.W., Rosebrugh, R.: Sets for Mathematics. Cambridge University Press, Cambridge (2003)
22. Lawvere, F.W., Schanuel, S.H.: Conceptual Mathematics. A First Introduction to Categories, second reprint. Cambridge University Press, Cambridge (2000)
23. Mac Lane, S., Moerdijk, I.: Sheaves in Geometry and Logic: A First Introduction to Topos Theory. Springer, New York (1992)
24. G. Malinowski. Q-consequence operation. Rep. Math. Log. **24**, 49–59 (1990)
25. Malinowski, G.: Towards the concept of logical many-valuedness. Folia Philos. **7**, 7–103 (1990)
26. McLarty, C.: Elementary Categories, Elementary Toposes. Oxford Clarendon Press, Toronto (1995)
27. McLarty, C.: Two constructivist aspects of category theory. Philos. Sci. **27** (Cahier spécial 6), 95–114 (2006)
28. Mortensen, C.: Inconsistent Mathematics. Kluwer Mathematics and Its Applications Series. Kluwer Academic Publishers, Dordrecht (1995)
29. Mortensen, C.: Closed set logic. In: Brady, R.T. (ed.) Relevant Logics and Their Rivals, vol. II, pp. 254–262. Ashgate Publishing, Aldershot (2003)
30. Reyes, G., Zawadowski, M.: Formal systems for modal operators on locales. Stud. Log. **52**(4), 595–613 (1993)
31. Reyes, G., Zolfaghari, H.: Bi-Heyting algebras, toposes and modalities. J. Philos. Log. **25**(1), 25–43 (1996)
32. Routley, R., Meyer, R.K.: Every sentential logic has a two-valued worlds semantics. Log. Anal. **19**(74–76), 345–365 (1976)
33. Scott, D.: Background to formalization. In: Leblanc, H. (ed.) Truth, Syntax and Modality, pp. 244–273. North Holland, Amsterdam (1973)

34. Scott, D.: Completeness and axiomatizability in many-valued logics. In: Henkin, L., et al. (eds.) Proceedings of the Tarski Symposium, pp. 411–436. American Mathematical Society, USA (1974)
35. Tsuji, M.: Many-valued logics and Suszko's thesis revisited. Stud. Log. **60**(2), 299–309 (1998)
36. van Inwagen, P.: Metaphysics, 3rd edn. Westview Press, Boulder, Colorado (2008)
37. Vasyukov, V.: Structuring the universe of universal logic. Log. Univers. **1**(2), 277–294 (2007)
38. Vigna, S.: A guided tour in the topos of graphs. Technical Report 199–7, Università di Milano, Dipartimento di Scienze dell'Informazione (1997). Available at http://vigna.dsi.unimi.it/papers.ToposGraphs.pdf
39. Wansing, H., Shramko, Y.: Suszko's thesis, inferential many-valuedness and the notion of logical system. Stud. Log. **88**(1), 405–429 (2008). See also the erratum in vol. 89, p. 147 (2008)

L. Estrada-González (✉)
Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México, Mexico City, Mexico
e-mail: loisayaxsegrob@gmail.com

A Note on the Internal Logic of Constructive Mathematics: The Gel'fond–Schneider Theorem in Transcendental Number Theory

Yvon Gauthier

Abstract The question of an internal logic of mathematical practice is examined from a finitist point of view. The Gel'fond–Schneider theorem in transcendental number theory serves as an instance of a proof-theoretical investigation motivated and justified by constructivist foundations of logic and mathematics. Constructivist notions are emphasized by contrasting the arithmetical proof procedure of infinite descent with the principle of transfinite induction. It is argued that intuitionistic logic cannot alone provide secure foundations for constructivist mathematics and a finitist logic is briefly sketched in the framework of polynomial arithmetic.

Keywords Constructive mathematics · Finitism · Infinite descent · Number theory

Mathematics Subject Classification (2000) Primary 03F50 · Secondary 11J81

1 Introduction. The Logical Problem

To illustrate the excluded middle principle of classical logic, Dummett [4, p. 6] cites the well-known example of a theorem in classical logic which gives a rational solution to two irrational numbers x^y:

There are solutions of $x^y = z$ with x and y irrational and z rational. Take $\sqrt{2}^{\sqrt{2}}$ as either rational or irrational:

If $\sqrt{2}^{\sqrt{2}}$ is rational, take (as irrational) $x = \sqrt{2}$ and $y = \sqrt{2}$, so that $z = \sqrt{2}^{\sqrt{2}}$ which by hypothesis is rational.

If on the other hand $\sqrt{2}^{\sqrt{2}}$ is irrational, put $x = \sqrt{2}^{\sqrt{2}}$ and $y = \sqrt{2}$ so that $z = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = (\sqrt{2})^2$, which is certainly rational.

Thus, in either case a solution exists, but the excluded middle principle is unable to tell us what is the case. Intuitionistic logic requires that one of the disjuncts of a disjunction be proven or instantiated (as for the existential quantifier). Of course, we know now that $\sqrt{2}^{\sqrt{2}}$ is irrational and even transcendental, due to the Gel'fond–Schneider theorem: here an intuitionist logician like A.S. Troelstra (see his *Proof theory and Constructivism* [26]) is content to say that there is a constructive proof and simply refers to the Gel'fond–Schneider theorem without further comment!

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_13

The Gel'fond–Schneider theorem (1934) is a solution to Hilbert's seventh problem:

For α and β algebraic numbers with $\alpha \neq 0, 1$ and β irrational, α^β is transcendental.

Schneider [23] puts the problem in the form

For ω an algebraic number $\neq 0, 1$ and θ irrational, ω^θ is transcendental with $\theta = \log\ \eta / \log\ \omega$ for $\eta = \omega^\theta$.

The proof proceeds by logarithmic approximations to algebraic numbers and concludes that any logarithm of an algebraic number with an algebraic base must be a transcendental or a rational number. By *reductio ad absurdum*, ω^θ is transcendental. With a similar procedure, Gel'fond [11] comes to the conclusion that the logarithms of algebraic numbers with an algebraic base are transcendental or rational numbers.

Both proofs are (partially) constructive in the sense that they extract arithmetical content (that is logarithmic approximations, minorizations and majorizations, effective bounds, etc.) from analytical methods (infinite series or power series, periodic functions, etc.); they use polynomial inequalities for the rational values of an analytic function $f(x)$ and end up by contradicting an assumption to the effect that finite values make it vanish identically $f(x) = 0$.

Baker [1] extended and generalized those results in transcendental number theory using auxiliary functions or polynomials – which he calls fundamental polynomials – linear forms and logarithms for approximations of algebraic numbers in order to establish algebraic independence by contradiction or *reductio ad absurdum*.

It is only in 1962 that Gel'fond produced an elementary (constructive) proof for real algebraic numbers $\omega, a > 0, b$ for e^ω with e the base of natural logarithms and a^b (see [13, Chap. 12]. Gel'fond relates that he used only Rolle's theorem as an analytical tool – here a radical constructivist could mention that a constructive version of Rolle's theorem is to be found in [2]. Essentially, Rolle's theorem says classically that a continuous function of a real variable on $[a, b]$ with $f(a) = f(b)$ for $a < b$ has a derivative $f'(c) = 0$. Bishop's constructive version introduces $|f'(x)| \leq \epsilon$ with $\epsilon > 0$ for moduli of continuity of f' and differentiability of f. In other words, Bishop defines more precisely the limits of the real interval $[a, b]$ much in the manner of Kronecker for Bolzano's theorem on intermediate values (see [7]). Bishop has admitted ([3]) that his foundational project was closer to Kronecker's finitist programme than to Brouwer's intuitionism.

However, Gel'fond's work is in analytic number theory and constructive number theory, not in constructive analysis. His results in transcendental number theory are algebraic in nature. The main theorem in [13, Chap. 12] states that "if $\omega \neq 0$ is an algebraic real, than e^ω is not algebraic" and is couched in the language of algebraic integers in finite fields. A finite field is also the arena for an another elementary proof ([13, Chap. 10]), Hasse's theorem on integral solutions for the equation:

$$y^2 \equiv x^3 + ax + b \pmod{p} \tag{1.1}$$

for integers a, b and a prime $p > 3$. Gel'fond formulates his solution in terms of an inequality

$$|N - p| < 2\sqrt{p} \tag{1.2}$$

where N is the number of integral solutions of the equation. Here, the language used is the language of polynomials and divisors with an algebro-geometric interpretation and

Gel'fond quotes a major result of André Weil on Riemann's hypothesis in function fields (see [28]). Weil's result is a special case of the Riemann hypothesis for quadratic finite fields (with a finite number of elements or points on a projective surface) and Weil claims that his result is free of the transcendental (analytic) theory. The main arithmetical tool here is the theory of forms or homogeneous polynomials. That theory has been developed first in great generality by Kronecker in his "*Allgemeine Arithmetik*" or General Arithmetic and Weil has repeatedly refered to Kronecker as the founding father of algebraic – arithmetic geometry on finite fields. Kronecker's theory of forms is equivalently a divisor theory (of modular systems) for which infinite descent works, since homogeneous polynomials are finite (integral and rational) functions with integer coefficients and indeterminates (variables). Fermat's infinite or indefinite descent, as Fermat qualified it, is in fact a finite process.

Weil describes infinite descent in the following:

> Infinite descent *à la* Fermat depends ordinarily upon no more than the following simple observation: if the product $\alpha\beta$ of two ordinary integers (resp. two integers in an algebraic number-field) is equal to an mth power, and in the g.c.d. of α and β can take its values only in a given finite set of integers (resp. of ideals), then both α and β are mth powers up to factors which take their values only in some assignable finite set. For ordinary integers this is obvious; it is so for algebraic number-fields provided one takes for granted the finiteness of the number of ideal-classes and Dirichlet's theorem about units. In the case of a quadratic number-field $\mathbb{Q}(\sqrt{N})$, this can be replaced by equivalent statements about binary quadratic forms of discriminant N. (see [29, pp. 335–336])

What interests us here in this modern terminology is the finiteness results and the character of *effectivity* that attaches to the proofs by infinite descent. In that context, Fermat's infinite descent is a generalized Euclidean algorithm or a division algorithm in finite number fields and in finite function fields. The original idea of Fermat's infinite or indefinite descent appears in his 1670 commentary on Diophantus:

> Eodem ratiocinio dabitur et minor istâ inventa per viam prioris, et semper in infinitum minores invenientur numeri in integris idem praestantes. Quod impossibile est quia, dato numero quovis integro, non possunt dari infiniti in integris illo minores ([5]).

I translate the last quotation as

> By the same calculation it is supposed that a smaller number is found in a descending procedure and that one can always find numbers smaller than the preceding one *ad infinitum*, which is impossible, since for an arbitrary integer there cannot be found an infinity of smaller ones in integers.

Let us remark that the method of infinite descent can be applied to a variety of problems, starting with the proof of the irrationality of $\sqrt{2}$ or the impossibility of

$$x^4 + y^4 = z^2 \tag{1.3}$$

for all $z > 0$ and and $x, y \neq 0$. Infinite or indefinite descent is, in fact, finite; it does not transcend the finite and the *reductio ad absurdum* is innocuous here, since the ensuing double negation is finitary. The finiteness of the procedure is still more evident when it is applied to "positive"questions, as Fermat says. Take the theorem: "Any prime number which is greater than a multiple of 4 by one must be composed of two squares." If there was such a prime number greater than a multiple of 4 by 1, but which would not be composed of squares, there would be a smaller one of the same nature and still smaller ones till one reaches 5, which is the smallest number having the said property. One must then conclude that the theorem is true. What we have here is simply a generalization of Euclid's division algorithm, but it has been used variedly from Fermat to contemporary arithmetic geometry as a proof-theoretical device of reduction (Legendre's term), Kronecker's elimination theory or decomposition of forms (homogeneous polynomials) in divisor theory and diverse descent techniques as in Grothendieck's programme (see my forthcoming paper [10] for details). Of course, not all those techniques are effective, even in number theory, that is, they do not necessarily provide calculations with explicit bounds. In some cases, like in n-category theory (infinity or ω-categories), descent can be encapsulated as a fly down escape from the aether of higher dimensional categories, the totality of which live in the Ω-universe of all ordinals – described by Cantor as an absolutely inconsistent plurality "*eine absolut inkonsistente Vielheit*." But even without the Ω totality of universes, categorical foundations still need an inaccessible cardinal of higher set theory, that is transfinite induction beyond ϵ_0, as V. Voevodsky admits in his univalent foundations for homotopy type theory (see [27]). The same is true for Martin-Löf intuitionistic (or so-called constructive) type theory with dependent or contextual types together with the program proof assistant Coq also needed in Voevodsky's categorical foundations. However, descent is still lurking in the background under the clothings of the axiom of foundation introduced by von Neumann for the cumulative rank structure of axiomatic set theory. Mochizuki [22] is keen on keeping the axiom of foundation as a descent procedure in the set-theoretic foundations of his ambitious programme of interuniversal geometry for the putative proof of the so-called *ABC* conjecture in number theory.

I would qualify such foundational programmes as *descriptive* – like in descriptive set theory – as opposed to *reductive* foundational programmes, meaning that foundations should incorporate a critical evaluation and a justification for mathematical practice hopefully within mathematics itself with a minimal (constructivist) philosophy, not just a unifying language. Although descriptive theories may have a computational or algorithmic intent, as in Voevedsky's univalent foundations or in Mochizuky interuniversal geometry, the abstract or general framework in higher category theory or higher topos theory, which both need transfinite induction (and recursion), is not constructive or so feebly constructive that the computational output seems to be a by-product, rather than a natural outcome of the theory – the algorithmic results are most of the time grounded on polynomial arithmetic as a basis for higher (alien) structures or creatures ! Note that the *ABC* conjecture mentioned above has been demonstrated for polynomials constructively, that is by elementary (nonanalytic) means.

2 The Internal Logic

The logical outcome of this can be stated in a few words : both disjuncts in number theory must have a number-theoretic content, while intuitionistic logic requires only that one of the disjuncts be instantiated in order that a disjunction may have a truth-value or rather a verification value – the same for the existential quantifier. What our example shows is that a constructivist logic must be dependent upon an external resource, a numerical content, in order to be effective in arithmetic. This means that the logic is derived and comes after or is posterior to what Hilbert called '*inhaltliches logische Schließen*', a terminology I have translated as 'internal logic' (see [7] Gauthier 2002 for details). Formal logic, in Hilbert's view, was an external metamathematical means '*äußeres Handeln*' to treat the internal logic of mathematical theories or the inferences pertaining to mathematics proper, not to the metamathematics or the proof theory of formal systems.

But if we follow Hilbert's lead, formal logic must depend upon mathematics' inner workings and Brouwer, who is not a Hilbertian by any means, would follow suit and concur by saying that the excluded middle principle is not admissible particularly in the mathematical analysis of infinite sequences or infinitely proceeding sequences (of natural numbers). Kolmogorov was well aware of the significance of intuitionistic logic which he interpreted as a logic of problems (and solutions) and he made a distinction between real mathematics and the classical mathematics of pseudo-truth (*pseudoistinosti*) where transfinite induction is operative. Problems, in Kolmogorov's mind, are essentially well-posed mathematical problems, since in the Hilbertian spirit, they must have a solution. We know that Brouwer was suspicious of logic and we could ask if intuitionistic logic, following Heyting and his successors, is faithful to Brouwer's original intent. I give only one example or counterexample in line with Brouwer's practice, the attempt to confound Gentzen's transfinite induction with Fermat's infinite descent; the former is designed to provide a consistency proof for Peano arithmetic while, the latter is a constructive method of proof in classical number theory, from Fermat, Euler, Gauss, Lagrange, Legendre, Kummer, Kronecker to contemporary number theory and algebraic–arithmetic geometry in the hands of Mordell, Hermann Weyl, André Weil, Gert Faltings among others.

3 Descent or Descending Induction

Infinite descent in classical number theory from Fermat to Kronecker and Weil is not infinite descent in the set-theoretic setting of an infinite set of natural numbers. It is in fact a finite arithmetical procedure that has little to do with the transarithmetical process of transfinite induction.

As A. Baker explains in in his major work on transcendental number theory ([1]), Gel'fond's and Schneider's proofs proceed by construction of auxiliairy functions and polynomials and they then derive their results by induction on an arbitrary (finite) large integer n by assuming that if the result holds for $n - 1$, it holds for all n. In the same vein, J.-P. Serre defines descending induction as "acting on two (positive) integers m and N with $m > N$ descending to N" (see [21]). For integers or numerical predicates, the procedure looks like

$$\begin{gathered} Ax_n \\ Ax_{n-1} \\ \vdots \\ Ax_{n-(n-1)} \\ Ax_0 = Ax_{n-n}\,. \end{gathered} \tag{3.1}$$

Ax_n is what Bourbaki has called a general term. While classical infinite descent works with the descending sequence of finite ordinals (natural numbers), designated as *weak* well-ordering by Kreisel, transfinite induction calls for the *strong* well-ordering of transfinite ordinals up to ϵ_0 of Cantor's second number class. This well-ordering has to deal with the subsets of $\mathbb{N}$. Here the logician must pick up a certain quantifier-free subset of the ordinal ω hierarchy, because he knows that the set of all countable ordinals of the ω's including the ϵ's is the uncountable ω_1 corresponding to the cardinal $\aleph_1$ – Cantor suggested that the continuum c was $\aleph_1 = 2^{\aleph_0}$. The power set of the set of natural numbers is significant in that connexion, since the well-ordering principle says

$$\forall S \subseteq \mathbb{N}(S \neq 0 \wedge Ax(x \in S)) \rightarrow (\exists y < x \wedge y \in S)\,, \tag{3.2}$$

that is, there exists a strictly decreasing sequence for all elements of the subsets of the set $\mathbb{N}$ of natural numbers. This is the *strong strict* well-ordering of $\mathbb{N}$ requiring the excluded middle on the power set of $P(\mathbb{N}) = 2^{\aleph_0}$! Transfinite induction runs along an initial segment of the ω-sequence *beyond* the first ω up to its limit ϵ_0, while infinite descent starts with an arbitrary integer n *below* the first ω. Here is one main cleavage between finitist and infinitist proof theory, but Gentzen wanted to believe that classical infinite descent was a disguised form of complete induction in order to justify transfinite induction over the denumerable ordinals (see [14]). Following suit, Kreisel has simply noticed that it was a form of infinite descent that Gentzen had used (see [18]).

Polynomials as the finite support of infinite power series are the natural extension of natural numbers and provide with a finitist alternative to set-theoretic Peano–Dedekind arithmetic. For polynomials of the form

$$P(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \tag{3.3}$$

in decreasing powers for integer coefficients a's and indeterminates x's, one works with their degrees (exponents or powers) and heights (the maximum of the absolute values of the coefficients) and the idea is to come down to irreducible polynomials, i.e., polynomials that cannot be factorized over the integers.

I have emphasized the finitist character of infinite descent and I want now to contrast it with the infinitist dimension of transfinite induction.

4 Transfinite Induction

The transfinite induction principle says

Proof

$$\forall\sigma[(\forall\tau)(\tau \prec \sigma)A(\tau, x) \rightarrow A(\sigma, x)] \rightarrow \forall\sigma A(\sigma, x) \tag{4.1}$$

for σ and τ as ordinals in the ϵ_0 segment of Cantor's second number class defined by

$$\lim_{n\to\omega} \omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega}}}}}\Big\}^{n} = \epsilon_0 \tag{4.2}$$

where epsilon naught is the limit ordinal of the omega hierarchy with n tending to ω, the limit of finite natural numbers n. Infinite descent in set-theoretic terms with a universal quantifier on natural numbers boils down to

$$\forall x(Ax \to \exists y(y \prec x)Ay) \to \forall x \neg Ax\,. \tag{4.3}$$

By successive transformations (classical equivalences and tautologies), I get

$$\forall x \neg Ax \vee \exists y Ay \tag{4.4}$$

and

$$(\neg \forall x \neg Ax \wedge \neg \exists y Ay) \vee \forall x \neg Ax \tag{4.5}$$

and by

$$(\neg \forall x \neg Ax) \leftrightarrow \exists x Ax \tag{4.6}$$

I get

$$\forall x \neg Ax \vee \exists x Ax, \tag{4.7}$$

which is the excluded middle. What we end up with is a derivation of the excluded middle obtained from a double negation operation on the infinite set of natural numbers. This is equivalent to the double negation elimination rule, not admissible in intuitionistic logic where the excluded middle principle is prohibited in general and specifically in nonfinite situations. Kolmogorov thought that the principles of excluded middle and double negation were involved in some forms of transfinite induction ([16, pp. 666–667]). The fact that intuitionists after Heyting (see [25]) accept complete induction on the infinite set of natural numbers

$$\forall x(\forall y(y \prec x)Ay \to Ax) \to \forall x Ax \tag{4.8}$$

leads them naturally to accept transfinite induction, which is just complete induction on ordinals up to ϵ_0. A similar exercise can be made with the smallest number principle (equivalent to infinite descent as classically interpreted, but not intuitionistically valid)

$$\exists y Ay \to \exists y(Ay \wedge \forall z(z \prec y)\neg Az) \tag{4.9}$$

By substituting $\neg Ay$ to Ay, we have

$$\neg \exists y(Ay \wedge \forall z(z \prec y \to \neg Az)) \tag{4.10}$$

and I obtain by classical equivalences and MP (Modus Ponens)

$$\forall y(\neg Ay \wedge \forall z(z \prec y \to \neg Az)) \tag{4.11}$$

and

$$\forall y(\forall z(z \prec y \to \neg Az)) \to Ay \tag{4.12}$$

and

$$\forall y \neg Ay \tag{4.13}$$

which is the consequent of infinite descent derived from Peano's induction postulate

$$\forall y Ay(A0 \to (\forall y Ay \to ASy)) \to \forall y Ay \tag{4.14}$$

and the induction rule

$$A0 \text{ and } \forall y Ay \to ASy \tag{4.15}$$

by MP. □

5 Conclusion: A Finitist Logic for Constructive Mathematics

A simple idea for the internal logic of finite arithmetic is that such a logic should be arithmetical, that is, it should represent or translate the logic of arithmetic in arithmetical terms. The idea is to interpret logical operators, expressions, and formulas in arithmetic and that arithmetic is polynomial arithmetic. Logic can be embedded in polynomial arithmetic, since it is a larger arena than integral arithmetic with all the arithmetic operations and it constitutes a field, in particular a finite field where infinite (indefinite or finite descent) can be freely enacted. The main advantage of the polynomial translation is that it is not simply an assignment of integers to logical expressions (like Gödel numbers), but a direct interpretation of logic in a purely arithmetical syntax. Such a direct translation would look like the following:

$$a \wedge b := a \cdot b \tag{5.1}$$

$$a \vee b := a + b \tag{5.2}$$

$$\neg a := \bar{a} \text{ for } 1 - a \tag{5.3}$$

$$a \rightarrow b := \bar{a} + b \tag{5.4}$$

$$\exists x := \Sigma(a_1, a_2, a_3, \ldots, a_n) \tag{5.5}$$

$$\forall x := \Pi(a_1, a_2, a_3, \ldots, a_n) \tag{5.6}$$

$$\mp x = \Pi(a_1, a_2, a_3, \ldots, a_n, \ldots) \tag{5.7}$$

Remarks The new quantifier ∓ is meant to express quantification over the unlimited sequence of natural numbers or Brouwer's infinitely proceeding sequences beyond the finite sets subjected to the classical quantifiers. A way to formalize infinite descent in Fermat–Kronecker arithmetic would be to use the unlimited or "effinite quantifier" in the following formulation:

$$\mp x\{([Ax \wedge \exists y(y < x)Ay] \rightarrow \exists y \mp z(z < y)Az\} \rightarrow \mp x \neg Ax, \tag{5.8}$$

which is the "negative" version of infinite descent. This means unlimited or indefinite quantification on an arbitrary sequence (not an infinite set) of natural numbers. The positive formulation looks like this

$$\mp x\{([Ax \wedge \exists y(y < x)Ay] \rightarrow \exists y \forall z(z < y)Az\} \rightarrow \exists z(z = 0 \vee 1 \vee n)Az\} \rightarrow \mp x Ax \tag{5.9}$$

and it means that a descent could stop at 0 or 1 or any positive integer (like 5) in order to give way to an unlimited number of solutions for diophantine equations, for example. This quantification is not bounded quantification, nor quantification in predicative arithmetic *à la* Nelson, it just reveals a formal contrast to Peano's induction postulate or complete induction on the (completed) infinite set of natural numbers.

The numerical expression "1" refers to the unlimited arithmetical universe. The arithmetical expression $\bar{a}$ for $1 - a$ stands for a *local* negation instead of the set-theoretic topological relative complement. Such a logic could be called a modular polynomial logic if we add a *Modus Ponens* in the form of

$$1 - a_0 x \equiv b_0 x \pmod{a_0 x} \tag{5.10}$$

where a_0x and b_0x are monomials (see [8] for details) and from the point of view of Gentzen's sequent calculus, the cut rule equivalent to *MP* is innocuous, since it is *modular*, that is, taken into account and then discarded in a pure (cut-free) equational calculus. Such a calculus is a calculus of polynomial content and could be considered as an internal logic for Kronecker's theory of forms in his general arithmetic.

The arithmetical logic I have sketched should be finitely decidable – the theory of finite fields is decidable. Of course, the arithmetic in question is not Peano arithmetic (or Dedekind–Peano) with its induction postulate on the denumerable set of natural numbers, but rather Fermat–Kronecker arithmetic with infinite descent substituting for infinite induction and acting on forms, that is the homogeneous polynomials with integer coefficients and indeterminates of Kronecker's general arithmetic (*allgemeine Arithmetik*).

After the arithmetization of analysis by Cauchy and Weierstrass – one can also include Dedekind and Cantor – and the arithmetization of algebra by Kronecker, my contention is that Hilbert has inaugurated the arithmetization of logic pursued in the work of Skolem, Gentzen, and Gödel and particularly active in contemporary theoretical computer science after Turing.

As far as number theory is concerned, the Gel'fond–Schneider theorem is a revealing instance of constructive proofs. Liouville's and Lindemann's proofs on the existence of trancendental numbers were not constructive, Dirichlet's analytical proof on the infinity of prime numbers in arithmetic progressions had to wait for the 1949 Selberg's constructive proof. The French logician and number-theorist Jacques Herbrand, a follower of Hilbert, even formulated a general hypothesis to the effect that theorems in nonanalytic (elementary) number theory must have a nonanalytic, i.e. constructive proof ([15]). It is an intrinsic feature of transcendental number theory that the existence of transcendental numbers is only negatively or indirectly demonstrated by reductio ad absurdum, but direct elementary proofs provide more information on the content of a theorem in the case of constructive mathematics and more so in finitist foundations. The emphasis is on what Hilbert called "*Sicherheit*" and "*Sicherung*," certainty or certification of the tools and means of the mathematician or the logician who wants to count on more information in order to rely on the concrete proof procedures at work in constructivist mathematics and in its internal logic.

References

1. Baker, A.: Transcendental Number Theory. Cambridge University Press, London (1975)
2. Bishop, E.: Foundations of Constructive Analysis. McGraw-Hill, New York (1967)
3. Bishop, E.: Mathematics as a Numerical Language. In: Intuitionism and Proof Theory, pp. 53–71. North-Holland, Amsterdam, New York (1970)
4. Dummett, M.: Elements of Intuitionism. Clarendon Press, Oxford (2000)
5. de Fermat, P.: Oeuvres, vol. 2, Gauthier-Villars, Paris (1894)
6. Gauthier, Y.:. Hilbert and the Internal Logic of Mathematics. Synthese **101**, 1–14 (1994)
7. Gauthier, Y.: Internal Logic. Foundations of Mathematics from Kronecker to Hilbert. Kluwer, Dordrecht, Boston, London (2002)

8. Gauthier, Y.: Logique arithmétique. L'arithmétisation de la logique. Presses de l'Université Québec, Laval (2010)
9. Gauthier, Y.:. Hilbert Programme and Applied Proof Theory. Log. Anal. **213**, 40–68 (2011)
10. Gauthier, Y.: Kronecker in Contemporary Mathematics. General Arithmetic as a Foundational Programme. Rep. Math. Log., 48, 37–65 (2013)
11. Gel'fond, A.O.: Sur le septieme Probleme de Hilbert. C. r. Acad. Sci. URSS Mosc. **2** 1–6 (1934)
12. Gel'fond, A.O.: Sur le septieme Probleme de Hilbert. Bull. Acad. Sci. URSS Leningr. **7**, 623–634 (1934)
13. Gel'fond, A.O., Linnik, V.: Elementary Methods in Analytic Number Theory. Rand McNally (1965)
14. Gentzen, G.: Collected Papers, ed. by E. Szabo. North-Holland, Amsterdam (1969)
15. Herbrand, J.: É crits logiques, ed. by J. van Heijenoort. PUF, Paris (1968)
16. Kolmogorov, A.N.: O Prinzipe tertium non datur. Mat. Sb. **32**, 646–667 (1925)
17. Kreisel, G.: Ordinal logics and the Characterization of Informal Concepts of Proof. Proc. 1958 Int. Congr. Math., pp. 289–299. Cambridge University Press, Cambridge (1960)
18. Kreisel, G.: What Have We Learnt From Hilbert's Second Problem? Mathematical Developments Arising From Hilbert's Problems, pp. 93–130. American Mathematical Society, Providence, Rhode Island (1976)
19. Kronecker, L.: Grundzüge einer arithmetischen Theorie der algebraischen Größen. Werke, vol. III, pp. 245–387, ed. by K. Hensel, 5 vols. Chelsea, New York (1968)
20. Kronecker, L.: Zur Theorie der Formen höherer Stufe. Werke, vol. II, pp. 419–424, ed. by K. Hensel, 5 vols. Chelsea, New York (1968)
21. Serre, J.-P.: How to use finite fields for problems concerning infinite fields. Proc. Conf. Marseille-Luminy, Cont. Math. Series, AMS (2007)
22. Mochizuki, S.: Inter-universal Teichmüller Theory IV. Log-volume Computations and Set-theoretic Foundations. Homepage of S. Mochizuki, August 2012 (2012)
23. Schneider, T.: Transzendenzuntersuchungen periodischer Funktionen. J. Reine Angew. Math. **172**, 65–69 (1934)
24. Schneider, T.: Transzendenzuntersuchungen periodischer Funktionen II. J. Reine Angew. Math. **172**, 70–74 (1934)
25. Troelstra, A.S., van Dalen, D.: Constructivism in Mathematics, vol. 1. North-Holland, Amsterdam (1988)
26. Troelstra, A.S.: Constructivism and Proof Theory. ILLC, University van Amsterdam (2003)
27. Voevodsky, V.: *Univalent Foundations Project*. A modified version of an NSF grant application. October 1, 2010.
28. Weil, A.: On the Riemann Hypothesis in Function Fields. Proc. Natl. Acad. Sci. USA **27**, 345–347 (1941)
29. Weil, A.: Number Theory. An Approach Through History. From Hammourabi to Legendre. Birkhäuser, Basel (1984)

Y. Gauthier (✉)
Department of Philosophy, University of Montreal, Montreal, Canada
e-mail: yvon.gauthier@umontreal.ca

Is Logic Universal or Hierarchical?

Ivor Grattan-Guinness

Abstract Some conceptions of logic claim that they are universal. By contrast, I assume that the applications of any logic are central to its conception, so that it has to comprise a hierarchy of its metalogics, metametalogics, ..., indefinitely extended but never capped off with some universal logic. I also advocate for the distinction between parts and moments of a multitude as key to this conception, and I query the assumption that set theory provides the most general means of handling collections of objects.

Keywords Logics · Hierarchies of metalogics · Logical pluralism · Self-reference · Multisets · Universal logic · Paraconsistent logics

Mathematics Subject Classification (2010) 03-03, 03-99, 03A05

> [...] the attempt to formulate the foundations of logic is rendered arduous by a corresponding 'logocentric' predicament. *In order to give an account of logic, we must presuppose and employ logic.*
> —Henry Sheffer [22, pp. 227–228]

1 On Logical Knowledge

Logic is a strange subject. It seems to be present, to a greater or lesser extent, in all discourses on any topic, especially concerning the development of a line of reasoning and comparisons between rival lines; yet it is elusive, hard to disentangle from the discourse itself, which usually is not logic but anything else (for example, proofs of theorems in geometry, or assessing the sweetness of wines, or studying the history of violin-making in Belgium ...). I take logic to be concerned with the properties of correct and incorrect reasoning from premises to conclusions in a discourse, whether in a formalised or a natural language. My approach may be called 'naturalistic', since I am much influenced by previous attempts to characterise logic; every suggestion that I offer below has been advocated somewhere. However, the combination that I offer might be new, especially for this reason.

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_14

I do not ask 'what is logic?', for it is subject to the usual criticisms concerning essentialism; instead I pose the much more modest question 'where is logic?', especially in discourses and in applications (for which I use the word 'settings'). The difference between the two questions is *very* considerable; in particular, the fact that the role of logic in its settings is as important as its 'pure' characteristics makes this approach quite different from others. Another basic difference is that I reject the common claim that 'logic dictates that ...' to justify some line of reasoning in a setting; for, with the important exception presented in Sect. 3, in my characterisation logic contains *none* of its settings, and so cannot dictate anything about any of them. Naturally this dependent status does not discourage or invalidate the development of logic 'purely' on its own, without appeal to a setting. However, I am struck by the paucity or superficiality of the discussion of logic as such in many logic textbooks!

Some assumptions and limitations need to be confessed before the characterisation of logic is described. Firstly, the word 'logic' refers to the subject, not to other uses such as in 'the logic of the situation' or 'logistics'. Secondly, well-formed propositions are translinguistic objects; languages have 'sentences', which are uttered as 'statements'. Thirdly, I shall consider only propositions in declarative mood, ignoring logics concerned with questions or commands (which are ignored too often). Fourthly, I grant epistemological priority to propositions over terms. Fifthly, I do not treat probability logic, often called 'inductive logic' in the older literature, for it is only one (or some) important logic among many. Sixthly, logical pluralism obtains, so that I use 'logics', 'a logic' and 'any logic', and 'logical knowledge' for the total ensemble of logics.

2 Specification

2.1 List

Although I do not try to 'capture' logic, I need to furnish a working list of topics that a logic should contain. Each topic is necessary for logichood, none is sufficient.

1) Well-formed propositions as translinguistic objects and their calculi (languages have 'sentences', uttered as 'statements').
2) Connectives, including their inter-definability.
3) Propositional functions, covering also relations and attributes, and their calculi.
4) Existential and universal quantification of individuals, propositions and propositional functions, to any order; each quantifier to reign over some specified universe of discourse.
5) Assertion or denial of the truth-values assigned to its propositions.
6) (In)valid reasoning, covering (at least) deduction, implication, entailment, inference, derivation, proof, natural deduction and logical consequence.

2.2 Comments

Many issues, especially philosophical ones, play important roles in discussions of logic. Cases include whether a logic is construed epistemologically or ontologically or conventionally or predicatively ...; whether propositions are to be understood positivistically or Platonically or ...; whether quantification is construed objectually or substitutionally: the deployment or avoidance of non-monotonic reasoning; the statuses that may be imposed upon symbols (schematic letters, quantifiable variables, constants, semiotic families of signs, and so on); on the extra-logical factors that may be associated with the specification of truth-values (for example, empirical testing). I hope that my modest question allows me to remain neutral on them.

I also hope that the various kinds of definitions needed in logics (which again are often ignored) can be formulated; not only nominal ones but also contextual definitions, definitions under hypothesis, creative definitions within an axiom system (that can help prove an otherwise unprovable theorem that contains neither the defined term nor any term dependent upon them), definition by mathematical induction and more generally by recursion, and allegedly essentialist ones.

3 Logics as Momental Theories

Logics have long been considered as concerned with forms; important examples include the classification of (in)valid syllogistic inferences in Aristotle and many later authors, and the distinction between form and matter that was emphasised by various logicians in the 19th century. ‘Schema(ta)’ and ‘logical skeleton’ have also been used; also ‘structure’ is sometimes used, although there are various theories outside logics that exhibit structures; for example, abstract algebras in mathematics. This status is often exhibited by writing a proposition in forms such as ‘all ... are ...’.

This approach is attractive, but lacks a clear means of distinguishing forms from non-forms in the first place; so I call upon a powerful philosophical distinction made by phenomenologists between independent and dependent parts (or ‘moments’) of a totality. It was applied widely to logics and mathematics especially by the philosopher Edmund Husserl from the 1890s, and is explored in detail in [24]. An example is the tail of a dog and its weight: the tail can be considered on its own, but its weight *necessarily* pertains to it as the totality. Both parts and moments may have their own parts and moments (the fur on the tail, the weight of that fur, the owner’s surprise over that weight, and so on). While parts of parts of ... of a totality can all be expressed in set theory using membership and inclusion; moments cannot, for they do not belong to the totality in the same way. However, one can speak of a set of moments for a given totality.

Here a logic is taken to be ‘momental’ relative to its setting and always subordinate to it in the way just described; talk of ‘forms’ and ‘structures’ is momental. To emulate a characterisation of philosophy made by Popper [17, p. 73], any logic is ‘deeply rooted in non-[logical] problems’, namely, its settings. This status helps to explain *why* a logic is so elusive; although it is omnipresent, it is always momentally dependent upon its settings for expression.

The distinction between parts and moments occurs also within a logic. For instance, the propositions 'P' and 'Q' are parts of their conjunction 'P & Q', but '&' is a moment of it. So are all other connectives, and also quantifiers, with their laws of combination indicated by their scopes.

4 The Central Importance of Hierarchies

Unlike a part, no 'host' logic L exists in isolation. Instead it *unavoidably* requires a metalogic $\mathrm{M}(L)$ which, among other duties, is to express inference and assertion, and to (fail to) show consistency and completeness. When it takes itself as a setting, it is the site not only for metapropositions such as the consistency of L but also, for example, for paradoxes like 'all propositions are untrue'. I use the plural 'metalogics' which allows for the fact that L can generate several: the formalist kind that David Hilbert developed is a major one, but there are others, often not formalised, such as comparisons with other logic, ways of teaching L, and its history. One reason why hierarchies cannot be avoided is that they are needed already in the list; assertion and inference belong to $\mathrm{M}(L)$. Each $\mathrm{M}(L)$ has metametalogics $\mathrm{MM}(L), \ldots$, and so on up, thereby generating a denumerable 'first-order' hierarchy

$$\mathrm{H}(L) := \langle \mathrm{M}(L), \mathrm{MM}(L), \ldots \rangle.$$

$\mathrm{H}(L)$ is a well-ordered sequence of collections of logics; while each member is a logic momental of all its predecessors. It is not itself a logic, nor is any subset. It is discussed in its own momental metatheory $\mathrm{MH}(L)$, which generates a 'second-order' hierarchy

$$\mathrm{HH}(L) := \langle \mathrm{MH}(L), \mathrm{MMH}(L), \ldots \rangle, \quad \text{and so on,}$$

in which, again, each member logic is a moment of all its predecessors. 'Meta…logic' names any logic.

As usual, the logic that forms each meta…logic in a hierarchy can, but does not have to, differ from each other. Among examples, the metalogic of bivalent logic is usually also bivalent and that of fuzzy logic is fuzzy; but the 'mathematics of the second order' of L.E.J. Brouwer's intuitionistic mathematics was bivalent [4, p. 61].

The assumed existence of the hierarchy will please supporters of Tarskian theories, Tr of truth, who will use it to lay out their solutions to paradoxes. However, it does not endorse or reject any proposed solutions to paradoxes; all may be tried, not only Tr but also types, say, or limitation of size. Each logic in a hierarchy may contain paradoxes of its own, awaiting attention of some kind.

5 No Universal Logic(s)

Some logicians advocate a 'universal' logic U, which is maximal, incapable of extension, embracing all known logic; but there cannot be such a logic, *since U cannot contain* $\mathrm{H}(U)$. The reply might be that $\mathrm{H}(U)$ is universal; but this is no defense, since H(U) has to be talked about, and this can only be done in its own momental metalogic M(U), which generates its own hierarchy HH(U) of H(U), and so on forever, but with *no* universal theory at the 'end'. Each logic has limits of various kinds (of settings, for example), but logical knowledge has no limits: the only appropriate talk of universes in logics is about expanding ones!

My doubts about universal logic resemble the difference between Georg Cantor and Bertrand Russell on transfinite arithmetic. In his theory of transfinite ordinals and cardinals Cantor found it basically a mistake to imagine that there could be a largest of either kind of number; neither has an immediate predecessor, just as no $(\aleph_0 - 1)$ or $(\omega - 1)$ can be found just before his transfinite sequences. The reason is that if one offered him any set S, then he would invoke his power-set proof-method and exhibit its power set, which has a larger cardinality or ordinality than S. It was Russell who proposed the (sic) set of all sets as a universal set, applied Cantor's power-set argument to it, and thereby found his paradox [8]. Cantor's reply to him could have been: 'serve you right'. In the same way, *whatever* logic a universal logician proposes as universal, I invoke its own hierarchy, which it cannot contain.

There is also no possibility of finding some universal category that embraces all the roles that were assigned to the list, as some Kantian and Hegelian logicians hoped to find in judgements. On the contrary, the roles include both logical and metalogical notions, so that the hierarchy is already in use.

6 Logics and/in Mathematics

A major concern for many logics, especially since the 1870s or so, is their relationships to branches of mathematics, in particular, those with a strong philosophical component and/or some generality. An interest among universal logicians is the characterisation of logics relative to mathematical theories: '*Universal logic* [...] is not new logic, but a general theory of logics, considered as mathematical structures' [3, p. vii]. However, take a logic L and consider the mathematical theories T relative to which it is considered: they have *their own* metatheories, which are bound to involve some logical notions. Mathematicians seldom explicitly list the logical needs of their theories, but in this context it is essential, and the charge of vicious circles must be examined; any item in the hierarchy H(T) of T cannot be in L. In particular, it is likely that H(T) includes some basic gear from bivalent (meta)logic, such as consistency and the law of excluded middle, and/or some tools from set theory. For example, Lutz Straßburger [25] seeks to characterise logic (and also proof) by attractive means E, which include ordering and several elements from set theory; but they appear to assume bivalence, and so cannot capture any logic (or theory of proof) that contains elements of E.

A reply might appeal to a distinction between 'working' everyday arithmetic that, among other duties, numbers the pages of a book on the foundations of arithmetic in which, without begging any questions, the author lays out on page 58 some foundational definition of cardinals from 0, including 58. One might make an analogous distinction between everyday set theory and topology and their foundational formulations, and remove them from the logic underlying T. However, none of these every-day theories in the elements of E helps the universalists: for on the contrary, the influence must go in the opposite direction, *from L to* them.

7 On Paraconsistent Logics: Gaps or Gluts?

Those logics with aspirations for universality need special attention. One influential case is paraconsistency, an example of the 'truth-glut' kind of logic, where in the context of bivalent logics some propositions, especially those that generate paradoxes, are both true and untrue. The reaction to paradoxes is accommodation: rejected is the 'explosion' rule

$$P \text{ and not } P; \quad \text{therefore } Q$$

for any propositions P and Q. An impressive range of theories has been found [2]. 'Dialethism' is the name given to the version of these theories that holds further that contradictions and paradoxes are both true and untrue, and permits P to be substituted by (P is true) and vice versa in well-formed contexts. Self-reference is handled similarly: the property 'is untrue of itself' is true of itself and also untrue of itself. While generality as such is not a prime property or claim for either logic, the scale of their potential application is extensive enough to excite quiet widespread doubts similar to those over universal logics: if a paraconsistent logic is to be really logical, then it has to be accompanied by its hierarchy, just like any other logic, including metaproposals such as paradoxes being both true and untrue. It is advocated especially by Graham Priest ([18] and elsewhere).

'Suppose, then', says the dialethist, 'that we allow sentences to take as semantic values one of the corresponding four *subsets* of the set {true, false}'; then the liar paradox quickly shows that if any proposition a is both true and untrue, then so are both not a and (a and not a) [18, pp. 4–5]. However, the assumption of truth-glut, especially in dialethic contexts, is dubious. The 'truth-gap' alternative, where some propositions are neither true nor untrue, is much more convincing: from 'it is true that this proposition is untrue' we deduce that 'it is untrue that this proposition is untrue' and also vice versa, suggesting that *neither* of these propositions take any truth-value, although each of them is interpretable in its metalogic. In a fine comparison of truth-glut and truth-gap theories Hartrey Field is strongly critical of the gluttists, especially the dialethists [6, especially Chapts. 7–8 and Part 5].

8 Sortality and Syntax

The preference for gaps over gluts benefits two important but often-overlooked aspects of the predicate calculus.

Firstly, gaphood accords with the theory of 'sortal terms', which Russell captured beautifully in a letter to Louis Couturat in 1906: 'To say, for example, that the principle of the excluded middle is not red, would be to utter a nonsense, and not a truth' [23, p. 604]. That is, the subject of this ill-suited proposition does not belong to the range of significance (or 'sort') of its predicate. The same can be said of the negation of the proposition, so that neither proposition takes a truth-value. Elaborating sortality is a challenging task, but it constitutes an important point of interaction between logics and semantics [15].

Secondly, gaps bear upon another much underrated aspect, namely, specifying ranges of significance of predicates called 'syntactic relativisation'. Here changing the range, or the universe of discourse of a logical argument, can convert an ill-sorted proposition to an untrue one.[1] For example, the pair of propositions 'all raccoons (dis)like Sundays' is ill-suited if its range of significance contains only raccoons; but if the range were extended to, say, all entities, then the propositions become 'for all x, if x is a raccoon, then x (does not hate) hates Sundays', which can be regarded as untrue.

9 On Collections, Especially Multisets

Throughout the history of logics collections have been handled. From antiquity up to the algebraic logicians such as George Boole and C.S. Peirce the usual theory was based upon some version of the part–whole distinction; the name 'mereology' is attached to some versions. I name this type of collection 'class'. From the late 19th century onwards it was challenged by Cantor's set theory, in which things belonged to sets and their membership was distinguished from the inclusion of subsets. A major difference was that in set theory an object is distinguished from its unit set, whereas in part–whole theory this distinction does not hold. Since around 1900 mathematical logicians have assumed that set theory is the only way to handle collections and they seem to have devised various axiomatic versions. However, other theories should be noted; not only mereologies but also, for instance, constructive set theories and fuzzy set theory.

Thus we have a meta-collection of theories of collections, which constitutes a body of knowledge of its own, *not* reducible to *any* logic. Indeed, as with mathematical theories, each will have to rely upon a logic and its hierarchy.

Thus I reject the claims of the universality of set theory, whether in general or embodied in laicisms, or whether related to mathematics ('set theory is the foundation of all mathematics [...]' according to Felix Hausdorff [12, p. 1], a noted despiser of logic!). However, my position does not rely on any of the considerations treated in this paper, but on pointing out a severe limitation in set theory itself: namely, that an object can be a member of a set *only once* (and analogously in mereologies, that a part is a part only once). This restriction severely limits the applicability of both kinds of theory, for it is very easy to find contexts in which multiple partship or membership is required, and no order is imposed upon the repeated occurrences: for example, repeated zeroes of polynomials and of latent roots of matrices, the repeated appearance as prime factors of many integers such as 16 and 28, and tallying the value of the coins in my purse. We cannot express such properties in terms of either theory; for example, the factors of 16 are 2, 2, 2 and 2, which forms neither a class

[1] See [21]. [13] contrasts the need for sortal terms in linguistics with their superfluity in mathematical logic; I regard them as relevant to any context involving predicates.

nor a set. It is just as easy to find cases in ordinary life or circumstances: for instance, words are multisets of letters and propositions of words.

Whatever logic is being developed with some wide-ranging generality or omnipresence intended, multisets *must* be used, that is, collections to which a member or part can pertain *more than once*. It is a total mystery to me that such an obvious point is not 'universally' known and accepted long ago; but multisets are still almost unknown among mathematicians and logicians [11].[2] Computer scientists are rather better served, due to the advocacy by Donald Knuth ([14] and elsewhere).

10 On Relevance Logics

One of Fields's criticisms of dialethism was Priest's rogue conditional [6, pp. 371–375]. This is unfortunate, because implication provides a link across to relevance logics, where the antecedent and consequent must be related in some way, such as having a proposition or a propositional function in common; for if in addition the antecedent is contradictory, then explosion cannot occur. The range of these logics provides a framework capable of formulating not only bivalent but also intuitionistic and modal logics[3] (always accompanied by hierarchies, of course); perhaps sufficiently large to accommodate a universal logic. They are paraconsistent logics, for the existence of a contradiction will not cause explosion. This follows from the fact that a conditional with a contradictory antecedent that does not share any propositional or predicate letters with the consequent cannot be true (or derivable).

Another aspect of relevance logic is its relationship to *logic diagrams*, which Dejnocka argues is much closer than has been realised [5]. They have a substantial history, and in recent years have gained a remarkably wide range of roles (see, for example, [1]). Some diagrams are partly representational (such as Euler and Venn diagrams), some are schematic (such as arrow diagrams). So what is the status of diagrams in logics?

Self-reference arises again, as *self-referencing self-reference*, which is so elusive that it seems to escape all hierarchies (of hierarchies of ...). Does it need a third dimension? If so, comprising what? Does it arise within some other logics that may have potential for universality?

11 On the History of Logical Hierarchies and Logical Pluralism

A.N. Whitehead and Russell were universal logicists in two respects: firstly, they regarded bivalent logic monistically as the only one; secondly, they made no effort to isolate its metalogical components (as we could construe them): [22] recognised the consequences at the heart of this article. Understandably puzzled, Ludwig Wittgenstein had already at-

[2] A good survey of applications of multisets is given in [27]. A short historical review of their presence and especially absence is provided in [11].
[3] A good formulation of relevance logics is available in [16] and [28].

tempted in his *Tractatus* to specify logic independently of the logicist thesis and imposed his own monism. Pondering the limitations of thought and logic, he denied that one could think beyond the limits; in particular, 'It is clear: the logical laws may not themselves fall again under logical laws'. He also distinguished between what can only be shown and what can be said (including this distinction, apparently!) [29, clause 6.123, see also especially clauses 4.1212, 5.6, 5.61, 5.621, 6.21, 7].

In his introduction to the essay, Russell noted this feature and proposed as an alternative position 'that every language has, as Mr. Wittgenstein says, a structure concerning which, *in the language*, nothing can be said, but that there may be another language dealing with the structure of the first language, and having itself a new structure, and that to this hierarchy of languages there may be no limit' [20, p. xxii]. He rejected the enterprise of thinking about limits and limitations; in contrast, Wittgenstein's commitment to it was to be an important motivation for forming dialethic logics in the 1980s, especially with Priest [18, especially Chapt. 12].

Wittgenstein was not the first philosopher to focus upon limitations of various kinds; in an excellent survey Priest discusses several predecessors, including Aristotle, Immanuel Kant and G.F. Hegel. In those days, the usual view of logics was monistic, in favour of some version of bivalent logic, and metalogics were not distinguished from their host logic. However, now that hierarchies are available, the doubt over a supposedly universal logic carries over to paraconsistent logics: each of them has its hierarchy, and there is no limiting case.

By contrast to Russell's bewilderment and Wittgenstein's disdain, great improvements were made by many others from the late 1920s and early 1930s, when it was recognized not only that metalogic was distinct from its host logic, but also that this distinction was of central importance. Among the leaders in this direction were David Hilbert with various papers in the 1920s, Kurt Gödel's paper [7] on incompletability (after reading it Rudolf Carnap proposed 'metalogic' as a technical term in 1931), and Alfred Tarski (and maybe some other Polish logicians as well) on 'metalanguage'.

It would be a pleasure to judge Russell as one of the father figures of this change, but this is not possible. His rebuff of Wittgenstein was one of his finest contributions to philosophy, but he *never* recognised its importance; in particular, it did not suggest to him any analogous hierarchy of logics and metalogics. Neither did his hierarchy of propositions in – that is, within – the theory of types that he used to solve the paradoxes. Thus he never understood Gödel's first (meta)theorem; not only did he always misstate it as applicable to all mathematical theories (a frequent error!) but especially he thought that his hierarchy of languages solved some 'puzzle' that it posed rather than being essential to its statement.[4] He always remained a logical monist with no hierarchy – one of the most striking ironies in the history of philosophy and logics.[5]

From the 1930s onwards metalogic, and metatheorising in general, became a central feature of logics and the foundations of mathematics. However, hardly any study seems to

[4] For discussion and evidence of this situation see [9, pp. 565, 592–593]. Russell seems to have ignored Gödel's second theorem.

[5] A similar position may be predicated of Gottlob Frege, but it differs from the absolutism imposed on him by Jean Van Heijenoort.

have been made of metametalogics, metametametalogics, . . . , metametametamathematics; search engines find very few occurrences of these words. I think this is the territory to explore; it is capable of indefinite extension and iteration, but with no 'universal' logic embracing it 'all' . . .

The reluctance of logicians to explore hierarchies contrasts strongly with the keenness for logical pluralism [10]. I share this enthusiasm, to the extent that I find the lateness of its arrival and slow progress in the early 1900s hard to understand.[6] The millennia-long history of over-looking temporal logics is especially extraordinary. Surely it is obvious, not only now but some millennia ago, that we very often make statements about sequences of actions and decisions effected *in some order in time*, with 'and' meaning 'and then'; for example, 'he opened the door and walked down the corridor', which is not logically equivalent to 'he walked down the corridor and opened the door'. However, this is not bivalent logic because conjunction is not commutative.[7] Yet temporal logics were not even among the non-bivalent logics to develop first.

12 The Glut of Foundational Notions

The main features of this conception are as follows. A logic has six offices to fulfil, involving propositions, connectives, propositional functions, quantification, deduction, assertion and truth-values; it comes with an attached hierarchy of meta. . . logics, which is momental internally and also relative to the setting in which it is used; there is no universal logic; the proposition is the prime concept; no logic includes any theory of collections (including multisets); logical pluralism abounds but hierarchies rarely feature above the meta level.

Numerous notions are now involved in the study of foundational theories in logics and mathematics. Taking in a few topics that have not been explicitly mentioned, at least the following are involved (where the grouping of topics is not necessarily defensible):

- logics: metalogics, hierarchies, form and matter, structure, parts and moments;
- logical monism and pluralism, universality, paraconsistency, generality;
- propositions, connectives, assertion, truth-values, theories of truth;
- implication, (rules of) inference, deduction, entailment, relevance, semantic tableaux;

[6] I am struck by the conformities between my advocacy of logical pluralism and hierarchies and the criticisms of the biologist Robert Rosen of the dominant status assigned to physics among the sciences, which for him forbade adequate engagement with biology and the life sciences in general. He associated physicalism with, for example, the foundationalists' preference for syntax over semantics and the advocacy of Hilbert's metamathematics. (He could have added the philosophy behind Russell's logicism, but not Whitehead's.) In his alternative approach Rosen used other devices from logic and foundations, especially relations and category theory, in order to specify minimal capacities that a material system should possess in order to be one of the simplest functional organisms normally regarded as being 'alive' [26].

[7] This difference is obtained independently of any construal of time itself (whether relative or absolute, for example, or as conceived in relativity theories). One must distinguish temporal logics from the expression of the temporal order of events in terms of ordered sets *within* bivalent logic. W.V. Quine does not do this in his treatment of tense in [19, pp. 30–31], the only attention paid to time in his book on the 'Philosophy of Logic'. He staunchly supported monism for bivalent logic while rejecting Russell's all-purpose 'implication'.

- non-monotonic reasoning, logical consequence;
- propositional functions, relations, attributes;
- quantification, relativisation of syntax, universe of discourse, (un)sortal predicates;
- self- and cross-reference, paradoxes, contradictions, fallacies, vicious circles;
- metamathematics, axiomatisation, proof theory, formalism;
- various senses of completeness, independence, consistency, constructivity;
- modality, impredicativity, temporality, vagueness;
- logicism, types, definite descriptions;
- various theories of collections, infinities, order-types, diagonalisation;
- vagueness, fuzziness;
- intension and extension;
- various senses of existence, judgments, analyticity, identity;
- model theory: categoricity, compactness, soundness;
- symbolism, well-formedness, semiotics, abstract algebras, operator algebras;
- syntax, logic diagrams, various kinds of definition;
- various kinds of foundations of mathematics.

The situation shows the merits of pluralism; but when there is *such* prolificacy, how do topics relate to each other?[8] Take a few of the principal topics: set theories, metamathematics, model theory, axiomatisation, mathematical logics, arithmetic, abstract and operator algebras and category theory. I have tried to find a comprehensible representation or summary of their relationships, but in vain; they seem to be extraordinarily complicated, with no powerful strategies for simplification.

Further, as was mentioned in Sect. 4 in connection with mathematics in logics and logics in mathematics, a serious danger lies in vicious circles, especially if a logic contains and also uses some logic. The only clear conclusion is that no universal logic will accommodate all, including itself. There does not appear to be much effort to tackle either question.

Acknowledgment For discussions or comments on the draft material I am indebted to Gregory Landini and Michael Scanlan.

References

1. Allwein, G., Barwise, J. (eds.): Logical Reasoning with Diagrams. Oxford University Press, Oxford (1996)
2. Béziau, J.-Y. et al. (eds.): Handbook of Paraconsistency. College Publications, London (2007)
3. Béziau, J.-Y. (ed.): Logica Universalis – Towards a General Theory of Logic, 2nd edn. Birkhäuser, Basel (2007)
4. Brouwer, L.E.J.: Over de Grondslagen der Wiskunde. Mass and van Suchtelen, Amsterdam (1907). Partial English translation, cited in: Collected works, vol. 1, pp. 11–101. North-Holland, Amsterdam (1975)

[8] While working on this point I received the last catalogue of books on 'Logic and Foundations' published by the American Mathematical Society of America. About 50 books were listed, from classics of the early 20th century to a section of recent works. The panorama of topics covered was beyond assessment.

5. Dejnožka, J.: The Concept of Relevance and the Logic Diagram Tradition. Dejnožka, J., Ann Arbor (2012)
6. Field, H.: Saving Truth from Paradox. Oxford University Press, Oxford (2008)
7. Gödel, K.: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. Monatsh. Math. Phys. **38**, 173–198 (1931)
8. Grattan-Guinness, I.: How Bertrand Russell discovered his paradox. Hist. Math. **5**, 127–137 (1978)
9. Grattan-Guinness, I.: The Search for Mathematical Roots, 1870–1940 – Logics, Set Theories and the Foundations of Mathematics from Cantor through Russell to Gödel. Princeton University Press, Princeton (2000)
10. Grattan-Guinness, I.: Manifestations of logical pluralism in the context of hierarchies, 1900s–1930s. In: Dutilh Novaes, C., Hjortland, O.T. (eds.) Insolubles and Consequences – Essays in Honour of Stephen Read, pp. 93–103. College Publications, London (2013)
11. Grattan-Guinness, I.: Despite Kempe, the modest place of multisets and multiclasses in foundational theories. In: Rowe, D.E., Horng, W.-S. (eds.) A Delicate Balance: Global Perspectives on Innovation and Tradition in the History of Mathematics. Birkhäuser, Basel (2014)
12. Hausdorff, F.: Grundzüge der Mengenlehre. De Gruyter, Leipzig (1914). Reprint: Chelsea, New York (1949)
13. Keenan, E.L.: Quantification in English is inherently sortal. Hist. Philos. Log. **20**, 251–265 (2000)
14. Knuth, D.E.: The Art of Computer Programming, vol. 2: Seminumerical Algorithms, 2nd edn. Addison-Wesley, Reading, MA (1981)
15. Lowe, E.J.: More Kinds of Being: a Further Study of Individuation, Identity, and the Logic of Sortal Terms, 2nd edn. Wiley-Blackwell, Chichester (2009)
16. Mares, E.: Relevance logic. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/archives/spr2014/entries/logic-relevance/ (2012)
17. Popper, K.R.: Conjectures and Refutations. Routledge and Kegan Paul, London (1963)
18. Priest, G.: Beyond the Limits of Thought. Cambridge University Press, Cambridge (1995)
19. Quine, W.V.: Philosophy of Logic.,2nd edn. Prentice Hall, Englewood Cliffs (1986)
20. Russell, B.A.W.: Introduction. In: Wittgenstein, L.: Tractatus Logico-Philosophicus, pp. 7–23. Kegan Paul, London (1922). Reprint: Collected Papers, vol. 9, pp. 96–112. Routledge, London (1988)
21. Sagüillo, J.M.: Domains of sciences, universes of discourse and omega arguments. Hist. Philos. Log. **20**, 267–290 (2000)
22. Sheffer, H.M.: Review of A.N. Whitehead and B.A.W. Russell's Principia Mathematica. Isis **8**, 226–231 (1926)
23. Schmid, A.F. (ed.): Bertrand Russell – correspondance avec Louis Couturat (1897–1913). Kimé, Paris (2001)
24. Smith, B. (ed.): Parts and Moments – Studies in Logic and Formal Ontology. Philosophia, Munich (1982)
25. Straßburger, L.: What is a logic, and what is a proof? In: Béziau, J.-Y. (ed.) Logica Universalis – Towards a General Theory of Logic, 2nd edn., pp. 135–152. Birkhäuser, Basel (2007)
26. Varenne, F.: Théorie mathématique des catégories en biologie et notion d'équivalence naturelle chez Robert Rosen. Rev. Hist. Sci. **66**, 167–197 (2013)
27. Wikipedia: Multiset. http://en.wikipedia.org/wiki/Multiset (2014)
28. Wikipedia: Relevance. http://en.wikipedia.org/wiki/Relevance_logic (2014)
29. Wittgenstein, L.: Tractatus Logico-Philosophicus. Kegan Paul, London (1922)

I. Grattan-Guinness (✉)
Middlesex University Business School, Hendon, London, UK
e-mail: eb7io6gg@waitrose.com

The English Tenses, Blanché and the Logical Kite

Dany Jaspers

Abstract There exists a very systematic limitation on natural compositional concept formation in natural language whose full complexity has not been laid bare. The first section of this chapter will revisit the lexical domains of logical constants, where conceptual "kinship and contrast" define a pattern of opposition first represented in the form of a kite in [21]. Next, I propose a kite analysis for the system of English finite tenses which has a partial precursor in a now-forgotten little article by Robert Blanché on the one hand and in the much more detailed system of binary temporal relations proposed by Vikner [23] on the other. From the former the kite structure for the finite tenses differs in that it focuses on the inviolability of prior binary divisions for later ones – which is the operation in this realm of progressive universe restriction (PUR) as defined in [21] – and on resulting asymmetries in the tense system; from the latter it differs in proposing a system of two pairs of binary oppositions, a temporal pair and an aspectual pair, yielding four binary relations in all (rather than three), with a surprising similarity between the internal architecture of the temporal and the aspectual pair of relations. On the whole, the pattern suggests that basic conceptual oppositions and lexicalization principles are guided by innate linguistic patterns of which binarity, opposition, and asymmetry are the central properties. If correct, the analysis proposed is striking confirmation in yet another semantic domain of the asymmetry approach to the square of opposition and its extensions of our earlier work.

Keywords Tenses · Square of opposition · Logical hexagon · Logical kite

Mathematics Subject Classification (2000) Primary 03B45, 03B10, 03B65, 03G05 · Secondary 05C99, 47H05

1 Introduction

> "It is as if as human beings – a particular biologically given organism – we have in our heads to start with a certain set of possible **intellectual structures**, possible sciences. Now in the lucky event that some aspect of reality happens to have the character of one of these structures in our mind, then we have a science: that is to say that, fortunately, the structure of our mind and the structure of some aspect of reality coincide sufficiently so that we develop an intelligible science. And it is because of this, it is because of this **limitation**, initial limitation in our minds to a certain

To Jean-Yves Béziau

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_15

> kind of possible science, it is precisely that that provides the tremendous richness and creativity of scientific knowledge. It is important to stress that. This has to do with your point about limitation and freedom. If it were not for these limitations, we would not have the creative act of going from a little bit of experience to a rich and highly articulated and complicated array of knowledge. (...) That doesn't mean that everything is ultimately going to fall within the domain of science. Personally I believe that many of the things we would like to understand, and maybe the things we would most like to understand, such as the nature of man, or the nature of a decent society, or lots of other things, might really fall outside the scope of possible human science."
> —Noam Chomsky, *Human Nature versus Power – Noam Chomsky debates with Michel Foucault* [boldface mine, DJ].

This chapter revisits a universal limitation on natural compositional concept formation and simplex lexicalization in natural language which grew out of a series of studies on the lexical semantics of logical operators [13, 21] and found a lot of inspiration in work by Seuren [19, 20], Horn [11], and Smessaert [22], but above all in a seminal monograph by the French philosopher Robert Blanché: *Structures Intellectuelles* [5], as well as in a little, now forgotten article by the same author [1]. It does so to extend the analysis to a new group of semantically cognate but contrasting concepts, namely those which define the English tense system and its expression in lexical predicates. To set the stage and characterize the nature of the binary concept formation constraint (CFC), Sect. 2 will review its genesis and historical antecedents and its effects in the lexical domain of logical operators.

The data set that our analysis will be built on consists of triadic sets of lexical items belonging to designated closed lexical domains:

$$\text{predicate calculus operators: } [[\textit{all some}]\ \textit{none}] \tag{1.1}$$

$$\text{propositional calculus operators: } [[\textit{and or}]\ \textit{nor}] \tag{1.2}$$

$$\text{tenses: } [[\textit{future present}]\ \textit{past}] \tag{1.3}$$

Not only is this triadic pattern surprisingly uniform – as proved for a wide range of cases in [21] and here extended to the tenses – the square brackets indicate that each triad contains two elements that stand out from the third by sharing a property that the third lacks: in the case of the operators, the bracketing sets off two affirmatives from a negative. In the case of the tenses, the future (in English and several other languages) is formed on the basis of present tense morphology, like the present itself, so that from a morphosyntactic perspective (but also for deeper conceptual reasons to be elucidated) the past is the lone third in this triad. But there is more to the square brackets uniting two elements of the triad: in each case the meaning expressed by the union of the denotations of the bracketed pair is itself lexicalised, carrying the same label as one of the bracketed elements and resulting in a lexical relation of colexemic hyponymy (or autohyponymy) between the contained and the containing colexemic item. This can be brought out by labeling the bracket in question and assigning the hyperonym the subscript 2 and the hyponym the subscript 1:

$$[[\textit{SOME}_2\ \textit{all some}_1]\ \textit{none}] \tag{1.4}$$

$$[[\textit{OR}_{2\ incl}\ \textit{and or}_1]\ \textit{nor}] \tag{1.5}$$

$$[[\textit{PRESENT}_2\ \textit{future present}_1]\ \textit{past}]. \tag{1.6}$$

In the predicate calculus, this introduces the unilateral operator $SOME_2$ *(= some maybe all)* over and above bilateral $some_1$ *(= some but not all)*; in the propositional calculus it gives a place in the system to inclusive $OR_{2\ incl}$ alongside exclusive or_1 *excl*. In the English finite tenses system, it introduces $PRESENT_2$, which occurs in sentences such as "If you work with John or Bill, I'll be delighted to learn that", where *work* may ambiguously refer to a situation that already obtains at present or whose future realization will lead to happiness. The concept formation constraint consists in the fact that while it is possible to combine the bracketed, semantically most closely related pair into a hyperonym yielding the fully natural lexicalizations just mentioned at level 2 (called level 2 and thus subscripted because they are combinations of two of the three level 1 predicates), the combination of the third element of the triad with one of the members of the natural pair is invariably nonnatural [21], 623. The level 2 bracketing that separates a pair from a third element somehow seems to preclude simplex lexicalization of any of its two hyponyms with the isolated third item. Thus, in the case of the predicate calculus operators, there exists no natural lexicalization **nall* to express the notion "$some_1$ *or none*" [8–11]; nor does there exist a natural lexicalization **allno* for the combination "*all or none*" (see [15, 16]). In the case of the propositional operators, there are no fully natural level 2 hyperonyms **nand* (= or_1 *or nor*) or **iff (= and-or-nor)*. In this chapter, we shall try to argue that in the system of the tenses, there is no simplex or periphrastic morphological lexicalization **nonpresent*$_1$ (= "*future or past*"), nor **nonfuture* (= "*present*$_1$ *or past*"), while there is a hyperonym $PRESENT_2$ as illustrated above.

Now that the data and the constraint to be fleshed out are in place, let us turn to the case originally made in [13] and [21], which forms the fertile soil for the new extension in the rest of the chapter.

2 The Concept Formation Constraint for Logical Operators

In *De Interpretatione* 6–7, 17b.17–26 and *Prior Analytics* I.2, 25a.1-25, Aristotle set up a classification of propositions in terms of two binary criteria: a Quality opposition (affirmative versus negative) and a Quantity opposition (universal versus particular) (Fig. 1).

The oppositional nature ("versus") of the two qualities and two quantities is reflected in the fact that each pair of propositions exhibits one of four possible metalogical relations of opposition (Table 1).

The first diagram in history to represent these relations is in book III of *De dogmate platonis philosophi*, attributed to Lucius Apuleius (c.124–c.170) [5, 22]. Much later the pattern was formalized by the late Roman philosopher and statesman A.M.S. Boethius (480–524) in the well-known Square of Opposition (with full lines for contradiction, arrows for subaltern entailment, a dashed line for contrariety, and a dotted line for subcontrariety, see Fig. 2).

To understand how the analysis developed in this paper relates to the Aristotelian-Boethian conception, it is important to be clear from the start about the status of three things: (1) the metalogical relations M in Table 1, (2) the propositions Q and R between which they obtain, and (3) the meanings (conceptual content) of the logical operators

	Affirmative	Negative
Universal	All citizens are proud	All citizens are not proud “No citizens are proud”
Particular	Some citizens are proud	Some citizens are not proud “Not all citizens are proud”

Fig. 1 Aristotle’s quality and quantity oppositions

“all,” “some” and “no” which are the ultimate defining factors of the relations that obtain, given that the propositions related are identical except for the operators used. To clarify our stance on these matters, we first consider the following example of a sentence in which another relation between two entities is expressed: “Mary is taller than Bill.” The relation established here clearly has no independent mind-external reality out there, it is an exclusively mind-internally assigned attribute. All there arguably is out there is the persons Mary and Bill – arguably, since we have no access to reality “an und für sich,” but we are nonetheless confident they exist. Viewing them from the perspective of their respective sizes and then relating those sizes by taking Mary’s as the first size considered and Bill’s as the size Mary’s size is compared to, and concluding that the relation involved is “taller than,” is all mind-internal activity involving mental representations. Note that for exactly the same reality out there we can switch the arguments of the proposition around and turn the relation “is taller than” into its converse “is shorter than” and get an equally valid mental perspective on exactly the same reality out there: “Bill is shorter than Mary.” With all this in mind, we return to the metalogical relations M, the propositions Q and R and

Table 1 Metalogical relations

Metalogical relations	Description	Example
Contradiction	Q and R cannot be true at the same time nor false at the same time	Q: *no citizens are proud* vs R: *some citizens are proud*
Subaltern entailment	Whenever Q is true, R is also necessarily true	Q: *all citizens are proud* R: *some citizens are proud*
Contrariety	Q and R cannot be true at the same time, but can be false at the same time	Q: *all citizens are proud* vs R: *no citizens are proud*
Subcontrariety	Q and R can be true at the same time but cannot be false at the same time	Q: *some citizens are proud* vs R: *some citizens are not proud*

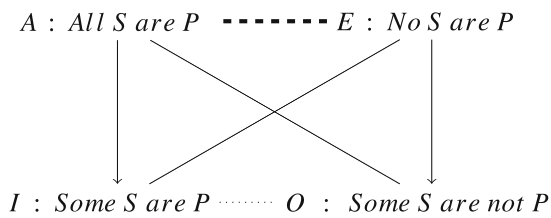


Fig. 2 Boethius' square of opposition

the logical operators. Here we posit that the relations M are once again purely derivative. Specifically, they are a function of the internal properties of the much more tangible real entities Q and R that are related. The nature of these entities is different from that of Bill and Mary in the earlier example in that they are themselves mental representations. They are real enough, but unlike Bill and Mary they are transient objects (propositions) in the mind/brain, not in the mind-external world. Moreover, they are themselves still complex entities further decomposable into less transient and more minimal mental representations called propositional functions, i.e., predicates over variables, such as ALL, SOME, or NO in ALL(S,P), SOME(S,P), and NO(S,P). It is these semantically cognate lexical predicates – more specifically the differences between their respective semantic contents – which are the reality that defines the set of metalogical relations of opposition and hence the nature of the logic involved.

Two important conclusions follow from this outlook. First, if one should decide to consciously play around with the meaning of one or more of the operators involved, one will thereby affect the nature of the relations and hence create a new logic [21]. In earlier work ([14], 558) it was argued that the choice to consciously modify the semantic content of operators is what distinguishes mathematical logic from the natural logic of ordinary language, which we claim runs on the innately specified constrained pattern that is described in [21] and further elaborated in this article rather than on consciously constructed operators. Secondly, our outlook implies that in the logic of natural language the relations of contradiction, entailment, contrariety and subcontrariety which hold between two propositional functions are themselves purely derivative: they are solely what they are in virtue of the actual existence of the mentally encoded natural meanings of the words employed. The task of the linguist is therefore to be a good lexical semanticist and decompose with formal precision the molecular meanings of the small set of cognate logical predicates in the propositions in the vertices and thereby explain why they form a closed lexical field exhibiting – in Wittgenstein's words – both "kinship and contrast."

But before embarking on that task, it is important to return to a point made at the outset and realize that from the viewpoint of natural language and more specifically its expressibility, there is something anomalous about the O-corner of the Boethian square [8, 9, 11]: there seems to be no simplex lexicalization for it. While *all*, *some* and *no* are natural lexicalizations, one has to resort to the syntactic combination *some (...) not* or *not all* for the O-corner, in the absence of a simplex form such as **nall*. This problem, perhaps originally raised for Latin predicate calculus operators by Thomas Aquinas ([9], p. 253), was to my knowledge first stated explicitly in more recent times by Blanché [4], pp. 95–96 (our translation, DJ) in the following passage:

> "The common language, for its part, has at its disposal no more than three words: *tout*, *nul*, *quelque*; it lacks for the particular concept the splitting in two that characterizes the universal. Lacuna or indifferentiation?" Is it the negative form that is lacking, or the distinction between affirmation and negation? In still other terms: are we dealing with an incomplete and irregular tetrad stripped of its O-corner or of a system that is regular and complete in its kind, but simplified and of a manifestly triadic form, in which the third element falls neither precisely in I or in O, but, so to speak, somewhere in between these two?"

Elsewhere ([3], p. 370), he generalizes the remark to modals and proposes a revision of the square to accommodate his finding:

> "Now it is important to notice – a remark which has general value, and will hold good in modal systems as well as in the quantified one – that it often suffices to consider, instead of this tetrad [A, I, E, O, (DJ)], a mere triad. Since the two subcontraries [I and O, (DJ)] may be, and in fact often happen to be, both true, we possibly need not – or perhaps cannot – distinguish between them: we do not think of either separately, but of both together. Let us designate by Y this third (or fifth) case, which results from the conjunction or logical product of I and O. (...) The third possible case is thus formed by the simultaneous rejection of the first two, i.e., by the negation of their disjunction or, what amounts to the same, by the conjunction of their negations. We thus fall again into the case designated by Y, since I and O are the negations of E and A."

This results in a pentagonal system (Fig. 3) in which "we have five positions to consider and to designate," namely the following extension of the Square, with (as anyone can check) the appropriately added dotted lines of the Hamiltonian triangle of contraries AYE, as well as entailment arrows from Y to I and O, a subscript 1 for the Y reading of *some* to distinguish it from the I-reading (subscript 2), and the corner without natural simplex lexicalization boxed.

Adding Y is a useful step in virtue of the fact that in ordinary language many propositions with *some* indeed have the reading Y:$some_1$ $(not\ all)$: it would clearly be felt to be deceptive to say *some flags are green* when one in fact knows that A:*all* are. Similarly, when we say that *not all flags are green*, the normal ordinary language interpretation is that Y:$some_1$ $(not\ all)$ are. Yet, just adding Y does not resolve the O-corner problem as long as just a single case can be found where the meanings of I and O are still attested. And they are. In *If you find some articles on Elisabeth of Spalbeek, please send them to me*, the word *some* clearly means I:$some_2$ $(or\ all)$, as the injunction includes the request to send the articles if you find all of them (the set of existing articles on Elisabeth of Spalbeek is fairly small). And in *If not all forms are filled in correctly, you will not get a visa* the applicant is not entitled to claim a visa if he has filled in *none* of the forms correctly. So the

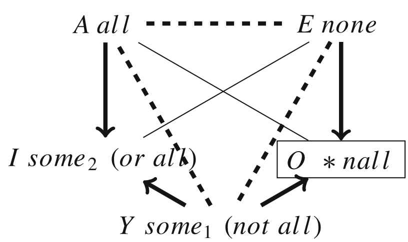


Fig. 3 Blanché's pentagon of opposition

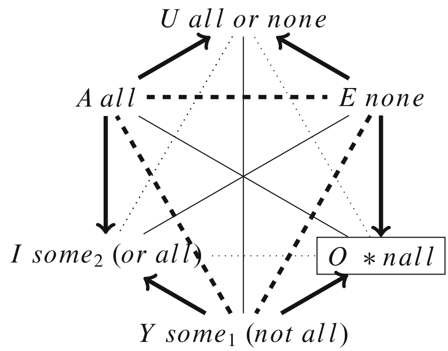


Fig. 4 The hexagon of opposition (Jacoby, Sesmat, Blanché)

interpretation of *not all* here clearly does include E:*none* and hence has a full-blown O-corner interpretation rather than Y. In sum, the mere addition of Y does not suffice to solve the O-corner problem. In hindsight, it must be said that although he had revived Aquinas' O-corner problem, Blanché did not pursue the question in sufficient depth, mainly because this linguistic asymmetry problem came second for him to a seemingly conflicting more pressing endeavour, namely to promote further extension of his logico-conceptual pentagon into a fully symmetrical bitriangular hexagon of opposition (first discovered by the American logician Jacoby [12]) by adding a final vertex U to the system. Its meaning is the contradictory of Y, and hence amounts to the disjunction of A and E (Fig. 4).

In several ways, Horn's work [8–11] marks important progress in the study of the O-corner problem. Not only did he bring the work of Blanché to the attention of the linguistic community [10], he also provided a Neo-Gricean pragmatic analysis which he (i.a.) used to address the remaining problem posed by the kinds of conditional sentences that still plagued Blanché's proposal. Moreover, he generalized the pattern further by including binary connectives, i.e., the propositional operators *and*, *or*, *nor* and **nand* in the story (see the scheme in [11], p. 404). This extension proved crucial, since it yielded a paradigm of operators which turned out more easily amenable to a decomposition analysis than the other lexical fields under study (cf. [13], chapter 4).

In Horn's Neo-Gricean account, bilateral Y:$some_1$ ($not\ all$) and unilateral I:*some*$_2$ (*or all*) are not two distinct logical operators, but the Y-case is considered a pragmatic restriction of the I-operator, from which its bilateral interpretation is derived in terms of the Gricean maxim of quantity. Not only I, but also O gets such a bilateral reading on account of that maxim, which means that its interpretation is then identical to that of bilateralized I. This identity "results in the superfluity of one of these subcontraries for lexical realization; given the functional markedness of negation (see Horn 1989 for a comprehensive review), the superfluous, unlexicalized subcontrary will always be O rather than I." [11], p. 404.

As said above, this type of pragmatic analysis can be put to use to provide a solution to the abovementioned problem of the I-reading (*some*$_2$ (*or all*)) in the protasis of certain conditionals, e.g. *If you find some articles on Elisabeth of Spalbeek, please send them to* *me*, where the unilateral lexical reading survives.[1] Given that the use of the conditional in the simple present here signals that the speaker cannot yet have certain or full knowledge about the reality of the proposition, the conditions for Gricean implicatures are not met and hence the use of *some* in this case does not trigger the implicature *but not all*.

However, although this line of reasoning can help to predict when the I-reading survives, the solution is not unproblematic: it comes at the cost of acknowledging that O and I do not always get a bilateral reading and are in those cases not equivalent. From this, Hoeksema [7] concluded: "if I and O are often not even pragmatically equivalent, because the conditions for Gricean implicatures are not met, then why should O be superfluous?"

In view of similar considerations, the view developed in [18, 20] and [13] was that the ban on O-corner lexicalization is a matter of lexical semantics, to be sought in a constraint on intralexical concept formation below word level rather than in a general pragmatic blocking constraint. Specifically, the argument in [13] was that the foundational opposition in the square and its offshoots is the E–I contradictory opposition. This fundamentum divisionis is inviolable for further concept formation in the same lexical field. This amounts to the claim that the lexical predicate in A, carved out on the affirmative side of the E–I opposition, contains I as a presupposition (thereby also lexically encoding that in natural language the subject set in A propositions is interpreted as nonnull – known as the source of the existential import problem). The consequence at the propositional level is that both A:*all flags are green* and Y:*some flags are green* contain the presupposition *some*$_2$. And consequently, (presupposition-preserving) negation of A keeps the presupposition intact in its natural reading, so that *Not all flags are green* normally does not get an O-interpretation, but an Y-interpretation. This does not, however, preclude an O-corner reading for the phrasal construct *not all* in *If not all forms are filled in correctly, you will not get a visa* provided the lexical presupposition can be blocked from projecting. Such blocking is actually quite common in conditionals. Thus, the factive verb *discover*, which normally entails the truth of its sentential object, is no longer factive in

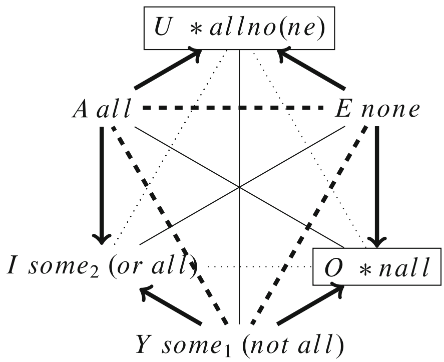
$$\text{If I discover that Bill is in New York, there will be trouble.} \qquad (2.1)$$

The picture that emerges is different from Horn's Neo-Gricean analysis: *some* is just another case of lexically polysemy – a widespread phenomenon in the lexicon anyway –, and the unilateral interpretation is restricted to contexts of (disjunctive) incomplete knowledge.

An argument in favor of this analysis is that it predicts that just as O is inaccessible for simplex lexicalization due to breaking out of the I subuniverse, U is similarly inaccessible. U-concepts would also combine a proper subpart of I, namely A, with E, which is outside I, a combination which our theory prohibits ([15], 13, and [21]). As far as we have been able to ascertain, this prediction is borne out by the facts: not only are there no natural O-corner lexicalizations, there are no natural U-lexicalizations either (Fig. 5). For this universal, which falls out automatically from our concept formation constraint, there seems to be no plausible analysis in terms of scalar implicatures.

[1] The matter of quantifier readings in the protasis is complex, as there are many other types than future (predictive) conditionals, cf. [2].

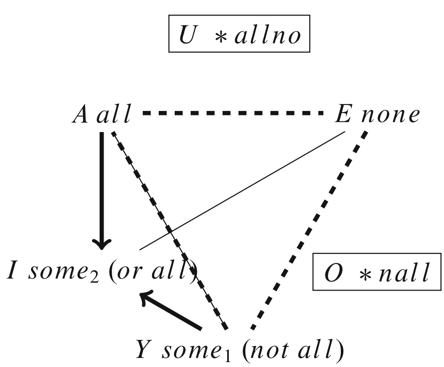
Fig. 5 The four-cornered hexagon of opposition



When the O and U vertices and the relations of opposition involving them are removed from the Hexagon, we end up with a quadrilateral (Fig. 6) in the form of a *Kite* [21], p. 621.

With respect to the latter, a caveat is in order concerning its scope. Specifically, it should be borne in mind throughout the rest of the chapter that the constraint is strictest for concept formation involving functional lexis. As the conditional sentences discussed earlier showed, the Hexagon remains in place as a logical substrate for the interpretation of certain supralexical syntactic constructs which contain more than functional lexis alone and for full utterances, in which lexical presuppositions can be overridden and O-interpretations for instance are not impossible. Furthermore, it also remains in full force with all relations intact when one decides to consciously give up the natural bounds to construct less mundane lexicalizations (such as the O-corner lexicalization *nand* widely used in the context of engineering and computing, but also the non-conventionalized forms **nall* and **allno*).

Fig. 6 The kite of opposition for the predicate calculus



3 Time and Tense

Having briefly sketched work extending the Square of Opposition and dealing with the O- and U-corner problems, the next section is devoted to the English finite tenses system, intended to pave the way for a final section which will show its relation to the logical kite.

3.1 Time in the Mind

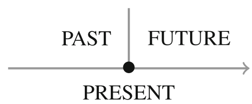
Although our cognition perceives time and space as different things, science knows better. In the natural world time is not a separate thing, as classical mechanics had it. Rather, it is a fourth coordinate which is required, along with three spatial ones, to specify an event. That four-dimensional continuum, called space–time, is a scientific construct which our mind needs considerable conscious training and learning to grasp: apparently our natural, innate "intellectual structures" are limited such that time and space are invariably experienced as separate – and for our purposes crucially bidyadic kite-structured – systems, not as one four-dimensional whole. Space itself is felt as a stable three-dimensional expanse in which we can move about, in thematic terms a stable LOCATION within which and relative to whose points of location the experiencing subject is either a non-motional THEME, c.q. a motional "running" THEME moving from a SOURCE LOCATION to a GOAL LOCATION (Gruber 1965).

$$[_{TH}\ \text{John}]\ \text{is}\ [_{LOC}\ \text{in the garden}] \tag{3.1}$$

$$[_{TH}\ \text{Mary}]\ \text{travelled}\ [_{SOURCE\ LOC}\ \text{from Spain}]\ [_{GOAL\ LOC}\ \text{to Italy}] \tag{3.2}$$

Time, for its part, has two converse forms of triadic conceptual organization. One conception is that of human beings taking steps from the past into the future, i.e., with human beings as themes, the past as their source location and the future as the goal location. Yet, instead of looking at the future and the past as the locations, there is a competing experience which turns the tables and conversely conceives of time as fleeting, as the "running" theme that moves (as expressed by the Dutch noun "toekomst" (means "coming-toward")) toward an immobile experiencing subject, or reaches and hence coincides with it (= present) and then passes the subject by ("past"), leaving him behind. Relative to that fleeting time, the subject is the immobile LOCATION, eternally locked in the now-moment of experience and incapable of time-travel. Apparently, our cognition is unable by its triadic design to experience things other than from these two angles. Clearly, our common sense cognitive experience of time and space does not change because physics tells us that our perception is not scientifically accurate. Just as we still say that the sun rises and sets and can accurately calculate the times at which this happens even though we know that it is really the earth which moves, our cognition unfailingly thinks of space as a stable unlimited three-dimensional expanse with three coordinates (= axes).

Interestingly, the description above of experiential time indicates that our mind experiences time just as three-dimensionally as space: past, present and future. Specifically, it constructs time as a continuum which – since it is intangible – is often abstractly represented as a unidirectional line on which each temporal point is in one of three possible relations relative to the fixed here-and-now point: before now (future), coinciding with now (present), behind now (past). These three form a triad of incompatible contrary predicates. The segment to the right of the now-dot is the future time sphere (the arrow indicates that the future is what the present moves into, transforming it former self into a past in the process), the segment to its left is the past time sphere:



This is a representation of our perception and conception of time, i.e., of time as we experience it. A curious property of this depiction is that we use objects in space (a line and a point) to capture something which we so clearly experience as different from space. It is Gruber (1965) who first demonstrated in great detail that our mind has adapted "insofar as possible, the independently motivated algebra of spatial concepts to (...) new purposes" (Jackendoff 1983:188) including time. Linguistically, this is evidenced by the fact that prepositions of time are on the whole identical to spatial expressions:

at the corner, at 6:00 o'clock (3.3)

from Lubbock to Amarillo, from Monday to Friday (3.4)

Actually, it is probably due to the concepts our mind uses that we have the intuitive feeling that time (which is felt to be different from space as it is felt to be fleeting) is somehow harder to come to terms with than space. In fact, we feel we can travel in space, visit the same place repeatedly, etc., but cannot really revive the past or pay an actual visit to the future. Maybe our illusion that we should somehow be able to travel in time (given that we can at least conceive of time travel) is caused by the fact that "traveling" is a spatial notion, and spatial concepts – though very useful – are not necessarily in all respects fully appropriate in non-spatial semantic fields. In any case, time in everyday cognition, which we call conceptual time, is experienced as being different from time in natural science.

3.2 Time in Speech: The Situation of Discourse and the Tense System

From natural time and experiential time, we now move to the expression of conceptual time in language. The prototypical counterpart of conceptual time in natural languages is the tense system.[2] Although that system is invariably predicated on experiential time, some additional apparatus is required. More specifically, perceptual time enters into the description of a situation of discourse, but the latter is arguably more asymmetric due to

[2] The term prototypical is used to remind us (i) that time is often indicated by means of other words than tensed verbs ("yesterday," "at five," "week," etc.) and (ii) that tenses can be used for non-temporal purposes ("Could you do me a favor?": use of the past tense to make the utterance sound more polite).

Fig. 7 Diagram of the past and the nonpast tense

a binarity requirement. First of all, the ever-shifting present moment is more narrowly defined as the Moment of (de)coding (M), which usually is the moment of speaking (but can also be the moment of decoding of a written message). From this vantage point we have a past-oriented binary choice to make: [+*retroject*] (past) or [−*retroject*] (nonpast). In the former case we jump away from M to a Retrojected Temporal Reference Point (R) preceding the moment of speech M, resulting in reference to the past time-sphere (R < M), in the latter we choose [-*retroject*], landing R at a point simultaneous with M (R = M), i.e., in the nonpast time sphere (Fig. 7).

Note that at this point there is no specific future time reference yet, which accommodates the fact that finite tenses in English are morphologically invariably either past or present. Indeed, linguistic means do not give immediate simplex morphological "flesh" to the future sphere of the threefold semantic opposition (past, present, future) of our world of temporal experience. There is in English no specific bound morpheme to mark the future form of the verb. Of course, this does not mean that English has no way of referring to future time. But the prototypical construction referring to future time and sometimes called a compound future tense is a combination of present tense with the auxiliary verb "will/shall + Vn," as in

One day, I shall/will move into my new house (3.5)

The crucial point supporting the initial past–nonpast time sphere divide is then that even the future tense is morphologically a present tense: it uses the present tense of the modal *shall/will* to obtain its semantic future time reference. This is why we are justified in postulating the drawings for the binary past–nonpast tense opposition in Fig. 7 above as the most primitive and for further divisions inviolable binary basis for the English finite tense system. Whatever choice is made, it leads to universe restriction (PUR) for further binary subdivisions. Vikner ([23], pp. 84), who proposed the same initial binarity constraint, formulates the binary choice and its morphosyntax as follows:

The first element of a past tense forms ends in the morpheme *-ed* (3.6)

The first element of a non-past tense form ends in the morpheme *-s* or Ø (3.7)

To obtain a real future and a present-only tense a further specification is required that will yield within the nonpast time sphere the opposition between the actual present time reference to M, and the actual future tense. These effects are realised by turning around from a past-orientation at M to a future-orientation at R and using R as the vantage point for a second, this time future-oriented but once again binary choice: [+*project*] (simple future) or [−*project*] (simple present). In other words, a new arrow either projects away

Fig. 8 Diagram of the future and the present tense

from the temporal reference point R (the retrojection point temporally coinciding with M) to a projection point P beyond R and M (P > R = M), into the future, or – in case of [−*project*] – it refrains from doing so, resulting in P = R = M.

Vikner [23, p. 84] describes this second binarity in the following morphosyntactic terms, illustrating that the second, future-oriented binarity turns on whether use is made of *will/shall* or not:

The first element of a future tense form is a form of *will/shall* (3.8)

This is not the case with a non-future tense form (3.9)

Since we defined R as the vantage point for the second, future-oriented binary choice between [+*project*] and [−*project*], this process can apply no less to a past R than to the nonpast one, which yields the configuration for the conditional and the definite past tense (as in *John left yesterday*), respectively (Fig. 9).

The dashed arrow in the diagram of the conditional represents the fact that this is the only tense where the position of P relative to M is indeterminate, a consequence of the fact that retroject and project both apply so that this is the only of the four tenses reached at this point (past, present, future, and conditional) which has temporal reference steps jumping in opposite directions. It seems probable that the fact that of these four tenses the conditional is also experienced as the odd one out is a consequence of this indeterminacy. In the three other tenses, there is never doubt about the position of P relative to M.

Last night at 7, John said he would do the dishes. (3.10)

. . . and he did at 9 (3.11)

. . . and he is doing them now (3.12)

. . . but he hasn't done them yet (3.13)

Given that the conditional tense form is itself indeterminate about the location of P, the three continuations of the first sentence above are all conceivable.

Fig. 9 Diagram of the conditional and the definite past tense

Let us now turn from the four basic tenses to aspect, starting with the perfect. Contrary to Vikner [23] and Declerck [6], I do not consider perfective (have + -ed) as a subcategory of tense and different from the progressive in that respect. Although Vikner is right that the binarity of the perfect ([+*perfect*] or [−*perfect*]) shows resemblance to the two tense binarities R and P, and that it allows reference to a situation that precedes P, I believe there is a major difference between temporal reference points like R and P and the perfect. While the former regulate reference to a temporal reference point at which an entire situation gets located, they impose no constraint on how much time gets wrapped into that point: it may be an instant, but equally as well a century or whatever in its entirety. This is why when P coincides with R and M and we get a present tense without further aspectual embellishment, the present time reference may be to a timeless present. The reference point is the entirety of the situation, which may encompass even all time that includes the present.

Water freezes at 0 degrees Centigrade. (3.14)

Two and two is four. (3.15)

What the two aspects do, then, is not retroject or project to a reference point of a whole situation of any possible length of internal temporal duration, but they restrict the time relevant for a main verb situation to a segment of the time line that includes the point P. The situation expressed by the main verb is then located at one or more points on that limited segment which ends (perfect) or begins (progressive) in P, as represented in the Fig. 10.

In other words, the perfect stretches the time zone within which the situation denoted by the main verb is located from P back into the past. The progressive on the other hand stretches the relevant time segment forward into the future from P, with the effect that at P the situation is often still in the middle of occurring and hence unfinished. This accounts for the following examples:

John has been with us since yesterday (3.16)

John has lost his key (3.17)

Mary is playing tennis now, so don't disturb her. (3.18)

Mary is playing tennis this afternoon. (3.19)

The time adverbials here indicate that some constraint is typically imposed on the time line, reducing it to a segment which precedes and leads up to P in the case of the perfect, or which starts from P and includes a limited period beyond P into the future in the case

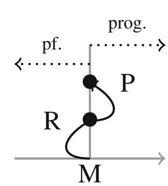


Fig. 10 Diagram of the perfect and the progressive aspects

Table 2 The 16 active finite English tenses

Name	Example	Past (R)	Fut (P)	Perf	Prog
Simple **pres**	Works	–	–	–	–
Pres prog	Is working	–	–	–	+
Pres perf	Has worked	–	–	+	–
Pres perf prog	Has been working	–	–	+	+
Simple **fut**	Will work	–	+	–	–
Fut prog	Will be working	–	+	–	+
Fut perf	Will have worked	–	+	+	–
Fut perf prog	Will have been working	–	+	+	+
Simple **past**	Worked	+	–	–	–
Past prog	Was working	+	–	–	+
Past perf	Had worked	+	–	+	–
Past perf prog	Had been working	+	–	+	+
Simple **cond**	Would work	+	+	–	–
Cond prog	Would be working	+	+	–	+
Cond perf	Would have worked	+	+	+	–
Cond perf prog	Would have been working	+	+	+	+

of the progressive. The second and fourth examples illustrat that the main verb situation located in the aspectually limited segment need not necessarily include P.

Observe that if one maintains with most traditional accounts that perfect is as much an aspect as progressive, we note a surprising parallel between the demarcation of temporal reference points R and P on the one hand and the segment demarcation by perfect and progressive respectively. Just as in the temporal reference point pair of binary relations, a past-oriented choice precedes a future-oriented choice in the aspectual pair of relations as well: first the perfect, then the progressive, in that order. The only difference between the two pairs is in the nature of the relation. The steps to R and P involve the choice to jump or not to a point, the perfect and progressive involve stretching out a period from P to a point earlier, respectively later than P, resulting in a line segment. Though R and P go in different directions, they can be combined (conditional). The same is true for the perfect and the progressive: different directions of operation, but not incompatible with one another. In sum, by keeping to the tradition, one gets two pairs of choices with the same order in the processes that apply. The only difference is in the nature of what happens. All in all, the four binary choices result in $2 \times 2 \times 2 \times 2 = 16$ active finite English tenses (Table 2).

4 Blanché's "note Sure Les Relations d'ordre" and the Tenses Kite

Given that the pair of binary tense choices R and P and the Logical Kite both involve a dyad of binary oppositions, with the second choice constrained by the result of the first, we will take that as a first indication that there is a connection between the tenses and the Kite. The required link between the two was observed in a totally forgotten little note

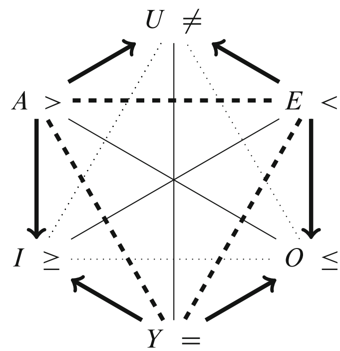


Fig. 11 The hexagon of opposition (Jacoby, Sesmat, Blanché)

from 1975 by Robert Blanché [1]. In this note, he returns to the hexagon of arithmetical relations which he discussed in par. 14 of [5] and was originally proposed by Sesmat [17].

Rather than using it to compare quantities, he now wants to apply it to serial order: from cardinalities to ordinals. He states that ordinal numbers are the prototypical case of linear order and changes the meanings of the symbols for arithmetical relations to accommodate the system: $<$ is interpreted as *precedes*, $>$ as *follows* and $=$ as *coincides with*, giving the foundational triad of contraries of linear order. He then goes on to point out that the prime concrete model of linear order is temporal succession: "before and after, anterior and posterior" (my translation, DJ), with simultaneity in between the two. Given the associations defined, *before* (and hence the past) ends up in the E-corner, *after* (and the future) in the A-corner and *simultaneity* (hence the present) in Y, exactly as needed for the tense kite if it is to do justice to the fact that the future has present tense morphology and is hence to be located in the nonpast subuniverse $PRESENT_2$, with the non-future present in Y.

$$[[PRESENT_2 \; future \; present_1] \; past] \tag{4.1}$$

This provides a model for a double Kite-structure for the tenses, within which we have made provisions for the conditional tense in a second kite that is connected to the topright E-corner of the leftmost kite (Fig. 12). It is that leftmost kite which gives expression to the arrangement of the present, future and past as proposed by Blanché. These tenses are here represented by means of the more detailed and compositional system of arrows elaborated in Sect. 3. In the A-corner of the leftmost kite, we find the simple future, with the specification [−R, +P]; in the Y-corner of the same kite resides the instantaneous simple present, with the specifications [−R, −P]. In the I-corner of the same kite, we find the kind of present ($PRESENT_2$) which occurs in sentences such as "If you work with John or Bill, I'll be delighted to learn that," where *work* can get an interpretation like that of the A-corner (future). This I-A shift [13] is typical of many other I-corner predicates: note that the interpretation of I-corner *or* in the propositional calculus, which is used in the if-sentence above also has a meaning which includes that of A-corner *and*: the meaning is "if you work with John, or Bill, or both John and Bill." In the E-corner of the leftmost kite we find a past with the single specification [+R]. This is not yet the actual definite past with the specification [+R, −P], which is in the Y-corner at the bottom of the second kite. What then is the nature of the past in the E-corner of the leftmost kite? The answer

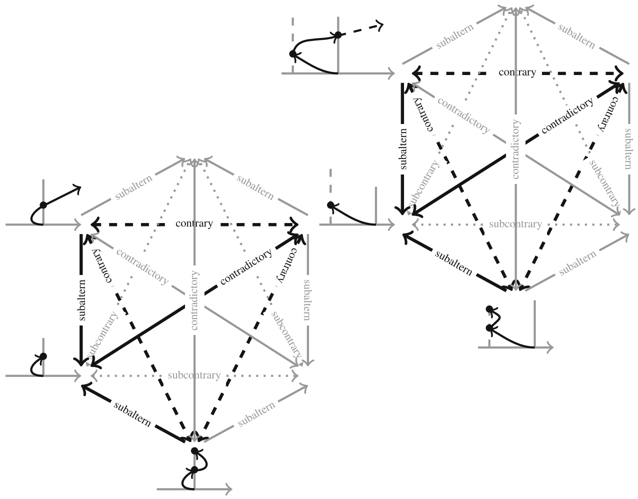


Fig. 12 Kite structure with the tense system

is straightforward once it is realized that this form is the I-corner of the second kite: it is the modal past found (i.a.) in conditionals.

$$\text{If you worked with Bill or John, I would be very happy} \tag{4.2}$$

In this case too, just as in the I-corner of the leftmost kite, the I-corner element includes the meaning of its A-corner (conditional tense), hence is interpreted as a conditional, which is supported by the fact that a possible Dutch translation of this past tense is

$$\text{Als je met Bill of John zou werken (= would work), zou ik zeer gelukkig zijn} \tag{4.3}$$

The actual definite past tense with regular past meaning, is located in the Y-corner of the rightmost kite, completing the pattern.

It will be clear from the analysis provided that the system of English tenses has no lexicalizations for the O- and U-corners of the two kites, exactly as predicted in [21]. The initial E–I opposition of the leftmost kite turns out to be equally as inviolable for tense systems as for logical operators and a wealth of other lexical fields. We take this to be strong support for both an analysis of tenses as envisaged here and the kite-system developed earlier. This does not mean there are no questions left of course. First of all, our kite structure does not yet provide for the aspectual binarities. How are they to be included, given that their structure is so parallel to that of the two temporal reference points? Such questions, which may involve 3D-extensions of the present model, are left for future work. What is provided here should be viewed as the backbone of our analysis of tenses, but since the literature on the topic is enormous and the data to be covered extremely rich, we can only hope that it is tried out on different languages.

5 Conclusion

This chapter was an exercise in extending a surprising pattern of concept formation constraints in closed lexical fields from the realm of logical operators to the triadic past–present–future structure of the English finite tense system, though it might prove to be much more generally valid given the conceptual generality – arguably universality – of the Kite pattern.

The foundations for this development were laid in [15], and [16], where a cognitive parallellism was established between opposition relations among primary (Red, Green, Blue, or RGB) and secondary (Yellow, Magenta, Cyan, or YMC) colour percepts on the one hand and oppositions between logical operators (*all*, *some-maybe-all*, *some-but-not-all*, and *none*) on the other. In that context, a constraint on lexical concept formation originally proposed in [13] was shown to affect both lexical fields alike, resulting in each case in a bifurcation between natural (RGBY and *all*, *some-maybe-all*, *some-but-not-all*, and *none* respectively) and constructed (MC and **nall*, **allno*) lexical items.

The first half of the logical variant of that constraint, namely the resistance to natural simplex lexicalization of the so-called O-corner of the Square of Opposition (the **nall*-problem) was embryonically introduced in the twentieth century by Von Wright [24], then generalized by Blanché [5], and consequently further generalized and analyzed in Neo-Gricean pragmatic terms by Horn [8–11]. The generalization involved extension of the O-corner problem to a wider number of logical calculi – such as the operators of the propositional calculus, where alongside the natural operators *or*, *nor*, *and* there is the constructed term **nand* – and beyond.

The second part of the constraint emerged in the context of an extension of the Square of Opposition into a logical Hexagon, first developed by the American philosopher Paul Jacoby [12] and consequently adopted and elaborated by [4, 10, 17] and several other authors. In the context of this Logical Hexagon, the universal ban on simplex lexicalization does not remain restricted to the O-corner, but affects another vertex, namely the so-called U-corner of the logical Hexagon [15, 16, 21]: there is no natural simplex lexicalization for an operator expressing the meaning *all-or-no*, the **allno*-problem. It is in light of this new observation that an alternative to the Neo-Gricean O-corner analysis was worked out [16, 21], which captures the universal ban on lexicalization of the two vertices in terms of a single constraint, arguably a more economical analysis than the Neo-Gricean alternative. The constraint in question consists in the claim that natural concept formation involves recursive binary opposition with progressive universe restriction (*PUR*) [21], 610). This mechanism operates below the level of standard (big) syntax. Its operation arguably exhausts the space of possible concepts available in purely functional fields such as those of the logical operators or the tenses, but does not in domains that are root [21] or semi-lexical. Although in the latter cases part of the space is still organised by the same "logically structured" substratum, another part appears to be supplied through the operation of other capacities.

Acknowledgment Thanks to Koen Roelandt of KU Leuven campus Brussels for crucial help with the tense figures. I dedicate this contribution to Jean-Yves Béziau, whose organizational skills and foresight about the relevance of Square studies for logic, philosophy and linguistics have done much to create

a vibrant Square studies community. I hope this chapter adequately testifies to my gratitude and respect. Ad multos annos!

References

1. Note sur les relations d'ordre. Rev. Philos. **4**, 431–433 (1975)
2. Abbott, B.: Part V. Back to Grice: Conditionals in English and FOPL. In: Shu, D., Turner, K. (eds.) Contrasting Meanings in Languages of the East and West. Peter Lang, Oxford (2009)
3. Blanché, R.: Quantity, modality, and other kindred systems of categories. Mind **61**(243), 369–375 (1952)
4. Blanché, R.: Sur l'opposition des concepts. Theoria **19**, 89–130 (1953)
5. Blanché, R.: Structures intellectuelles. Essai sur l'organisation systématique des concepts. Vrin, Paris (1969)
6. Declerck, R.: The Grammar of the English Tense System: A Comprehensive Analysis. Mouton de Gruyter, Berlin/New York (2006)
7. Hoeksema, J.: Blocking effects and polar sensitivity. In: JFAK. Essays dedicated to Johan van Benthem on the Occasion of his 50th Birthday. Vossiuspers/Amsterdam University Press, Amsterdam (1999)
8. Horn, L.: On the semantic properties of logical operators in English. Doctoral dissertation, University of California, Los Angeles (1972)
9. Horn, L.: A Natural History of Negation. University of Chicago Press, Chicago (1989)
10. Horn, L.: Hamburgers and truth: Why Gricean inference is Gricean. Berkeley Linguist. Soc. **16**, 454–471 (1990)
11. 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. Lang, Bern (2012)
12. Jacoby, P. A triangle of opposition in Aristotelian logic. New Scholast. **24**, 32–56 (1950)
13. Jaspers, D.: Operators in the Lexicon – On the Negative Logic of Natural Language. LOT Dissertation Series 117, Utrecht Institute of Linguistics / LOT Netherlands Graduate School of Linguistics, Utrecht (2005)
14. Jaspers, D.: The language of science and the science of language: On natural words and nonnatural terminology. In: Jaspers, D., Vermoortel, P. (eds.) Business and Language – Essays in Honour of Wilfried Janssens, pp. 567–594. Peeters Publishers, Leuven (2009)
15. Jaspers, D.: Logic of colours in historical perspective. HUB Res. Pap. **3** (2011)
16. Jaspers, D.: Logic and colour. Log. Univers. **6**(1–2), 227–248 (2012)
17. Sesmat, A.: Logique – II. Les raisonnements, la logistique. Hermann, Paris (1951)
18. Seuren, P.: The logic of thinking. Koninklijke Nederlandse Academie van Wetenschappen, Mededelingen van de afdeling Letterkunde, Nieuwe Reeks **65**(9) (2002)
19. Seuren, P.: The natural logic of language and cognition. Pragmatics **16**(1), 103–138 (2006)
20. Seuren, P.: Language From Within, vol. 2. The logic of language. Oxford University Press, Oxford (2010)
21. Seuren, P., Jaspers, D.: Logico-cognitive structure in the lexicon. Language **90**(3), 607–643 (2014)
22. Smessaert, H.: On the 3d visualisation of logical relations. Log. Univers. **3**, 303–332 (2009)
23. Vikner, S.: Reichenbach revisited: One, two or three temporal relations? Acta Linguist. Hafniensia **19**, 81–95 (1985)
24. Von Wright, G.: An Essay in Modal Logic. North-Holland, Amsterdam (1951)

D. Jaspers (✉)
KU Leuven campus Brussels, Brussels, Belgium
e-mail: dany.jaspers@kuleuven.be

Topological Aspects of Matrix Abduction 1

Azriel Laufer and Dov M. Gabbay

Abstract A new method of abduction, matrix abduction, has been introduced in Abraham, M., Gabbay, D., Schild, U.: Talmudic argumentum a fortiori inference rule (Kal Vachomer) using matrix abduction. Studia Logica **92**(3), 281–364 (2009). This method describes the Kal Vachomer and the Binyan Abh rules by using microscopic parameters which exist in the inputs of these rules. In order to find these parameters the method needs to calculate the minimal number of parameters that will describe the logical rule. In the current chapter, the matrix abduction method is formulated by Partially Orderd Sets (Posets). Consequently it is shown that the minimal number of parameters similarly defined to the dimension and k-dimension of Posets and a new poset dimension is defined which is the Kal Vachomer Dimension. In addition, several theorems and bounds of this dimension are shown.

Keywords Partially ordered sets · Poset dimension · Graph theory · Matrix abduction · Kal Vachomer

Mathematics Subject Classification Primary 06A07 · Secondary 03B48

1 Background and Orientation

The article "Analysis of the Talmudic argumentum a fortiori inference rule (Kal Vachomer) using matrix abduction" [1] suggests an analysis of the nondeductive rules, analogy (Binyan Abh) and argumentum a fortiori (Kal Vachomer) by using logical model named "matrix abduction." These two logical rules are represented by binary matrix with missing value which defined the matrix abduction problem. The way to find this missing value is based on the assumption that microscopic parameters are exist in the logical rules data.

In order to find the missing value two alternatives are compared: filling with 1 value and filling with 0 value. The nicer alternative is the correct value. In order to decide which alternative is nicer, the matrix is generates a graph and four characteristics of this graph are used as criteria. Three of these characteristics are easy to calculate: the number of points on the graph, the number of connected components (connectivity) and the number of changes of direction. The fourth characteristic is the minimal number of microscopic parameters

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_16

that describes the logical rules (named the Kal Vachomer dimension). In contrast to the first three characteristics, the calculation of this characteristic is not trivial.

In this chapter, the dimension calculation problem is mathematically defined using the partially ordered set dimension term. Several theorems are shown in order to find bounds to the dimension and to learn about its properties. Furthermore, there is a reference to the algorithm shown in [1] and a disproof is shown in this article.

2 Matrix Abduction

The *matrix abduction problem* is defined as follows [1]: Let $\mathbf{A} = [a_{i,j}]$ be 0–1 matrix, where $a_{i,j} \in \{0, 1, ?\}$, $i = 1, \ldots, m$ (m rows), $j = 1, \ldots, n$ (n columns) such that the following holds:

a. $m \leq n$.
b. Exactly one a_{i_0,j_0} is undecided, all the others are in $\{0, 1\}$.

The matrix abduction problem is to devise some algorithm which can decide whether $a_{i_0,j_0} =?$ should be 1 or $a_{i_0,j_0} =?$ should be 0 or $a_{i_0,j_0} =?$ must remain undecided.

This matrix consists of n Boolean vectors of length m, i.e., $\boldsymbol{V}_j = (a_{1,j}, a_{2,j}, \ldots, a_{m,j})$. An order is defined on two vectors $\boldsymbol{V}, \boldsymbol{V}'$ by comparing their coordinates in this way: $\boldsymbol{V} \leq \boldsymbol{V}' \Leftrightarrow$ for all i: $v_i \leq v'_i$, where $\boldsymbol{V} = (v_1, \ldots, v_m)$, $\boldsymbol{V}' = (v'_1, \ldots, v'_m)$. This order is indicated by writing $\boldsymbol{V} \rightarrow \boldsymbol{V}'$. By using the arrow one can draw a graph which describes the relation between the matrix columns (every column is represented by a vertex).

For example, given matrix **A**:

$$\mathbf{A} = \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} & \boldsymbol{c} & \boldsymbol{d} & \boldsymbol{e} & \boldsymbol{f} & \boldsymbol{g} \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The appropriate graph G for **A** is shown in Fig. 1.

The graph is generated for two matrices: one represents the matrix with $a_{i_0,j_0} = 1$ and the other represents the matrix with $a_{i_0,j_0} = 0$. The algorithm is deciding which answer is nicer according to four criteria:

1. Number of vertices (the fewer vertexes the nicer graph).
2. Connectivity (the less connected components the nicer graph).
3. Changes of direction (the fewer changes in the graph the nicer graph).
4. Kal Vachomer dimension (lower dimension is better).

This article will focus on the fourth criterion, i.e., the Kal Vachomer dimension.

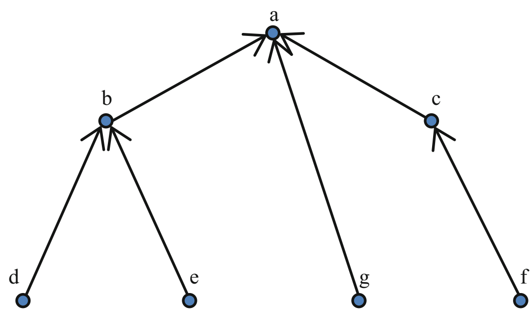


Fig. 1 Graph G that described matrix abduction **A**

3 Kal Vachomer Dimension

The graph dimension is defined by multisets in the following way: Let L be a set of labels $L = \{\alpha_1, \alpha_2, \ldots\}$. Let $\mathbb{M}(\mathbb{L})$ be the family of all multisets based on $\mathbb{L}$, so these are subsets with copies from $\mathbb{L}$. Let $\mathbb{M}_1(\mathbb{L})$ be all multisets of the form $\{m\alpha, \beta_1, \ldots, \beta_{k-1}\}$, i.e., at most one element appears with more than one copy. The elements of $\mathbb{M}_1(\mathbb{L})$ are the microscopic parameters mentioned before. In order to give a set of microscopic parameters for every vertex, the function $f : V(G) \rightarrow \mathbb{M}_1(\mathbb{L})$ is used. The function constraint is that assuming two column vectors $\boldsymbol{V}, \boldsymbol{V}' \in \mathbf{A}$ the following condition should satisfy: $\boldsymbol{V} \leq \boldsymbol{V}' \Leftrightarrow f(\boldsymbol{V}') \subseteq f(\boldsymbol{V})$. The Kal Vachomer dimension is defined as the minimal number of labels satisfying this condition. One can see an example for graph G in Fig. 2.

Following the matrix abduction and the Kal Vachomer dimension will be described as partially ordered set and partially ordered set dimension.

4 Partially Ordered Set

A *partially ordered set* (*poset*) is a pair (X, P) where X is a set and P is a reflexive, antisymmetric, and transitive binary relation on X. Elements of the set X are called *points*. The poset is *finite* if the set X is finite, this article will discuss only finite posets. When $(\boldsymbol{x}, \boldsymbol{y}) \in P$ we denote it by writing $\boldsymbol{x} \leq \boldsymbol{y}$ in P or $\boldsymbol{y} \geq \boldsymbol{x}$ in P. Let $(\boldsymbol{x}, \boldsymbol{y}) \in P$ with $\boldsymbol{x} \neq \boldsymbol{y}$, we say $\boldsymbol{x}$ and $\boldsymbol{y}$ are *comparable* in P when either $\boldsymbol{x} < \boldsymbol{y}$ in P or $\boldsymbol{y} < \boldsymbol{x}$ in P. On the other hand, $\boldsymbol{x}$ and $\boldsymbol{y}$ are *incomparable* in P if neither $\boldsymbol{x} < \boldsymbol{y}$ in P nor $\boldsymbol{x} > \boldsymbol{y}$ in P. We say $\boldsymbol{x}$ is *covered* by $\boldsymbol{y}$ in P, denoted by $\boldsymbol{x} <: \boldsymbol{y}$ in P, when $\boldsymbol{x} < \boldsymbol{y}$ in P and there is no point $\boldsymbol{z} \in X$ for which $\boldsymbol{x} < \boldsymbol{z}$ in P and $\boldsymbol{z} < \boldsymbol{y}$ in P. A poset (X, P) is called a *chain* if every distinct pair of points from X is comparable in P, when (X, P) is a chain, P is called *linear order*. Similarly, a poset (X, P) is called an *antichain* if every distinct pair of points from X is incomparable in P.

A point $\boldsymbol{x} \in X$ is called *maximal* point (respectively, *minimal* point) if there is no point $\boldsymbol{y} \in X$ with $\boldsymbol{x} < \boldsymbol{y}$ in P (respectively, $\boldsymbol{x} > \boldsymbol{y}$ in P). An element $\boldsymbol{x} \in X$ is called a *maximum* point of P if $\boldsymbol{y} \leq \boldsymbol{x}$ in P for every $\boldsymbol{y} \in X$. Similarly, $\boldsymbol{x} \in X$ is called a *minimum* point of P if $\boldsymbol{y} \geq \boldsymbol{x}$ in P for every $\boldsymbol{y} \in X$. The set of all chains in a poset

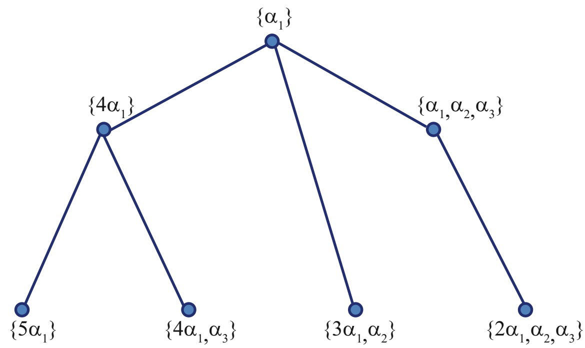


Fig. 2 Microscopic parameters for matrix **A**

(X, P) is partially ordered by set inclusion and the maximal elements in this poset are called *maximal chains*. A chain C is *maximum chain* if no other chain contains more points than C. *Maximal* and *maximum antichain*s are defined analogously. The *height* of a poset (X, P) is the number of points in a maximum chain and the *length* is one less than the height. The *width* of a poset (X, P) is the number of points in a maximum antichain.

When (X, P) and (Y, Q) are posets, a bijection $f\colon X \to Y$ so that $\boldsymbol{x}_1 \le \boldsymbol{x}_2$ in $P \Leftrightarrow f(\boldsymbol{x}_1) \le f(\boldsymbol{x}_2)$ in Q defines an *isomorphism* from (X, P) to (Y, Q). An isomorphism from (X, P) to a subposet of (Y, Q) is called an *embedding* of (X, P) in (Y, Q). The *dual* of a partial order P on a set X is denoted by P^{d} and is defined by $P^{\mathrm{d}} = \{(\boldsymbol{y}, \boldsymbol{x}) \mid (\boldsymbol{x}, \boldsymbol{y}) \in P$. The *dual* of the poset (X, P) is defined by (X, P^{d}).

A *cover* graph is associated with a poset (X, P). The edges of the cover graph consist of those pairs $(\boldsymbol{x}, \boldsymbol{y}) \in P$ for which $\boldsymbol{x} <: \boldsymbol{y}$ in P or $\boldsymbol{y} <: \boldsymbol{x}$ in P. It is often convenient to specify a finite poset by means of suitably drawn diagram of the cover graph in the Euclidean plane. A vertical coordinate system in the plane is chosen and the requirement is that the vertical coordinate of the point corresponding to $\boldsymbol{y}$ be larger than the vertical coordinate of the point corresponding to $\boldsymbol{x}$ whenever $\boldsymbol{x} <: \boldsymbol{y}$ in P. Such diagrams are called *Hasse diagrams*.

5 Representation of Matrix Abduction by Poset

As shown before the matrix abduction consists of n Boolean vectors of length m which define an order between themselves. Assuming two columns (vectors) $\boldsymbol{V}, \boldsymbol{V}' \in \mathbf{A}$ the order is defined by the vectors coordinates: $\boldsymbol{V} \le \boldsymbol{V}' \Leftrightarrow$ for all i: $v_i \le v_i'$. It is easy to see that this order is reflexive, antisymmetric, and transitive binary relation. According to the poset definition, one can notice that the matrix abduction is poset and the columns are the poset elements. We will denote it by $(\mathbb{Z}_2^m, P)$. Furthermore, the appropriate graph G for matrix $\mathbf{A}$ shown above is the Hasse diagram for $(\mathbb{Z}_2^m, P)$. The use of the Euclidean coordinate system allows us to omit the arrows from the graph.

The function $f\colon V(G) \to \mathbb{M}_1(\mathbb{L})$ used for giving microscopic parameters to the graph vertices can be defined as embedding of $(\mathbb{Z}_2^m, P)$ in $(\mathbb{N} \times \mathbb{Z}_2^d, P)$ under the constraint

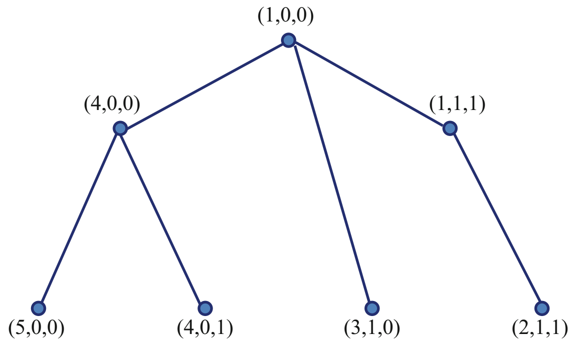


Fig. 3 Hasse diagram for matrix **A**

mentioned before ($V \leq V' \Leftrightarrow f(V') \subseteq f(V)$) One can see an example for graph G in Fig. 3.

The minimal number of parameters $(d + 1)$ that satisfies this embedding is the Kal Vachomer dimension and it will be defined by using the poset dimension term.

6 Poset Dimension

Let P and Q be order sets on the set X. Q is an *extension* of P if $P \subseteq Q$, i.e., if $\boldsymbol{x} \geq \boldsymbol{y}$ in P then $\boldsymbol{x} \geq \boldsymbol{y}$ in Q for all $\boldsymbol{x}, \boldsymbol{y} \in X$. The set of all extensions of P is partially ordered and the maximal elements of this set are linear orders on X. These elements are called *linear extensions* of P and the set of all linear extensions of P is denoted by $\varepsilon(P)$. It is easy to see that $P = \bigcap \varepsilon(P)$. The *dimension* of a poset (X, P) was defined by Dushnik and Miller [4] as the least positive integer t for which there exists a family $R = \{L_1, L_2, \ldots, L_t\}$ of linear extensions of P so that $P = \bigcap R = \bigcap_{i=1}^{t} L_i$, denoted by $\dim(X, P)$.

Ore [8] has defined alternatively the dimension of (X, P) as the least positive integer t for which (X, P) can be embedded in the Cartesian product $C_1 \times C_2 \times \cdots \times C_t$ of finite chains. There is no restriction on the sizes of the chains. Golumbic [5] has explained it by representing each element $\boldsymbol{x} \in X$ with the vector $(x_1, x_2, \ldots, x_t)$. If $\boldsymbol{x} \leq \boldsymbol{y}$ in P then the vectors are satisfying partial order in the following way: $x_i \leq y_i \Leftrightarrow (x_1, x_2, \ldots, x_t) \leq (y_1, y_2, \ldots, y_t)$ for all i when equality is not satisfying for all indexes.

Trotter [9] and Novak [7] have defined the k-*dimension* of (X, P) as the least positive integer t for which (X, P) can be embedded in the Cartesian product $C_1 \times C_2 \times \cdots \times C_t$ of finite chains, where $|C_i| = k$ for all i. It is denoted by $\dim_k(X, P)$. Some studies have focused on the case where $k = 2$ [6, 10, 11]. By this definition the Kal Vachomer dimension can be defined similarly [12].

Definition (Kal Vachomer Dimension) Assuming a poset (X, P) that was generated from matrix abduction A, the *Kal Vachomer dimension* of (X, P) is the least positive integer t for which (X, P) can be embedded in the Cartesian product $C_1 \times C_2 \times \cdots \times C_t$

of finite chains where $|C_i| = 2$ for all $i > 1$, there is no restriction on $|C_1|$. We denote it $\dim^*(X, P)$.

7 Kal Vachomer Dimension Bounds

Studies have been conducted [10] in order to find bounds to the poset dimension and to the k-dimension. It is easy to see that the relation between $\dim^*(X, P)$, $\dim_2(X, P)$ and $\dim(X, P)$ is: $\dim(X, P) \leq \dim^*(X, P) \leq \dim_2(X, P)$. This fact means that every lower bound for $\dim(X, P)$ is also a lower bound for $\dim^*(X, P)$. Similarly, every upper bound for $\dim_2(X, P)$ is also upper bound for $\dim^*(X, P)$. This section includes some theorems about the bounds and of the Kal Vachomer dimension and a few other characteristics of it. Further research should be conducted in order to find connections between these dimensions.

Assuming an antichain with n elements $A_n = (X, P)$ *then*: $\dim^*(A_n) = \lceil \log_2 n \rceil + 1$.

Theorem 7.1 (Dimension of Antichain) *Assuming an antichain with n elements* $A_n = (X, P)$ *then*: $\dim^*(A_n) = \lceil \log_2 n \rceil + 1$.

Proof Let us denote $\dim^*(A_n) = d + 1$. We will find the maximal antichain which can be embedded with $d + 1$ parameters by combinatorial considerations. The first element of the antichain will look like this: $\underbrace{(0, 1, \ldots, 1)}_{d+1}$. In the next elements, the first parameter will increase by 1 and in order to save the incomparable other parameter will be 0: $\underbrace{(1, 1, \ldots, 1, 0)}_{d+1}$. Since we can set 0 to more parameters, we can describe the next element as $\underbrace{(1, 1, \ldots, 1, 0, 1)}_{d+1}$ and so on for all the parameters. After using all the parameters, we will increase the first parameter again and two other parameters will be 0: $\underbrace{(2, 1, \ldots, 1, 0, 0)}_{d+1}$. This process can continue until the first parameter will be d and all the others are 0: $\underbrace{(d, 0, \ldots, 0)}_{d+1}$. The number of elements from this construction is n and we can calculate it in the following way:

$$n = \binom{d}{d} + \binom{d}{d-1} + \binom{d}{d-2} + \binom{d}{d-3} + \cdots + \binom{d}{1} + \binom{d}{0} = \sum_{k=0}^{d} \binom{d}{k} = 2^d.$$

The right-hand side of the equality can be proved by Newton's binomial formula. Now we will find d in order to know how much parameters we will need for maximal number of elements n:

$$d = \log_2 n,$$

i.e., with d parameters we can build an antichain with n elements at most. Generally, the minimal d for given n elements which are incomparable is $d = \lceil \log_2 n \rceil$. Hence

$$\dim^*(A_n) = \lceil \log_2 n \rceil + 1. \qquad \square$$

Theorem 7.2 (Dimension of a Chain) *Assuming a chain with n elements* $C_n = (X, P)$ *then*: $\dim^*(C_n) = 1$.

Proof Since there is no restriction on $|C_1|$ one can define a bijection $f: X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\} \to C_1 = \{1, 2, \ldots, n\}$ in the following way: $f(\boldsymbol{x}_i) = i$. This means that the least positive integer t is 1, i.e. $\dim^*(C_n) = 1$. $\square$

Theorem 7.3 *Assuming a poset* (X, P) *then* $\lceil \log_2 \text{width}(X, P) \rceil + 1 \leq \dim^*(X, P)$.

Proof Assuming the opposite $\dim^*(X, P) < \lceil \log_2 \text{width}(X, P) \rceil + 1$. Since $\dim^*(X, P)$ is a positive integer then $\dim^*(X, P) = \lceil \log_2 \text{width}(X, P) \rceil$. According to the assumption we can choose a maximal antichain with $\text{width}(X, P)$ elements and embed it with $\lceil \log_2 \text{width}(X, P) \rceil$ parameters, in contradiction to Theorem 7.1. $\square$

Theorem 7.4 *Assuming a poset* (X, P) *with n elements then* $\dim^*(X, P) \leq n$.

Proof In [9] Trotter has proved that $\dim_2(X, P) \leq n$. Since $\dim^*(X, P) \leq \dim_2(X, P)$ then $\dim^*(X, P) \leq n$. $\square$

Theorem 7.5 *Assuming a poset* (X, P) *that was generated from matrix abduction with m rows and n columns then* $\dim^*(X, P) \leq m$.

Proof The matrix abduction columns define subposet of Boolean poset, where the size of each vector is m. Let us define a bijection $f: (X, P) \to (Y, Q)$, where X is in the form $\{(x_1, x_2, \ldots, x_m)\}$ and Y is $\{(y_1, y_2, \ldots, y_m)\}$, the bijection will be $f(x_1, x_2, \ldots, x_m) = (2 - x_1, 1 - x_2, \ldots, 1 - x_m)$. The image of the bijection is embedding of the poset under the constraints mentioned before. Thus the upper bound for the Kal Vachomer dimension is m. $\square$

In order to formulate the next theorem several notations will be defined.

Vertices Level The set of vertices from *level* 1 of poset (X, P) is defined as

$$\text{level } 1 = \{\boldsymbol{x} \in X \mid \nexists \boldsymbol{y} \in X : \boldsymbol{y} > \boldsymbol{x}\}.$$

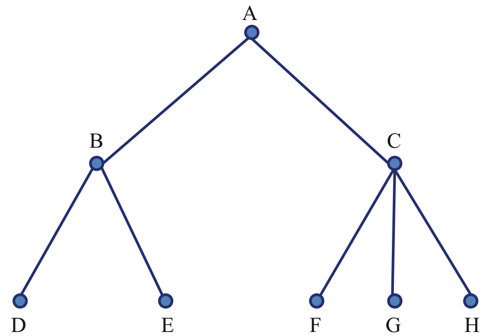
For each element from this set $\boldsymbol{x} \in$ level 1 we will denote $\text{level}(\boldsymbol{x}) = 1$.

The set of vertices from *level* $n + 1$ of poset (X, P) is defined as:

$$\text{level}(n + 1) = \{\boldsymbol{x} \in X \mid \forall \boldsymbol{y} \in P_r(\boldsymbol{x}) : \text{level}(\boldsymbol{y}) \leq n \text{ and } \exists \boldsymbol{y} \in P_r(\boldsymbol{x}) : \text{level}(\boldsymbol{y}) = n\}.$$

For each element from this set we will denote $\text{level}(\boldsymbol{x}) = n + 1$.

Fig. 4 Tree T



Antichain Sets The set of all antichains of poset (X, P) is also poset [2] and is denoted by $\mathfrak{A}(X, P)$. A subposet of $\mathfrak{A}(X, P)$ will be defined from all the antichains that have a common cover element:

$$\mathfrak{A}^{\mathrm{T}}(X, P) = \{A \in \mathfrak{A}(X, P) \mid \forall \boldsymbol{x}, \boldsymbol{y} \in A\colon \exists \boldsymbol{z} \in X\colon \boldsymbol{y} <: \boldsymbol{z} \Leftrightarrow \boldsymbol{x} <: \boldsymbol{z}\}.$$

Now, we will make a subposet of $\mathfrak{A}^{\mathrm{T}}(X, P)$, which consists of all the antichains such that the common cover element is from level i:

$$\mathfrak{A}^{\mathrm{T}}_i(X, P) = \{A \in \mathfrak{A}^{\mathrm{T}}(X, P) \mid \mathrm{level}(\boldsymbol{z}) = i, i = 1, \ldots, l(X, P)\}.$$

The maximal element of this subposet is the antichain with maximal number of elements, we will denote it by

$$\max \mathfrak{A}^{\mathrm{T}}_i(X, P) = \{A \in \mathfrak{A}^{\mathrm{T}}_i(X, P) \mid \forall B \in \mathfrak{A}^{\mathrm{T}}_i(X, P)\colon |A| \geq |B|\}.$$

One can see an example for these definitions on the tree T in Fig. 4.

We will divide the vertices to levels:

$$\mathrm{level}\,1 = \{\boldsymbol{A}\}, \quad \mathrm{level}\,2 = \{\boldsymbol{B}, \boldsymbol{C}\}, \quad \mathrm{level}\,3 = \{\boldsymbol{D}, \boldsymbol{E}, \boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}\}.$$

All the antichains with common cover element are

$$\mathfrak{A}^{\mathrm{T}}(T) = \{\{\boldsymbol{B}, \boldsymbol{C}\}, \{\boldsymbol{D}, \boldsymbol{E}\}, \{\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}\}\}.$$

The set which consists all the level 1 antichains is:

$$\mathfrak{A}^{\mathrm{T}}_1(T) = \{\{\boldsymbol{B}, \boldsymbol{C}\}\}.$$

The set which consists all the level 2 antichains is

$$\mathfrak{A}^{\mathrm{T}}_2(T) = \{\{\boldsymbol{D}, \boldsymbol{E}\}, \{\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}\}\}.$$

The maximal antichain in this set is: $\max \mathfrak{A}^{\mathrm{T}}_2(T) = \{\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}\}$.

Now we can formulate a bound for a special case of poset – a tree.

Theorem 7.6 (Dimension of a Tree) *Assuming a poset* (X, P) *whose Hasse diagram is a tree then*:

$$\dim^*(X, P) \leq \sum_{i=1}^{\text{length}(P)} \lceil \log_2 |\max \mathfrak{A}_i^{\mathrm{T}}(X, P)| \rceil + 1$$

Proof Poset (X, P) has just one element with level 1 (it is a tree), this element is the maximum vertex. In order to embed this element we will need one parameter.

- *Antichains from level* 1: The elements covered by the maximum vertex are generate antichain. Since this antichain is the only antichain that covered by the maximum vertex we can denote it by: $\max \mathfrak{A}_1^{\mathrm{T}}(T)$. According to Theorem 7.1 we can embed this antichain by $\lceil \log_2 |\max \mathfrak{A}_1^{\mathrm{T}}(T)| \rceil$ parameters besides the first parameter (from the maximum vertex) that will increase in order to contain the maximum vertex. The parameters of every element will inherit to all of its sons.
- *Antichains from level* 2: The quantity of level 2 antichains is $|\mathfrak{A}_2^{\mathrm{T}}(T)|$. For each antichain we will add other parameters (beside the parameters that inherited). The parameters number determine according to the biggest antichain. Thus, the added parameters quantity is $\lceil \log_2 |\max \mathfrak{A}_2^{\mathrm{T}}(T)| \rceil$.

We will continue in the same way until we get to the last antichains set.

- *Antichains from level* $l(T)$: We will add $\lceil \log_2 |\max \mathfrak{A}_{l(T)}^{\mathrm{T}}(T)| \rceil$ parameters analogously to the previous step.

The number of parameters that is requested in order to embed the tree is:

$$\begin{aligned} &1 + \lceil \log_2 |\max \mathfrak{A}_1^{\mathrm{T}}(T)| \rceil + \lceil \log_2 |\max \mathfrak{A}_2^{\mathrm{T}}(T)| \rceil + \cdots + \lceil \log_2 |\max \mathfrak{A}_{\mathbf{l}(T)}^{\mathrm{T}}(T)| \rceil \\ &= \sum_{i=1}^{l(T)} \lceil \log_2 |\max \mathfrak{A}_i^{\mathrm{T}}(T)| \rceil + 1. \end{aligned}$$

In other words, we have generated an embedding such that

$$\dim^*(T) \leq \sum_{i=1}^{l(T)} \lceil \log_2 |\max \mathfrak{A}_i^{\mathrm{T}}(T)| \rceil + 1. \qquad \square$$

Example (Theorem 7.6) One can see in Fig. 5 embedding that generated according to Theorem 7.6 proof.

The upper bound can be calculated according to Theorem 7.6:

$$\begin{aligned} \dim^*(T) &\leq \sum_{i=1}^{2} \lceil \log_2 |\max \mathfrak{A}_i^{\mathrm{T}}(T)| \rceil + 1 \\ &= \lceil \log_2 |\max \mathfrak{A}_1^{\mathrm{T}}(T)| \rceil + \lceil \log_2 |\max \mathfrak{A}_2^{\mathrm{T}}(T)| \rceil + 1 \\ &= \lceil \log_2 2 \rceil + \lceil \log_2 3 \rceil + 1 = 1 + 2 + 1 = 4. \end{aligned}$$

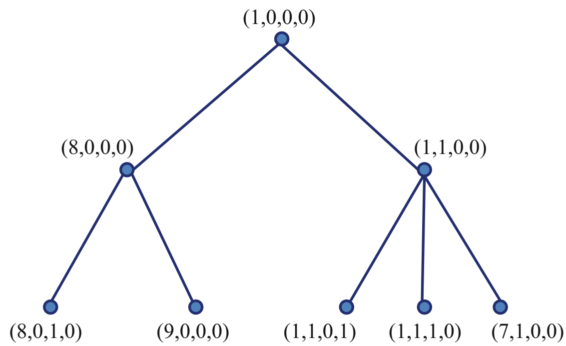


Fig. 5 Embedding of tree according to theorem 7.6 proof

The following conjecture is based on Dilworth's theorem, therefore this theorem will be shown (proof in [10]).

Dilworth's Theorem *Assuming a poset* (X, P) *with* $\text{height}(X, P) = n$, *then there exists a partition* $X = A_1 \cup A_2 \cup \cdots \cup A_n$, *where* A_i *is an antichain for* $i = 1, 2, \ldots, n$.

Conjecture *Assuming a poset* (X, P) *with* $\text{height}(X, P) = n$. *Let there be a partition into n antichains, according to Dilworth's theorem, then*:

$$\dim^*(X, P) \leq \sum_{i=1}^{n} \lceil \log_2 \text{width}(A_i) \rceil + 1.$$

Disproof The crown S_n^k is defined as a poset with $n+k$ maximal elements $a_1, a_2, \ldots, a_{n+k}$ and $n+k$ minimal elements $b_1, b_2, \ldots, b_{n+k}$, i.e. $\text{height}(S_n^k) = 2$. Each b_i is incomparable with $a_i, a_{i+1}, \ldots, a_{i+k}$ and less than the remaining $n-1$ maximal elements. In [9] Trotter has shown that $\dim_2(S_n^k) = n + k$ for every $n \geq 3, k \geq 0$. In case that $k = 0$ the crown is denoted by S_n. Let us check the S_8 graph. According to the conjecture the Kal Vachomer dimension is $\dim^*(S_8) \leq \sum_{i=1}^{2} \lceil \log_2 \text{width}(A_i) \rceil + 1 = \lceil \log_2 8 \rceil + \lceil \log_2 8 \rceil + 1 = 7$. But according to Trotter's theorem $\dim^*(S_8) = 8$ in contrast to the conjecture. □

Theorem 7.7 *Assuming a poset* (X, P) *then* $\dim^*(X, P) = \dim^*(X, P^{\mathrm{d}})$.

Proof Assuming $\dim^*(X, P) = d$. First we will see that (X, P^{d}) can be embedded with d parameters. Finally, we will show that this is the minimal embedding. Let us denote the maximal value of the vectors first parameter as $\max C_1(X, P)$ and define a bijection from the embedding of (X, P) to those of (X, P^{d}) $f\colon (\mathbb{N} \times \mathbb{Z}_2^{d-1}, P) \to (\mathbb{N} \times \mathbb{Z}_2^{d-1}, P^{\mathrm{d}})$, where the elements of the posets are in the form $(x_1, x_2, \ldots, x_d)$ and $(y_1, y_2, \ldots, y_d)$, the bijection will be $f(x_1, x_2, \ldots, x_d) = (\max C_1(X, P) - x_1, 1 - x_2, \ldots, 1 - x_d)$. We can see that $\dim^*(X, P^{\mathrm{d}}) \leq d$.

Now we will see that we cannot embed (X, P^{d}) with less that d parameters. Assuming the opposite that (X, P^{d}) can be embedded with $d - 1$ parameters. Using the bijection

f defined above one can see that (X, P) can be embedded with $d - 1$ parameters, in contradiction to the assumption that $\dim^*(X, P) = d$. Hence $\dim^*(X, P^{\mathrm{d}}) \geq d \Rightarrow \dim^*(X, P) = d = \dim^*(X, P^{\mathrm{d}})$. □

8 The Kal Vachomer Algorithm

An algorithm to minimally embed a graph under the conditions mentioned before is shown in [1]. This algorithm and comments about it will be shown in this chapter. In order to present the algorithm some definitions will be defined.

Let poset (X, P), a *projection* of maximal chain $C \subseteq X$ is the set $C^* = \{\boldsymbol{y} \mid \boldsymbol{y} \geq \boldsymbol{c}, \boldsymbol{c} \in C\}$. A maximal chain $C_1 \subseteq X$ is a *maximal thin path* if there is no other maximal chain $C_2 \subseteq X$ such that C_2^* has less elements than C_1^*.

An element $\boldsymbol{z} \in X$ is a critical point if the following holds:

a. $\boldsymbol{z}$ has at least two predecessors.
b. There exist $\boldsymbol{y} \in X$ such that $\boldsymbol{y} \not< \boldsymbol{z}$ and all predecessors of $\boldsymbol{z}$ are also predecessors of $\boldsymbol{y}$.

The level partition is done as above.

Step 0 Identify and choose one maximal thin path in (X, P), call it T. Also identify all critical points.

Step 1 Consider all maximal points (level 1 points). One of them, say $\boldsymbol{x}_1$ is in T. We will embed to this point one component vector $f(\boldsymbol{x}_1) = (1)$. If the other maximal points are $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ then n more components are added to the vector while an embed to element $\boldsymbol{y}_i$ is defined by f:

$$f(\boldsymbol{y}_i)(j+1) = \begin{cases} 0, & j \neq i, \\ 1, & j = i, \end{cases} \quad j = 1, 2, \ldots n,\ i = 1, 2 \ldots k.$$

The embedding of $\boldsymbol{x}_1$ will be updated according to n and will be: $f(\boldsymbol{x}_1) = \underbrace{(1, 0, \ldots, 0)}_{k+1 \text{ components}}$.

Step $n + 1$ Consider all level $n + 1$ points. One of them, say $\boldsymbol{e}$ is in T. Let the others be $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$. There are the following possibilities for a level $n + 1$ point $\boldsymbol{e}$.

a. $\boldsymbol{e}$ has only one predecessor and $\boldsymbol{e}$ is in T.
b. $\boldsymbol{e}$ has only one predecessor and $\boldsymbol{e}$ is not in T.
c. $\boldsymbol{e}$ has several predecessors and $\boldsymbol{e}$ is the only one with these predecessors. We have two subcases:
   1. $\boldsymbol{e}$ is a critical point.
   2. $\boldsymbol{e}$ is not critical.

d. $\boldsymbol{e}$ and $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ have the same set of predecessors and $\boldsymbol{e} \notin T$.
e. $\boldsymbol{e}$ and $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ have the same set of predecessors and $\boldsymbol{e} \in T$.

We now will embed vectors for all the elements according to different cases (the embedding of a cover element $\boldsymbol{y}$ is $f(\boldsymbol{y}) = (y_1, \dots, y_n)$).

a. Let $\boldsymbol{y}$ be the single predecessor of $\boldsymbol{e}$. Then since $\boldsymbol{e} \in T$, we also have $\boldsymbol{y} \in T$. We distinguish two subcases:

   1. $\boldsymbol{e}$ is the only element covered by $\boldsymbol{y}$. Let: $f(\boldsymbol{e}) = (y_1 + 1, \dots, y_n)$.
   2. $\boldsymbol{y}$ covers more elements apart from $\boldsymbol{e}$. Let $m(\boldsymbol{y})$ be the number of nonembedded elements in chain T. Then: $f(\boldsymbol{e}) = (y_1 + m(\boldsymbol{y}), \dots, y_n)$.

b. Let $\boldsymbol{y}$ be the single predecessor of $\boldsymbol{e}$. Since $\boldsymbol{e} \notin T$ then $\boldsymbol{y} \notin T$. We distinguish two subcases:

   1. The first component in the embedding of $\boldsymbol{y}$ is not 0, i.e., $f(\boldsymbol{y}) = (y_1 \neq 0, y_2 \dots, y_n)$. Then, $f(\boldsymbol{e}) = (y_1 + 1, \dots, y_n)$.
   2. The first component in the embedding of $\boldsymbol{y}$ is 0, i.e., $f(\boldsymbol{y}) = (0, y_2 \dots, y_n)$. Then, more component is added to the vector: $f(\boldsymbol{e}) = (y_1, \dots, y_n, 1)$.
   3. The first component in the embedding of y is not 0 and there exists an element $\boldsymbol{z}$ smaller than $\boldsymbol{y}$ ($\boldsymbol{z} < \boldsymbol{y}$ in P) and also smaller than some point in T. Then, more component is added to the vector: $f(\boldsymbol{e}) = (y_1, \dots, y_n, 1)$.

c. Let the predecessors of $\boldsymbol{e}$ be $\boldsymbol{y}_1, \dots, \boldsymbol{y}_k$. We distinguish two subcases:

   1. $\boldsymbol{e}$ is not critical. Then,

$$f(\boldsymbol{e}) = (\max(\boldsymbol{y}_i)_1, \dots, \max(\boldsymbol{y}_i)_n), \quad i = 1, \dots, k.$$

   2. $\boldsymbol{e}$ is critical. Then, more component is added to the vector:

$$f(\boldsymbol{e}) = (\max(\boldsymbol{y}_i)_1, \dots, \max(\boldsymbol{y}_i)_n, 1), \quad i = 1, \dots, k.$$

d. Let $\boldsymbol{y}, \boldsymbol{y}_1, \dots, \boldsymbol{y}_k$ be the predecessors, then the embedding for elements $\boldsymbol{e}_i$ is

$$f(\boldsymbol{e}_i) = (\max(\boldsymbol{y}_i)_1, \dots, \max(\boldsymbol{y}_i)_n, 0, \dots, \underbrace{1}_{n+1+i}, \dots 0), \quad i = 1, \dots, k.$$

   The embedding for element $\boldsymbol{e}$ is similar:

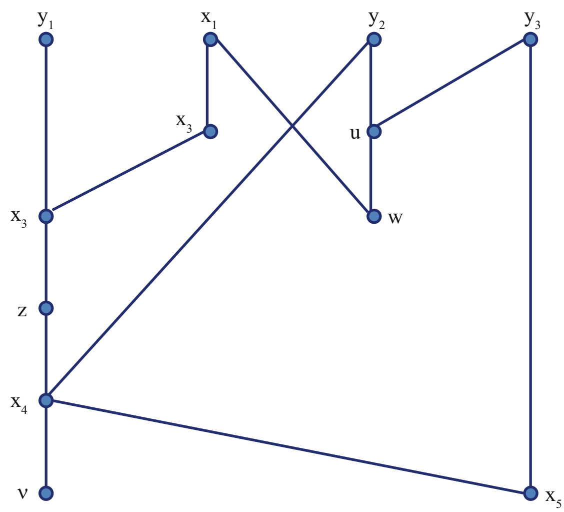
$$f(\boldsymbol{e}) = (\max(\boldsymbol{y}_i)_1, \dots, \max(\boldsymbol{y}_i)_n, \dots, 0, \dots, \underbrace{1}_{n+1}, \dots, 0).$$

e. Same as case d. besides the embedding of element $\boldsymbol{e}$:

$$f(\boldsymbol{e}) = (m(\boldsymbol{y}) + \max(\boldsymbol{y}_i)_1, \dots, \max(\boldsymbol{y}_i)_n, \dots, 0, \dots, \underbrace{1}_{n+1}, \dots, 0).$$

As a last step, we compare the number of vector components to the biggest vector by adding zeros.

Fig. 6 Graph to demonstrate the Kal Vachomer Algorithm



Example to the Algorithm An example to the algorithm is shown in [1] (diagram (X, P) in Fig. 6).

Step 0 Set of all critical points: $C = \{\boldsymbol{u}\}$.
Maximal thin path: $T = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{z}, \boldsymbol{x}_4, \boldsymbol{v}\}$.

Step 1 The set of level 1 points is: $\{\boldsymbol{x}_1, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3\}$. The size of the embedded vector is 4, as the number of level 1 points. $\boldsymbol{x}_1$ is in T and therefore the first component of its vector will be 1, the embedding to the points will be

$$f(\boldsymbol{x}_1) = (1, 0, 0, 0),$$
$$f(\boldsymbol{y}_1) = (0, 1, 0, 0),$$
$$f(\boldsymbol{y}_2) = (0, 0, 1, 0),$$
$$f(\boldsymbol{y}_3) = (0, 0, 0, 1).$$

Step 2 Level 2 points: $\{\boldsymbol{x}_2, \boldsymbol{u}\}$, $\boldsymbol{x}_2 \in T$, $\boldsymbol{u} \in C$. The embedding will be

$$f(\boldsymbol{x}_2) = (5, 0, 0, 0).$$

When we embed $\boldsymbol{u}$ more component is added to the vector

$$f(\boldsymbol{u}) = (0, 0, 1, 1, 1).$$

Step 3 Level 3 points: $\{\boldsymbol{x}_3, \boldsymbol{w}\}$. The embedding will be

$$f(\boldsymbol{x}_3) = (5, 1, 0, 0, 0),$$
$$f(\boldsymbol{w}) = (1, 0, 1, 1, 1).$$

Step 4 Only $\boldsymbol{z}$ is level 4 point. It is on T and therefore the embedding will be

$$f(\boldsymbol{z}) = (6, 1, 0, 0, 0).$$

Step 5 Only $\boldsymbol{x}_4$ is a level 5 point. The embedding will be

$$f(\boldsymbol{x}_4) = (6, 1, 1, 0, 0).$$

Step 6 Level 6 points: $\{\boldsymbol{x}_5, \boldsymbol{v}\}$. The embedding will be

$$f(\boldsymbol{x}_5) = (6, 1, 1, 1, 0),$$
$$f(\boldsymbol{v}) = (7, 1, 1, 0, 0).$$

A 0 component will be added to the first five vectors in order to complete it to 5 components. According to the algorithm one can notice that $\dim^*(X, P) = 5$.

9 A Note About the Algorithm Optimality

It can be distinguished that already in level1 the algorithm used too much components. For four points the algorithm used four components when three components are sufficient if combinations are included. For example (to the level 1 points):

$$f(\boldsymbol{x}_1) = (0, 1, 1),$$
$$f(\boldsymbol{y}_1) = (1, 1, 0),$$
$$f(\boldsymbol{y}_2) = (1, 0, 1),$$
$$f(\boldsymbol{y}_3) = (2, 0, 0).$$

This fact raises the concern that the algorithm is not optimal. According to Theorem 7.3, a lower bound can be calculated: $\dim^*(X, P) \geq \lceil \log_2 4 \rceil + 1 = 3$. Despite it is not possible to embed the diagram with three components. By brute force algorithm [3] we can find that $\dim^*(X, P) = 4$. The embedding is shown in Fig. 7.

From this example, one can see that the algorithm is not optimal. A correction to an embedding shown in [1] is written in the appendix.

10 Appendix – Correction for Graph

The following matrix abduction is shown in [1]:

$$A = \begin{array}{c} \begin{array}{ccccccc} \boldsymbol{K} & \boldsymbol{G} & \boldsymbol{H} & \boldsymbol{Y} & \boldsymbol{P} & \boldsymbol{A} & \boldsymbol{N} \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & ? & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

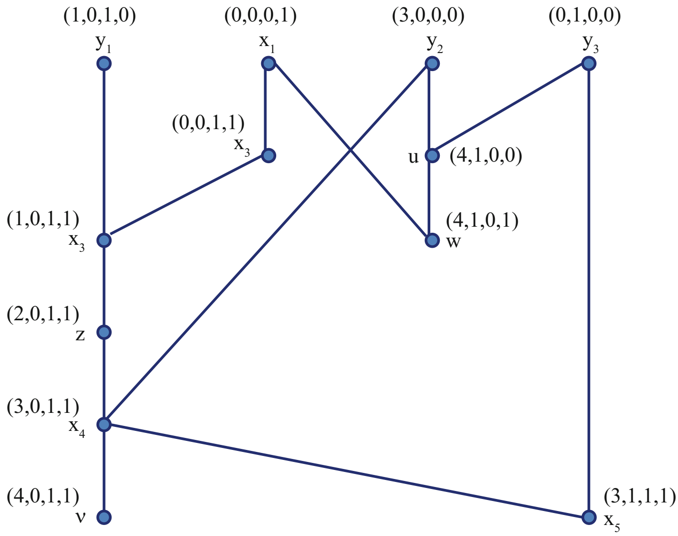


Fig. 7 Embedding with only 4 parameters

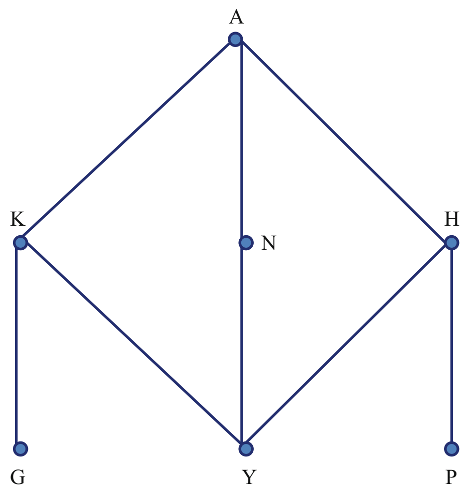


Fig. 8 Appropriate graph for matrix **A** with undecided value equal 1

The appropriate graph for this matrix such that the undecided value is 1 is shown in Fig. 8.

The minimal embedding that was calculated in [1] is written alongside the vertices (Fig. 9).

According to this solution the Kal Vachomer dimension of the graph is 4.

In contrast to this calculation one can see that three parameters are enough as shown in Fig. 10.

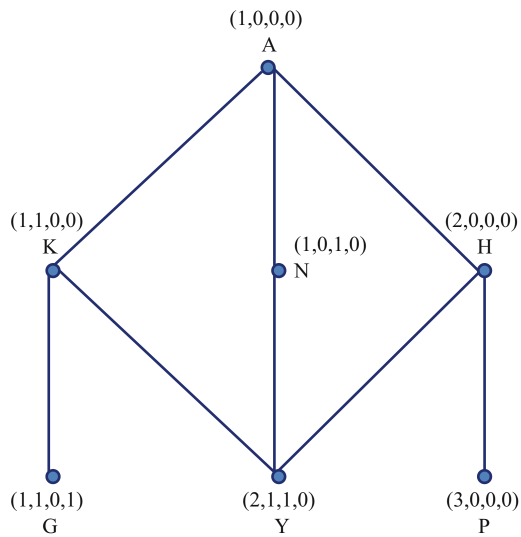


Fig. 9 Embedding matrix **A** with 4 parameters

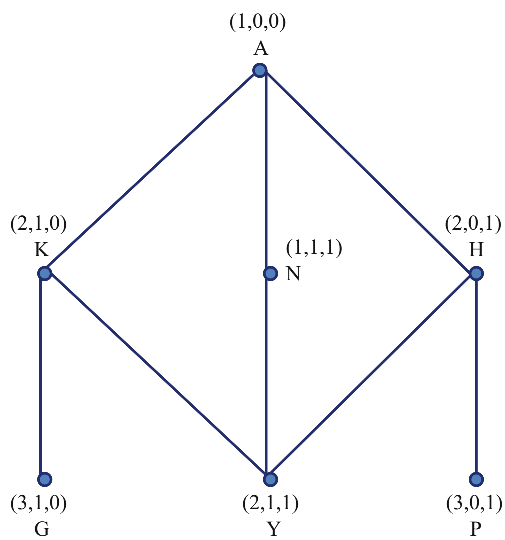


Fig. 10 Minimal embedding for matrix **A** with only 3 parameters

The width of this graph is 3. According to Theorem 7.3, the next inequality holds:

$$\dim^* \geq \lceil \log_2 3 \rceil + 1 = 3.$$

Therefore, the Kal Vachomer dimension of this graph is 3.

It can be shown similarly that the Kal Vachomer dimension of the matrix with 0 value in the undecided component is 3 and not 4.

References

1. Abraham, M., Gabbay, D., Schild, U.: Analysis of the Talmudic argumentum a fortiori inference rule (Kal Vachomer) using matrix abduction. Stud. Log. **92**(3), 281–364 (2009)
2. Anderson, I.: Combinatorics of Finite Sets. Clarendon Press, Oxford University Press, Oxford (1987)
3. Ben-Adar Bessos, M.: Improving the practical ability to perform matrix abduction. Research proposal for Master degree. Department of Computer Science, Bar Ilan University (2013)
4. Dushnik, B., Miller, E.W.: Partially ordered sets. Am. J. Math. **63**(3), 600–610 (1941)
5. Golumbic, M.C.: Algorithmic Graph Theory and Perfect Graphs. Academic Press, New York (1980)
6. Markowski, G.: The representation of posets and lattices by sets. Algebr. Univers. **11**(1), 173–192 (1980)
7. Novak, V.: On the pseudo-dimension of ordered sets. Czechoslov. Math. J. **13**, 587–598 (1963)
8. Ore, O.: Theory of Graphs. AMS Colloq. Publ., vol. 38. Am. Mathematical Society (1962)
9. Trotter, W.T.: Embedding finite posets in cubes. Discret. Math. **12**(2), 165–172 (1975)
10. Trotter, W.T.: Combinatorics and Partially Ordered Sets. The John Hopkins University Press (1992)
11. Trotter, W.T.: Graphs and partially ordered sets: recent results and new directions. Technical Report, Arizona State University (1995)
12. Trotter, W.T.: Personal communication (2012)

A. Laufer (✉)
Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel
e-mail: azriel.laufer@gmail.com

D. M. Gabbay
Department of Computer Science, King's College, London, UK
Bar-Ilan University, Ramat Gan, Israel

Topological Aspects of Matrix Abduction 2

Mai Ben Adar Bessos and Dov M. Gabbay

Abstract In the last few years, several publications discussed a new method of matrix completion, called *matrix abduction*, which is particularly useful in the context of Talmudic logic and legal ruling systems in general. Given a matrix **A** with entries that are either in $\{0, 1\}$ or blank, the method allows us to decide whether each blank entry should be 0 or 1, or remain undecided. Unlike existing matrix completion methods, which are based on notions of analogy using distance (analogy to nearest neighbors), the new method takes a different approach and completes the matrix using only topological criteria; as a result, the outcome of the process is often significantly different. This chapter will focus specifically on one of these criteria, involving the representation of the finite partially ordered set as minimally generated multisets (ordered by inclusion) for which there is no known exact polynomial-time evaluation algorithm. We examine the existing efficient greedy algorithm for this criteria, its connection the similar known concepts dimension and two dimension of a poset (partial ordered set) then derive new useful properties and an algorithm. We start by proving that the exact evaluation of the criteria is NP-complete both in the general case and in the limited context of matrix abduction. We then discuss one-point removal properties of the problematic criteria, which may aid in solving specific cases and be utilized by new algorithms. Finally, we present a new greedy algorithm that offers a significant improvement over the existing one, and discuss the possibility of reusing existing algorithms that approximate 2-dimension of a poset.

Keywords Bit-vector encodings · Partially ordered sets · Matrix abduction · Greedy algorithms · Kal Vachomer

Mathematics Subject Classification (2010) Primary 03C45

1 Introduction and Orientation

Consider n vectors $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_n$ of values $\{0, 1\}$ of length $m + 1$, $m \geq n$. Let $\boldsymbol{B}$ be another vector of length $m + 1$ which has $\{0, 1\}$ in coordinates $1, \ldots, m$ and a "?" at place $m + 1$. The problem is to determine by analogy of $\boldsymbol{B}$ with $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n$ what value to give for "?" Current methods use distance. We consider the truncated vectors $\boldsymbol{A}'_1, \ldots, \boldsymbol{A}'_n, B'$ without the $(m + 1)$th coordinate, and choose the nearest vectors, say $\boldsymbol{A}'_{i_1}, \ldots, \boldsymbol{A}'_{i_k}$ to $\boldsymbol{B}$

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_17

and see what values $\boldsymbol{A}'_{i_1}, \ldots, \boldsymbol{A}'_{i_k}$ have at their respective $(m+1)$th coordinate. Accordingly, we decide what value to give to B at the $(m+1)$th coordinate, 0 or 1 or undecided. We can arrange the vectors as columns in the matrix $\mathbf{M} = (\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n, B)$ and call the problem the "matrix abduction problem." This problem arises in modeling Talmudic reasoning. To find an appropriate answer to the question of what value to give to "?," in the Talmudic case, it was found that the distance analogy method is not adequate, but instead a new topological method needs to be used. The topological method works as follows: let $\boldsymbol{B}^0$ and $\boldsymbol{B}^1$ be the vectors obtained from $\boldsymbol{B}$ by letting $? = 0$ and $? = 1$, respectively. Define an ordering on vectors $\boldsymbol{X} = (x_1, \ldots, x_{m+1})$ by $X \leq Y$ if and only if $x_i \leq y_i$ for all $i = 1, \ldots, m+1$. This yields two graphs, G^0 obtained using $\{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n, \boldsymbol{B}^0\}$ and $G^1 = \{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n, \boldsymbol{B}^1\}$. We compare the two graphs G^0, G^1 according to four topological criteria and depending on what we find we decide whether $? = 1$, $? = 0$ or $? =$ undecided. The topological method works correctly for Talmudic logic, and does not always agree with the distance method. One of the criteria required involves the problem of representing a graph $(G, \leq)$ by a certain type of multiset, generated by a minimal number of generators $\{\alpha, \beta_1, \beta_2, \ldots\}$. We let a function f be the representation with $f(\boldsymbol{x}) = E_{\boldsymbol{x}}$, $\boldsymbol{x}, \boldsymbol{y} \in G$, $E_{\boldsymbol{x}}$ a multiset, and we require that $\boldsymbol{x} \leq \boldsymbol{y}$ if and only if $E_{\boldsymbol{x}} \subseteq E_{\boldsymbol{y}}$. There is a problem of finding for a given graph G, representations f as above in terms of the minimal number of generators. Note that this is a purely combinational problem of representing finite graphs for partially ordered sets in terms of faithful multiset inclusion of certain types. This question has been considered in the general mathematical literature and seems to have arisen in modeling Talmudic logic. This chapter contributes toward the complexity problem of this question.

We now explain by example the context in which this problem arises. Suppose we want to buy a screen, considering few preferences. Our options are:

- Xerox XM7 24A,
- Viewsonic FHD VX 2640w,
- Nec WVX,
- Nec WMCX.

Table 1 describes how each screen matches our preferences; each column describes a different property and each row describes one of the compared items. Note that a value of "1" consistently indicates support in a certain direction. In this case, a value of "1" implies the screen offers better performance. The matrix abduction method assumes that such a direction exists in the given matrix and relies on it in order to deduce the missing value in "?".

From the mathematical point of view, the abduction is a problem of analogical reasoning. Many of the existing matrix completion algorithms aim to minimize the Euclidian distance or the rank of the recovered matrix (e.g., k-nearest neighbor and convex optimization [2, 10]). In a broader sense, there were many attempts for modeling analogy and new methods are still being developed today, based on slightly different approaches [8, 16] provide a thorough review of such approaches and corresponding methods). Moreover, even though not defined as such, many classification algorithms may be used for abductive inference, since the underlying goal is similar: deriving a conclusion regarding an occurrence based on pre-established reliable data. The need for different approaches stems from

Table 1 Comparison of different screens, considering specific preferences

	(1) Price over 450 £	(2) Self-collection	(3) Screen bigger than 24 inch	(4) Dot size less than 0.275	(5) Stereophonic	(6) Reaction time below 4 ms
Xerox	1	0	1	0	0	1
Nec WMCX	0	0	0	0	1	1
Nec WVX	1	1	0	1	0	1
Viewsonic	0	0	0	1	1	?

1 – yes, 0 – no, ? – no data given

Fig. 1 Resulting Poset Comparison

the multitude of contexts and forms of data representation. The matrix abduction method, arising from modeling the use of analogy in Talmudic logic, takes an entirely new approach and relies only on topological criteria, unlike the existing methods; as a result, the outcome of the process is often significantly different compared to other algorithms, and has better applicability in different cases. In its simplest form, the method takes as input a matrix with entries in $\{0, 1\}$ and several blank entries, then under the assumption that the data of the matrix meets certain properties, the method allows us to infer whether each blank entry is 0, 1, or should be left undecided. The decision is made by assigning every possible combination of values instead of the missing entries, then transforming each complete matrix into a poset (partial ordered set) which may be reexamined and evaluated using four specific quantitative topological criteria. For each complete matrix the criteria are evaluated and compared, and if a conclusive decision may be made then the preferable assignment is chosen. For example, assigning “0” and “1” instead of the missing values, results in the posets as illustrated in Fig. 1. Intuitively, assigning “1” gives us a simpler graph, and indeed when evaluated using the four topological criteria (further explained in Sect. 2) it is deemed preferable.

In more detail, the four criteria are:

- amount of elements,
- index of connectivity i.e. the amount of connected components,
- index of directional change (maximal amount of turns when walking in the graph),
- dim* – a new concept, discussed in more details in the next section.

In this example, dim* and the amount of elements are tied, but the index of connectivity and the index of directional change are in favor of assigning "1." In contrast, the result of comparing distances (k-nearest neighbor algorithm) would be assigning "0." Section 2 provides an additional, more detailed comparison of the two algorithms, after we fully define the new method.

Three of the above criteria are easy to evaluate using efficient algorithms, but dim* (also referred to as *logical model dimension*) cannot be evaluated as easily. The primary goal of this chapter is advancing toward a practical algorithm for evaluating the result of the matrix abduction method. In Sect. 2, we establish the necessary preliminaries and define the method itself. Section 3 gives a brief review of the related work. In Sect. 4, we examine the computational complexity of the matrix abduction method and the evaluation of dim* and prove the suspicion that both problems are NP-complete. Afterward, we will present multiple dim* point – removal properties that are useful when performing an abduction. Finally, in Sect. 5, we present an improved variation of the algorithm devised in [1], which allows us to produce a better evaluation of dim* and consequently also a better evaluation of the matrix abduction. In our experiments, we had nearly 800 comparisons, and the new algorithm has produced better results than the existing one in almost every test.

2 Preliminaries

In this section, we give some of the basic definitions that will be used in the scope of this chapter. Similar definitions may also be found in [1, 6, 10, 13, 18]. Additional propositions, definitions, and annotations will be given along with the context of their use.

Let $P = (X, \leq_P)$ be a partial order on the ground set X. We only consider finite orders and denote by $|P|$ the cardinality of X. Similarly, we may apply binary set operations on X (e.g., union, intersection, difference etc.), and use P instead of X to denote the operation. The same order relation $\leq_P$ restricted to a subset Y of X is called a *suborder* of P and also referred to as the *order induced* by P on Y. Let $\boldsymbol{x}, \boldsymbol{y} \in X$, $\boldsymbol{x} \neq \boldsymbol{y}$, then we say that $\boldsymbol{x}$ and $\boldsymbol{y}$ are *comparable* in P if either $\boldsymbol{x} \leq_P \boldsymbol{y}$ or $\boldsymbol{y} \leq_P \boldsymbol{x}$. Otherwise, we say that $\boldsymbol{x}$ and $\boldsymbol{y}$ are *incomparable*. Let $B, T \subseteq X$, we denote $B \leq_P T$ if for every $\boldsymbol{b} \in B$, $\boldsymbol{t} \in T$ it holds that $\boldsymbol{b} \leq_P \boldsymbol{t}$. A partial order where every pair of elements is comparable is called a *chain*. A partial order where every pair of different elements is incomparable is called an *antichain*. By extension, for the poset $P = (X, \leq_P)$, a nonempty subset Y of X is called a *chain* (resp. *antichain*) of P if every pair of different elements of Y is comparable (resp. incomparable) in P. Let Y_1, Y_2 be antichains of P. If $Y_1 \leq_P Y_2$, we say Y_1, Y_2 are *ordered antichains* of P. The maximum cardinality of a chain of P minus 1 is called the *height* of P, and is denoted by $h(P)$. The maximum cardinality of an antichain of P is called the *width* of P and is denoted by $w(P)$. An element $\boldsymbol{x} \in X$ is called the *maximum* (resp. *minimum*) of P if for all $\boldsymbol{y} \in X$, $\boldsymbol{y} \leq_P \boldsymbol{x}$ (resp. $\boldsymbol{x} \leq_P \boldsymbol{y}$). An element $\boldsymbol{x} \in X$ is *maximal* (resp. *minimal*) of P if for all $\boldsymbol{y} \in X$, $\boldsymbol{x} \leq_P \boldsymbol{y}$ (resp. $\boldsymbol{y} \leq_P \boldsymbol{x}$) does not hold. The dual poset of P, denoted by P^{d}, is a poset which consists of the same ground set X, and the order relation $\leq_{P^{\mathrm{d}}}$ such that for all $\boldsymbol{x}, \boldsymbol{y} \in X$, $\boldsymbol{x} \leq_P \boldsymbol{y}$ if and only if $\boldsymbol{y} \leq_{P^{\mathrm{d}}} \boldsymbol{x}$. The comparability graph of poset P, denoted $G(P)$, is the undirected graph whose vertices are the elements of X and where two elements are adjacent if they are comparable in P. The *strict order relation* for $P = (X, \leq_P)$ is denoted by $<_P$ and defined for all $\boldsymbol{x}, \boldsymbol{y} \in X$ as $\boldsymbol{x} <_P y$ if $\boldsymbol{x} \leq_P \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}$. For each $\boldsymbol{x} \in X$, the set of *predecessors* (resp. *successors*) of $\boldsymbol{x}$ in P is defined by $\text{Pred}_P(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid \boldsymbol{y} <_P \boldsymbol{x}\}$ (resp. $\text{Succ}_P(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid \boldsymbol{x} <_P \boldsymbol{y}\}$). Moreover, we say that $\boldsymbol{x}$ is covered by $\boldsymbol{y}$ in P, and denote $\boldsymbol{x} \prec_P \boldsymbol{y}$, if $\boldsymbol{x} <_P \boldsymbol{y}$ and there is no element $\boldsymbol{z} \in X$ such that $\boldsymbol{x} <_P \boldsymbol{z}$ and $\boldsymbol{z} <_P \boldsymbol{y}$. For each $\boldsymbol{x} \in X$, we define the set of *immediate predecessors* (resp. *immediate successors*) of $\boldsymbol{x}$ in P which is $\text{ImPred}(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid \boldsymbol{y} \prec_P \boldsymbol{x}\}$ (resp. $\text{ImSucc}(x) = \{\boldsymbol{y} \in X \mid \boldsymbol{x} \prec_P \boldsymbol{y}\}$). An order $T = (X, \leq_T)$ is called a *tree* if it has a minimum $\boldsymbol{m}$ and for every $\boldsymbol{x} \in X$, $\boldsymbol{x} \neq \boldsymbol{m}$, $\boldsymbol{x}$ has a unique immediate predecessor. Given a chain $C = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_h\}$ of P such that $h > 1$ and $\boldsymbol{c}_1 \leq_P \boldsymbol{c}_2 \leq_P \ldots \leq_P \boldsymbol{c}_h$, we say that C is a *thin chain* if for $1 \leq d < h$ it holds that $\boldsymbol{c}_d$ is a unique immediate predecessor of $\boldsymbol{c}_{d+1}$.

Definition 2.1 (Abduction Matrix) Given $m \times n$-matrix $\mathbf{A} = [a_{i,j}]$, where $a_{i,j} \in \{0, 1, ?\}$ $(1 \leq i \leq m, 1 \leq j \leq n)$, $\mathbf{A}$ is an *abduction matrix* if the following holds:

1. $m \leq n$ (i.e., no less columns than rows).
2. Except for $a_{i_0 j_0} = ?$, the rest of the values are in $\{0, 1\}$.

The abduction will be done by deciding whether the value of $a_{i_0 j_0}$ is 0, 1 or should stay? in cases it is not possible to decide. The data in the matrix is suitable for abduction if each row represents a subject for comparison, and each column represents a property. Additionally, it is assumed the data has some underlying direction, e.g., if some properties apply to a subject, it indicates that other properties also apply. Note that the definition of abduction matrix may be extended to include cases with multiple missing values and not necessarily binary, but such matrices will not be discussed in the scope of this chapter.

Definition 2.2 (Matrix Abduction Method) Given an abduction matrix $\mathbf{A}$, the method will decide whether the value of $a_{i_0 j_0}$ is one of 0, 1, or ?. Let $\mathbf{A}^1$ denote the matrix in which $a_{i_0 j_0} = 1$ and let $\mathbf{A}^0$ denote the matrix in which $a_{i_0 j_0} = 0$. Let Π_1 and Π_0 denote the poset of the columns of $\mathbf{A}^1$ and $\mathbf{A}^0$, respectively, where the columns are used as vectors and compared coordinate wise. The method decides which of Π_0 and Π_1 is preferable, using the criteria defined below. If Π_0 is preferable to Π_1, it is decided that $a_{i_0 j_0} = 0$. If Π_1 is preferable to Π_0, $a_{i_0 j_0} = 0$. Otherwise, $a_{i_0 j_0} = ?$.

Poset Preference Definition

1. Let $L = \{\beta_1, \beta_2, \ldots\}$ be a set of labels. We denote by $M_1(L)$ the family of multisets of the form $\{m\alpha, \beta_1, \ldots, \beta_{k-1}\}$, i.e., every multiset which consists of $m \in \{0\} \cup \mathbb{N}$ copies of the label α, and up to one copy of each of the other labels in L. We refer to k as the *dimension* and refer to m as the *index* of the multiset.
2. Let E be a finite subset of $M_1(L)$. We define the *dimension* and the *index* of E as the maximum of the respective dimension and index of its elements.
3. Let $(X, \leq_P)$ be a poset, and L a set of labels.

(a) We say the function $f: X \to M_1(L)$ is a (L, f)-*realization* of $(X, \leq_P)$ if and only if the condition $\boldsymbol{x} \leq_P y \Leftrightarrow f(\boldsymbol{x}) \subseteq f(\boldsymbol{y})$ holds for each $\boldsymbol{x}, \boldsymbol{y} \in X$, and we say that the label(s) in $r \in M_1(L)$ is (are) *assigned* to the element $\boldsymbol{x} \in X$ if and only if $r \subseteq f(\boldsymbol{x})$. We define the *dimension* and the *index* of f as the dimension and the index of $\{f(\boldsymbol{x}) \mid \boldsymbol{x} \in X\}$. We say (L, f) is *label minimal* if and only if there is no other realization (L, f') of smaller dimension. We also denote the dimension of the label-minimal realization as dim*.

(b) We define the *index of connectivity* of $(X, \leq_P)$ as the amount of connected components of $(X, \leq_P)$.

(c) For each $\boldsymbol{x}, \boldsymbol{y} \in X$, we define $\boldsymbol{x} R \boldsymbol{y}$ as $x <_P y \vee y <_P x$. We denote by R^* the transitive closure of R. If $\boldsymbol{x} R^* \boldsymbol{y}$, then there exist $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k \in X$ such that $\boldsymbol{x} R_1 \boldsymbol{z}_1 R_0 \boldsymbol{z}_2, \ldots, R_i \boldsymbol{z}_k R_{1-i} \boldsymbol{y}$, $R_1, R_0 \in \{<_p, >_p\}$ and $R_1 \neq R_0$. We denote by $\rho(x, y)$ the minimal k for which such a $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k$ sequence exists. We define the *index of directional change* as $\max_{x,y} \rho(x, y)$, i.e., the maximal amount of direction changes necessary to travel between any two points in X.

4. We define partial order preference by comparing the following evaluations, where smaller values are preferred:

- amount of elements in the poset,
- index of connectivity,
- index of directional change,
- dimension of a label-minimal realization.

When comparing the partial orders Π and Π', if Π is preferable in at least one of these evaluations and Π' is preferable in none, we say Π is preferable to Π'. Note that these criteria may not be comprehensive and additional evaluations may be considered, depending on the context of the abduction. Also note that while the above definition is limited to solving only a single missing value in the abduction matrix, it may be extended to allow multiple missing values (instead of generating the matrices $\mathbf{A}^1$ and $\mathbf{A}^0$, we generate a separate matrix for every possible completion combination). Other existing methods for abduction may also reach a tie between different decisions, and a common solution for this problem, which may also be used in this method, provides additional information, up to a point where the decision is clear. In fact, when applying the matrix abduction method in an actual debate, e.g., in the context of Talmudic reasoning, each side in the debate is expected in its turn to introduce new information in order to affect the result of the abduction; therefore, the construction of the matrix is often gradual (for a detailed example, see [1, p. 338]).

Comparison of Matrix Abduction with Existing Algorithms We now revisit the example given in the introduction (in Table 1), and demonstrate the difference between the matrix abduction method and decision by the Euclidian distance (k-nearest neighbor algorithm).

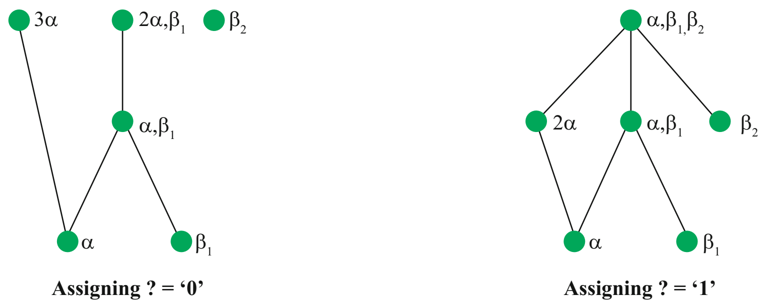


Fig. 2 Example minimal realizations

Euclidian Distance In this method we compare distances between the known values of column (6) and the other columns. Column (1) is the best match, and therefore we use 0 to complete the matrix.

(1)	(2)	(3)	(4)	(5)
2	1	1	1	1

Matrix Abduction In this method, we compare the evaluations of the aforementioned criteria.

Table 2 specifies the evaluations of the topological criteria over the posets in Fig. 2. As the table shows, the only difference between the two options is under the criteria of index of directional change, and since 1 is preferable, we use 1 to complete the matrix.

Definition 2.3 (Dimension and 2-dimension of a Poset) We denote the *dimension* of poset $P = (X, \leq_P)$ by $\dim(P)$, which is defined similar to $\dim^*$, with the exception that in this context the realization's range is a vector of natural numbers, i.e., $f: X \to \mathbb{N}^k, k \in \mathbb{N}$. We also similarly define the 2-*dimension* of P, denoted by $\dim_2(P)$, with the exception that in the context of $\dim_2$ the range is a vector of bits, i.e., $f: X \to \{0, 1\}^k, k \in \mathbb{N}$. In order to distinguish between different types of realizations, we use dim *realization*, $\dim^*$ *realization* and $\dim_2$ *realization*. Note that any $\dim^*$ realization is also a dim realization, and any $\dim_2$ realization is also a $\dim^*$ realization, therefore for any poset P, $\dim_2(P) \leq \dim^*(P) \leq \dim(P)$ holds. Also note that the evaluation of dim and $\dim_2$ is known to be NP-complete.

Table 2 Comparison of evaluated criteria for both possible values of a missing entry

Assumed value	Amount of elements	Index of connectivity	Index of directional change	$\dim^*$
0	6	1	2	2
1	6	1	1	2

3 Related Work and Algorithms

The first algorithm for evaluating $\dim^*$ was presented along with the concepts introduced in [1], but at the time of its publication, it was yet unknown whether the algorithm produces a precise evaluation of $\dim^*$ or not. Soon after, examples that contradict the algorithm's optimality were discovered and presented in [17] and made it clear that the algorithm may be improved in certain cases; moreover, Laufer identified and described in [13] the similarity between the definition of $\dim^*$ and the existing concept 2-dimension of a poset whose evaluation is a known NP-complete problem. This was a hint that the matrix abduction method is hard to evaluate, but also provided additional motivation for finding better $\dim^*$ approximations, since it may serve as an alternative to two dimensions in some of its existing practical uses.

Apart from [14], the connection between the matrix abduction method and existing common methods such as matrix completion algorithms or more general classification algorithms was yet to be thoroughly studied. At a glance, the assumptions that guide the completion process are useful in entirely different contexts and even demand different computational power to solve; therefore, if such a connection exists, it may not be simple to find.

4 Properties of Matrix Abduction

In this section, we provide several new theorems that may be useful when conducting an abduction, and for further research on $\dim^*$. We begin in Sect. 4.1 with additional preliminaries and known properties of $\dim^*$ which will be used implicitly in the following sections. Afterward, in Sect. 4.2 we provide proofs by reduction that both the exact evaluation of $\dim^*$ in the general case and the matrix abduction method are NP-complete. Finally, in Sect. 4.3 we discuss point-removal properties which may aid in the evaluation despite the aforementioned result. Note that in Sect. 4.2, it is assumed that the reader is familiar with basic topics in computational complexity theory, such as the 3-SAT problem, the NP-complete complexity class, and the process of proving NP-completeness by a polynomial-time reduction. (For a thorough discussion on these topics, see [11].)

4.1 Preliminaries

Definition 4.1 (Poset Levels) Let $P = (X, \leq_P)$ be a poset, and $M \subseteq X$ be the set of all minimal elements in P. We define the level of an element $\boldsymbol{x} \in X$ in P, denoted $\text{level}_P(x)$, by

$$\text{level}_P(\boldsymbol{x}) = \begin{cases} 1, & \boldsymbol{x} \in M, \\ 1 + \max_{\boldsymbol{e} \in \text{Pred}_P(\boldsymbol{x})}(\text{level}_P(e)), & \boldsymbol{x} \in X \setminus M, \end{cases}$$

i.e., the level of each element is greater by 1 the highest level of its predecessors in the poset. Note that all elements of the same level make an antichain.

dim* of Dual Posets [13] Given poset $P = (X, \leq_P)$, which was produced from the complete $m \times n$-abduction matrix $\mathbf{A} = [a_{i,j}]$ $(1 \leq i \leq m, 1 \leq j \leq n)$, as described in Definition 2.2, then P^{d} may be produced from the $m \times n$-matrix $\mathbf{A}^{\mathrm{d}} = [1 - a_{i,j}]$. Moreover, for any poset P, it holds that $\dim^*(P) = \dim^*(P^{\mathrm{d}})$.

Monotony of dim* Given poset $P = (X, \leq_P)$ and Q, a suborder of P, then it holds that $\dim^*(Q) \leq \dim^*(P)$. Proof is trivial, since every valid $\dim^*$ realization of P is also a valid realization of Q.

Definition 4.2 (Encoding) In the context of $\dim^*$ we abbreviate a set of assigned labels by pairing a natural number and bit vector encoding, e.g., for (L, f) $\dim^*$ realization, where $L = \{\alpha, \beta_1, \beta_2, \beta_3, \beta_4\}$, we abbreviate $\{3\alpha, \beta_1, \beta_3\}$ to 3,1010. We call the process of assigning labels to an element as *encoding* the element. Moreover, the encoding of multiple elements, which represents a label-minimal realization of the poset they make, is called *minimal encoding*.

Definition 4.3 (Encoding Cardinality) Given a poset $P = (X, \leq_P)$, $(L \cup \{\alpha\}, f)$ $\dim^*$ realization of P, and $\boldsymbol{x} \in X$, we denote by $f_\alpha(x)$ the amount of alpha label copies in $f(\boldsymbol{x})$. Also, $|f(x)|$ denotes the amount of distinct labels in $f(x)$, i.e., $|f(x)| = |L \cap f(x)| + k$, where k is 1 if $f_\alpha(x) \geq 1$ and is 0 otherwise.

Encoding Chains [13] Given a poset $P = (X, \leq_P)$ that is also a chain with $l \geq 1$ elements, then $\dim^*(P) = 1$ holds. Let $X = \{x_0, x_1, \ldots, x_{l-1}\}$ such that $x_0 < x_1 < \ldots < x_{l-1}$, then (L, f) $\dim^*$ realization of P, where $L = \{\alpha\}$, $f(x_i) = \{(i)\alpha\}$, provides a minimal encoding.

Encoding Antichains [13, 17] Given a poset $P = (X, \leq_P)$, and $A = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_l\} \subseteq X$, $L = \{\beta_1, \beta_2, \ldots, \beta_{\lceil \log_2(l) \rceil}\}$, we denote by $E_{L,A}(\boldsymbol{a}_i)$ the set of labels given by $\{\lfloor \log_2(l - i) \rfloor \alpha\} \cup \{\beta_{b+1} \mid \lfloor i/2^b \rfloor \bmod 2 = 1\}$, i.e., labels that correspond to the binary representation of each element's index, and the amount of α label copies is the minimal number that guarantees the encoding $E_{L,A}(\boldsymbol{a}_i)$ will not conflict with the encoding of any other encoding $E_{L,A}(\boldsymbol{x})$, $\boldsymbol{a}_i \neq \boldsymbol{x} \in X$. This method of encoding was first presented in [17], and proved to be optimal for antichains in [13], i.e., at least $\lceil \log_2(l) \rceil$ nonalpha labels must be modified in any encoding of the antichain. Moreover, given any poset P, it holds that $\dim^*(P) \geq 1 + \lceil \log_2(w(P)) \rceil$.

Lemma 4.4 (Encoding Ordered Antichains) *Given a poset P which consists of $h \geq 1$ ordered antichains $A_0, A_1, \ldots, A_{h-1}$, such that $A_0 \leq_P A_1 \leq_P \ldots \leq_P A_{h-1}$, then* $\dim^*(P) = 1 + \sum_{i=0}^{h-1} \lceil \log_2 |A_i| \rceil$.

Proof By induction:

For $h = 1$: Same as antichain encoding.

For any $h \geq 1$: We assume that for some $h \geq 1$, for any poset P_h which consists of h ordered antichains (as described above) it holds that $\dim^*(P_h) = 1 + \sum_{i=0}^{h-1} \lceil \log_2 |A_i| \rceil$ and $(L \cup \{\alpha\}, f)$ gives a minimal encoding.

For $h+1$: Consider the poset $P_{h+1} = (X \cup A_h \leq_{P_{h+1}})$, which is constructed by adding to P_h an antichain on the ground set A_h such that $A_0 \leq_{P_{h+1}} \dots \leq_{P_{h+1}} A_{h-1} \leq_{P_{h+1}} A_h$ holds. P_{h+1} may be encoded by the following $(L \cup L^{\text{opt}} f^{\text{opt}})$ $\dim^*$ realization:

$$f^{\text{opt}}(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}), & \boldsymbol{x} \in X, \\ L \cup \{(k)\alpha\} \cup E_{L^{\text{opt}}, A_h}(\boldsymbol{x}), & \boldsymbol{x} \in A_h, \end{cases}$$

where $k = \max_{\boldsymbol{x} \in X}(f_\alpha(\boldsymbol{x}))$ and $L^{\text{opt}} = \{\beta_1, \beta_2, \dots, \beta_{\lceil \log_2 |A_h| \rceil}\}$. Thus, it holds that $\dim^*(P_{h+1}) \leq \dim^*(P_h) + \lceil \log_2 |A_h| \rceil$. Moreover, for any $\boldsymbol{x} \in X$ and any $(L \cup L', f')$ $\dim^*$ realization of P_{h+1}, it holds that $f'(\boldsymbol{x}) \subseteq \bigcap_{\boldsymbol{y} \in A_h} f'(\boldsymbol{y})$. Since at least $\lceil \log_2 |A_h| \rceil$ labels must be modified in the encoding of A_h, then $\dim^*(P_{h+1}) \geq \dim^*(P_h) + \lceil \log_2 |A_{h+1}| \rceil$ must also hold. □

Lemma 4.5 (Encoding Minimal and Maximal Elements) *Given a poset* $(X, \leq_P)$, $(L \cup \{\alpha\}, f)$ $\dim^*$ *realization of* P, *and let* $M \subseteq X$ *be the set of all minimal* (*resp. maximal*) *elements in* P. *If* $|M| \geq 2$, *then for any* $x \in M$ *it holds that* $|f(\boldsymbol{x})| > 0$ (*resp.* $|f(\boldsymbol{x})| < |L| + 1$), *with the exception that for one maximal element* $\boldsymbol{x} \in X$, $|f(\boldsymbol{x})| = |L| + 1$ *may hold if* $f_\alpha(\boldsymbol{x}) < f_\alpha(\boldsymbol{y})$ *for any* $\boldsymbol{y} \in X$, $\boldsymbol{y} \neq \boldsymbol{x}$.

The proof is trivial, since contradicting these inequalities clearly causes a conflict.

4.2 Computational Complexity of Evaluating dim* *and Matrix Abduction*

Proposition 4.6 *The precise evaluation of* $\dim^*$ *is NP-complete*.

Proof Given a 3-SAT input, the CNF formula F which consists of $n \geq 3$ variables denoted v_i $(0 \leq i < n)$ and $k \geq 1$ clauses denoted by c_i $(0 \leq i < k)$, where each clause depends on exactly three variables, we construct the poset $R^*(F) = (X_{R^*}, \leq_{R^*})$ and show that defining (L, f) $\dim^*$ realization on $R^*(F)$ is as hard as finding an assignment which satisfies the given formula. The formula may be satisfied if and only if exists a $\dim^*$ realization such that $|L| = 4|n| + 1$.

Construction 4.7 (Suborders $\mathbf{VG}_i^*$) Let $\mathrm{VG}_0^*, \mathrm{VG}_1^*, \dots, \mathrm{VG}_{n-1}^*$ be similar, separate suborders of $R^*(F)$ each consisting of 10 elements. We denote the elements of suborder VG_i^* $\boldsymbol{eh}_i^*, \boldsymbol{el}_i^* \boldsymbol{f}, \boldsymbol{el}_i^* \boldsymbol{t}, \boldsymbol{el}_i^* \boldsymbol{w}_1, \dots, \boldsymbol{el}_i^* \boldsymbol{w}_7$, as illustrated in Fig. 3a.

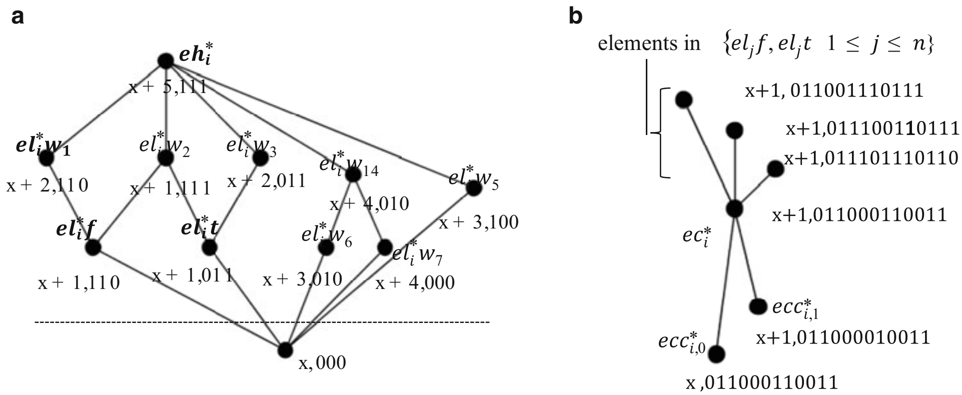


Fig. 3 Constructions VG_i^* (**a**) and CG_i^* (**b**)

Definition 4.8 (Spare Labels) Let $\boldsymbol{e} \in \{\boldsymbol{el}_i^* \boldsymbol{f}, \boldsymbol{el}_i^* \boldsymbol{t}\}$, poset P and a valid encoding f of P be given. We define $\mathrm{spare}(f(e)) = \{\beta_0, \beta_1, \ldots\}$ as the set of labels such that for any single label $\alpha \neq \beta \in \mathrm{spare}(f(e))$, if β is removed from $f(e)$ then f stays a valid encoding of P.

Remark 4.9 (Encoding Limitations of VG_i^*) Using an exhaustive scan of realizations, it is possible to prove that for every suborder VG_i^* it holds that $\dim^*(\mathrm{VG}_i^*) = 4$, and for every $0 \leq i < n$, either $\mathrm{spare}(f(\boldsymbol{el}_i^* \boldsymbol{f})) = \emptyset$ or $\mathrm{spare}(f(\boldsymbol{el}_i^* \boldsymbol{t})) = \emptyset$ for any minimal encoding of VG_i^* (for example, in the encoding illustrated in Fig. 3a, it is possible to remove a label that is used for encoding $\boldsymbol{el}_i^* \boldsymbol{t}$, but for $\boldsymbol{el}_i^* \boldsymbol{f}$ it is not). Note that every element $\boldsymbol{eh}_i^*$ is one of the multiple maximal elements in $R^*(F)$, along with another maximal element e for which $|f(e)| = \dim^*(R^*(F))$ (will be discussed later). Therefore, it holds that

$$|f(\boldsymbol{eh}_i^*)| < \dim^*(R^*(F)). \tag{4.1}$$

Construction 4.10 (Suborders $\mathbf{CG}_i^*$) Let $\mathrm{CG}_0^*, \mathrm{CG}_1^*, \ldots, \mathrm{CG}_{k-1}^*$ be similar, separate suborders of $R^*(F)$, each consists of three elements. We denote the elements of suborder CG_j^* by $\boldsymbol{ec}_j^*$, $\boldsymbol{ecc}_{j,0}^*$, and $\boldsymbol{ecc}_{j,1}^*$, as illustrated in Fig. 3b. The element $\boldsymbol{ec}_j^*$ is the immediate predecessor of $\boldsymbol{el}_i^* \boldsymbol{f}$ (resp. $\boldsymbol{el}_i^* \boldsymbol{t}$) if and only if c_j may be satisfied by the negative (resp. positive) literal of v_i.

Assuming each suborder VG_i^* is encoded by modifying distinct four nonalpha labels, then encoding every element $\boldsymbol{ec}_i^*$ by the labels shared by its predecessors will not pose a conflict between the elements $\boldsymbol{ec}_0^*, \ldots, \boldsymbol{ec}_{k-1}^*$. Additionally, since $\boldsymbol{ecc}_{i,0}^*$ and $\boldsymbol{ecc}_{i,1}^*$ make an antichain of size 2, at least one nonalpha must be modified in order to encode it. If for some immediate successor $\boldsymbol{e}$ of $\boldsymbol{ec}_i$ exists a label in $\mathrm{spare}(f(\boldsymbol{e}))$ it may be modified without causing conflicts between $\boldsymbol{ecc}_{i,0}^*$, $\boldsymbol{ecc}_{i,1}^*$ and any other element of any suborder VG_i^*.

Fig. 4 Procedure `set_realization*`

```
set_realisation*(WG*_i)
d ← 4(n − 1)
V ← {β_t | 0 ≤ t < 4i ∨ 4(i + 1) ≤ t < 4n}
while d > 2
  f(ew*_{i,d,0}) = V ∪ {(d − 1)α}
  r ← 4(n + i + 1) + d
  V ← V \ β_{r mod 4n}
  f(ew*_{i,d,1}) = V ∪ {(d)α}
  d ← d − 1
  f(ew*_{i,2}) = V ∪ {α}
  f(ew*_{i,1}) = V
```

Construction 4.11 **(Suborders WG_i^*)** Let $\mathrm{WG}_0^*, \mathrm{WG}_1^*, \ldots, \mathrm{WG}_{n-1}^*$ be similar, separate suborders of $R^*(F)$, which consist of $4(n-1)$-ordered antichains, as illustrated in Fig. 5. For WG_i^* the elements in levels $3, \ldots, 4(n-1)$ make antichains of width 2 and are denoted by $\boldsymbol{ew}_{i,d,0}^*$ and $\boldsymbol{ew}_{i,d,1}^*$ (where d corresponds the level of the element) and the two elements that are in levels 1 and 2, denoted by $\boldsymbol{ew}_{i,1}^*$ and $\boldsymbol{ew}_{i,2}^*$, make a chain. Each of these elements is a predecessor of all elements in VG_i^*, as illustrated in Fig. 4. We denote by $\boldsymbol{ewt}_i^*$ the set of seven elements in WG_i^* that are in levels 2–5, and denote by B_i^* the set of nonalpha labels that are used to encode the elements of $\boldsymbol{ewt}_i^*$. Encoding the elements of levels $3, \ldots, 4(n-1)$ in WG_i^* requires the modification of at least one nonalpha label which was not already modified for encoding elements in the higher levels. Note that every element $\boldsymbol{ew}_{i,1}^*$ is one of the multiple minimal elements in $R^*(F)$, along with another element $\boldsymbol{e}$ which is necessarily encoded by α labels only (will be discussed later). Therefore, $|f(\boldsymbol{ew}_{i,1}^*)| \geq 1$ and $f_\alpha(\boldsymbol{ew}_{i,1}^*) < f_\alpha(\boldsymbol{e})$ also holds. Consequently, the elements $\boldsymbol{ew}_{i,4(n-1),0}^*$ and $\boldsymbol{ew}_{i,4(n-1),1}^*$ must be encoded (together) using at least $4(n-1)$ nonalpha labels, and at least $4(n-1)$ copies of α label. Considering (4.1), and since suborders $\mathrm{VG}_0^*, \ldots, \mathrm{VG}_{n-1}^*$ are encoded by modifying at least three nonalpha labels, it holds that

$$\begin{aligned} 4(n-1)+1 &\leq \max\{|f(\boldsymbol{ew}_{i,4(n-1),0}^*)|, |f(\boldsymbol{ew}_{i,4(n-1),1}^*)|\} \\ &\leq |f(\boldsymbol{eh}_i^*)| - 3 \leq \dim^*(R^*(F)) - 4. \end{aligned} \tag{4.2}$$

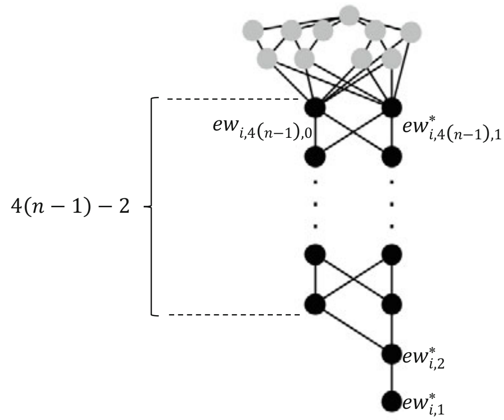
The elements of WG_i^* may be encoded using the labels $\beta_t | 0 \leq t < 4n \cup \{\alpha\}$, using the procedure `set_realization*`, as shown in Fig. 4.

Remark 4.12 (Correcting WG_i^* Instances) We assume the encoding defined in `set_realization*` will be a part of the minimal encoding, and that it affects accordingly on the encoding freedom of suborders $\mathrm{VG}_0^*, \ldots, \mathrm{VG}_{n-1}^*$. Additionally, we assume that for every element $\boldsymbol{ew}_{i,0,1}^*$, $\boldsymbol{ec}_j^*$, $\boldsymbol{ecc}_{j,0}^*$ and $\boldsymbol{ecc}_{j,1}^*$, the following holds:

1. $f_\alpha(\boldsymbol{ec}_j^*) = f_\alpha(\boldsymbol{ecc}_{j,1}^*) = 4(n-1)+1$.
2. $4(n-1) = f_\alpha(\boldsymbol{ecc}_{j,0}^*) = f_\alpha(\boldsymbol{ew}_{i,0,1}^*)$.

With these assumptions, we apply element contractions for elements that are given identical encoding, and fix the order between the elements to match the encoding. That is, we assume that the spare label that is used to encode $\boldsymbol{ec}_j^*$ is also used to encode $\boldsymbol{ecc}_{j,0}^*$; thus, the

Fig. 5 Construction WG_i^*



only immediate successor of $\boldsymbol{ecc}^*_{j,0}$ is $\boldsymbol{ec}^*_j$, and the predecessors of $\boldsymbol{ecc}^*_{j,0}$ and $\boldsymbol{ecc}^*_{j,1}$ depend solely on the labels that are necessarily shared by the successors of $\boldsymbol{ec}^*_j$. We now define an additional suborder which compels the minimal encoding to match these assumptions.

Construction 4.13 (Suborder AG) Let AG be a suborder of $R^*(F)$, which consists of $8n-7$-ordered antichains as follows:

1. The element in level $8n-3$, denoted $\boldsymbol{ea}^*_{8n-3}$, is a maximal element of $R^*(F)$.
2. The elements in levels $8n-4 \geq d > 4n-4$ make antichains of width 2 and are denoted $\boldsymbol{ea}^*_{d,0}$, $\boldsymbol{ea}^*_{d,1}$, where d corresponds to the level of the element.
3. The elements in levels $4n-4 \geq d \geq 1$ make a chain, and are denoted $\boldsymbol{ea}^*_d$, where d corresponds to the level of the element.

The purpose of the suborder AG is to limit the relative amount of α labels given to elements in suborders $\mathrm{WG}_0^*, \ldots, \mathrm{WG}_{n-1}^*$ by introducing a chain of height $4(n-1)-1$ that is necessarily encoded by α labels only.

The element $\boldsymbol{ea}^*_{8n-3}$ is a successor of every element in each of the suborders VG_0^*, $\ldots$, VG_{n-1}^* except for the elements $\boldsymbol{eh}^*_0, \ldots, \boldsymbol{eh}^*_{n-1}$. Similarly, each of the elements $\boldsymbol{ea}^*_{d,0}$ and $\boldsymbol{ea}^*_{d,1}$ is the successor of these elements, with the exception that for $7n-3 \leq d < 8n-3$, $\boldsymbol{ea}^*_{8n-4-i,1}$ and its predecessors are not successors of any element that its assumed encoding includes the label given by $f(\boldsymbol{ew}^*_{i,1})$ as assigned by `set_realization*`. Note that $\boldsymbol{ea}^*_{7n-3,1}$ is the successor of only $\boldsymbol{ea}^*_{7n-3-1,1}$, $\boldsymbol{ea}^*_{7n-3-1,0}$, and the added constraints are independent from asymmetric decisions. In order to limit the freedom of α labels given to element in WG_i^*, we assume that for $d \leq 4n-4$ any element $\boldsymbol{ea}^*_d$ is encoded by $(d)\alpha$ and that the assumptions described in suborders WG_i^* definition also hold; accordingly, we insert constraints to every element $\boldsymbol{ea}^*_d$ from elements in suborders $\mathrm{GG}_0^*, \ldots, \mathrm{GG}_{k-1}^*$, and $\mathrm{WG}_0^*, \ldots, \mathrm{WG}_{n-1}^*$. For $0 \leq i \leq n-1$, $1 \leq d \leq 4(n-1)$, these constraints imply that $\boldsymbol{ea}^*_d <_{R^*} \boldsymbol{ew}^*_{i,d,1}$ holds, but $\boldsymbol{ea}^*_d$ and $\boldsymbol{ew}^*_{i,d,0}$ are incomparable.

For each level $8n-4 \leq d \leq 4n-4$, at least one nonalpha label must be modified in order to encode the elements $\boldsymbol{ea}^*_{d,0}$ and $\boldsymbol{ea}^*_{d,1}$, and therefore $f(\boldsymbol{ea}^*_{8n-3}) \geq 4n+1$ must hold in any encoding. Since the element $\boldsymbol{ea}^*_{8n-3}$ is a maximal element, along with every element $\boldsymbol{eh}^*_i$, then under the assumption that $\dim^*(R^*) = 4n+1$, it holds that $f(\boldsymbol{ea}^*_{8n-3}) = 4n+1$, and for every element $\boldsymbol{eh}^*_i$, $f(\boldsymbol{eh}^*_i) = 4n$. Moreover, for $d \leq 4n-4$, no nonalpha labels may be used for encoding any element $\boldsymbol{ea}^*_d$. Since this also applies to $\boldsymbol{ea}^*_{8n-4}$, which is a minimal element along with every element $\boldsymbol{ew}^*_{i,4(n-1)}$, then $f_\alpha(\boldsymbol{ea}^*_{8n-4}) = 1$, $|f(\boldsymbol{ew}^*_{i,4(n-1)})| = 1$ and $f_\alpha(\boldsymbol{ew_{i,4(n-1)}}^*) = 0$ hold.

Construction 4.14 (Poset $R^*(F)$) The poset $R^*(F)$ is defined by the composition of:

1. $\mathrm{VG}^*_0, \ldots, \mathrm{VG}^*_{n-1}$, where every element $\boldsymbol{eh}^*_i$ is a maximal element of $R^*(F)$.
2. $\mathrm{CG}^*_0, \ldots, \mathrm{CG}^*_{k-1}$, connected to $\mathrm{VG}^*_0, \ldots, \mathrm{VG}^*_{n-1}$ according to the given formula.
3. $\mathrm{WG}^*_0, \ldots, \mathrm{WG}^*_{n-1}$ after contractions are applied, including the additional implied constraints, where every element $\boldsymbol{ew}^*_{i,1}$ is a minimal element of $R^*(F)$.
4. AG, where $\boldsymbol{ea}^*_{8n-3}$ is a maximal element of R^*, and $\boldsymbol{ea}^*_1$ is a minimal element of $R^*(F)$.

For any encoding f of $R^*(F)$ that makes use of no more than $4n$ nonalpha labels, the inequality (4.2) and the following limitations hold:

1. $f_\alpha(\boldsymbol{ew}^*_{i,d,0}) < f_\alpha(\boldsymbol{ew}^*_{i,d,1})$, since $\boldsymbol{ea}^*_d <_R \boldsymbol{ew}^*_{i,d,1}$ holds, but $\boldsymbol{ea}^*_d$ and $\boldsymbol{ew}^*_{i,d,0}$ are incomparable.
2. For any $i, j\, 0 \leq d < l \leq n-1$, it holds that $f_\alpha(\boldsymbol{ew}^*_{i,d,0}) < f_\alpha(\boldsymbol{ew}^*_{j,l,0})$ and $f_\alpha(\boldsymbol{ew}^*_{i,d,1}) < f_\alpha(\boldsymbol{ew}^*_{j,l,1})$, for a similar reason.
3. $|f(\boldsymbol{eh}^*_i)| = |f(\boldsymbol{ew}^*_{i,4(n-1),0})| + 3 = 4n$ (derived from (4.2) and the above limitations).
4. Given $\boldsymbol{ew}^*_{i,d,0}, \boldsymbol{ew}^*_{i,d-1,0}, \boldsymbol{ew}^*_{j,d2,0}$ such that $\boldsymbol{ew}^*_{i,d,0} \geq_{R^*} \boldsymbol{ew}^*_{j,d2,0} \wedge \boldsymbol{ew}^*_{i,d-1,0} \not\geq_{R^*} \boldsymbol{ew}^*_{j,d2,0}$ then $\exists \beta \in f(\boldsymbol{ew}^*_{j,d2,0}) \cap f(\boldsymbol{ew}^*_{i,d,0}) \wedge \beta \notin f(\boldsymbol{ew}^*_{i,d+1,0}) \wedge \beta \neq \alpha$. This also holds for $\boldsymbol{ew}^*_{j,d2}$ (i.e., the two minimal elements in each suborder WG^*_i).
5. Due to the definition of `set_realization*` for any two chains WG^*_i and WG^*_t, where $t \neq (n+i-1) \bmod n$, there are exactly four consecutive elements in WG^*_i, that are not in $\boldsymbol{ewt}^*_i$ and are direct successors of elements in $\boldsymbol{ewt}^*_t$, for which the condition described in the limitation 4 applies.[1] Consequently, the only labels that may be used for encoding $\boldsymbol{ewt}^*_i$ are completely distinct from the ones in $\boldsymbol{ewt}^*_t$, i.e., $B^*_t \cap B^*_i = \emptyset$. Clearly, for any $n \geq 3$ this applies to any two chains $\mathrm{WG}^*_i, \mathrm{WG}^*_t$.
6. For every $\boldsymbol{ew}^*_{i,4(n-1),0}, t$, where $0 \leq t < n$ and $t \neq (n+i-1) \bmod n$, it holds that $B^*_t \subset f(\boldsymbol{ew}^*_{i,4(n-1),0})$. Therefore, VG^*_i must be encoded by modifying only α labels, and nonalpha labels that are in $B^*_{(n+i-1) \bmod n}$, and $|f(\boldsymbol{eh}^*_i) \cup f(\boldsymbol{ew}^*_{(n+i-1) \bmod n,1,0})| = L+1$.
7. Given $\boldsymbol{ec}^*_i, \boldsymbol{ep}_1 \in \{\boldsymbol{el}^*_{p_1}\boldsymbol{f}, \boldsymbol{el}^*_{p_1}\boldsymbol{t}\}$, $\boldsymbol{ep}_2 \in \{\boldsymbol{el}^*_{p_2}\boldsymbol{f}, \boldsymbol{el}^*_{p_2}\boldsymbol{t}\}$, $\boldsymbol{ep}_3 \in \{\boldsymbol{el}^*_{p_3}\boldsymbol{f}, \boldsymbol{el}^*_{p_3}\boldsymbol{t}\}$, where $\boldsymbol{ep}_1, \boldsymbol{ep}_2, \boldsymbol{ep}_3$ are direct successors of $\boldsymbol{ec}^*_i$, then $\boldsymbol{eh}^*_{p1}$, $\boldsymbol{eh}^*_{p2}$ and $\boldsymbol{eh}^*_{p3}$ are also successors of $\boldsymbol{ec}^*_i$, and any other element in any suborder VG^*_j is incomparable with $\boldsymbol{ec}^*_i$, $\boldsymbol{ecc}^*_{i,0}$ and $\boldsymbol{ecc}^*_{i,1}$. Additionally, for some element $\boldsymbol{e} \in \{\boldsymbol{ep}_1, \boldsymbol{ep}_2, \boldsymbol{ep}_3\}$ $\mathrm{spare}(f(\boldsymbol{e})) \neq \emptyset$ must hold in order to avoid conflicting encodings with other elements.

[1] The only exception to this is when elements are shared (contracted), but the desired limitation still holds.

Considering these limitations, it is possible to define $(L \cup \{\alpha\}, f)$ dim* realization of $R^*(F)$, where $L = \{\beta_t \mid 0 \le t < 4n\}$ if and only if the formula may be satisfied by the assignment $S = \{r_i \mid 0 \le i < n\}$, and that realization is given by:

1. Encode WG_i^* instances as defined in `set_REALIZATION*`.
2. Encode AG according to the following rules:
   - $f(\boldsymbol{ea}^*_{8n-3}) = L \cup \{(8n-3)\alpha\}$.
   - $f(\boldsymbol{ea}^*_{8n-4,0}) = L \cup \{(8n-4)\alpha\}$.
   - $f(\boldsymbol{ea}^*_{8n-4,1}) = L \setminus \beta_1 \cup \{(8n-3)\alpha\}$.
   - For $4n-4 < d \le 8n-5$, $f(\boldsymbol{ea}^*_{d,0}) = f(\boldsymbol{ea}^*_{d-1,0}) \setminus \{\alpha, \beta_d\}$ and $f(\boldsymbol{ea}^*_{d,1}) = f(\boldsymbol{ea}^*_{d-1,1}) \setminus \{\alpha, \beta_d\}$.
   - For $1 \le d \le 4n-4$, $f(\boldsymbol{ea}^*_d) = \{(d)\alpha\}$.
3. Encode each suborder VG_i^* by modifying the nonalpha labels in $B^*_{(n+i-1) \bmod n}$, as illustrated in Fig. 3a, where $f(\boldsymbol{eh}^*_i) = L(8n-2)\alpha f(\boldsymbol{ew}^*_{(n+i-1) \bmod n,1})$ and we insure that spare$(f(\boldsymbol{el}^*_i \boldsymbol{f})) \ne \emptyset$, if $r_i =$ false, or that spare$(f(\boldsymbol{el}^*_i \boldsymbol{t})) \ne \emptyset$ otherwise.
4. Encode each element $\boldsymbol{ec}^*_i$, $\boldsymbol{ecc}^*_{i,0}$ and $\boldsymbol{ecc}^*_{i,1}$ as described in Construction 4.10. The element $\boldsymbol{ecc}^*_{i,0}$ and $\boldsymbol{ecc}^*_{i,1}$ may be encoded with no conflicts if and only if clause c_i may be satisfied by S.

If and only if the formula may be satisfied, we can derive a satisfying assignment for it from any valid encoding according to the amount of labels used to encode the elements in suborders $\mathrm{VG}_0^*, \ldots, \mathrm{VG}_{n-1}^*$. □

Remark 4.15 Note that the above reduction may also be used to prove hardness of additional, similar poset dimension types (e.g., dim$_2$) with relatively simple modifications.

Graph Representation Theorem [1] *Let $P = (S, <)$ be poset on the ground set $S = \{a_1, \ldots, a_m\}$. Then there exists a definite matrix with m columns and m rows such that the column ordering is the same as $(S, \le)$, and it is given by the matrix $\mathbf{A} = [a_{i,j}]$, $1 \le i, j \le m$, defined by $a_{i,j} = 1$ if and only if $a_i \le a_j$.*

Theorem 4.16 Given an abduction matrix $\mathbf{A}$ and two posets R_0 and R_1 produced by assigning 0 and 1, respectively, in the missing entry of $\mathbf{A}$, then the decision whether dim$^*(R_0)$ is either greater, less, or equal to dim$^*(R_1)$ is NP-complete.

Proof Let $F = (V, C)$ be a 3-SAT input, the CNF formula F (as given in the previous reduction). We construct an abduction matrix with a single missing value b, and we show that deciding whether $b = 0$ or should remain undecided, according to the relevant criteria values that depend on b (amount of elements in the poset, index of connectivity, index of directional change and dim*) is as hard as deciding whether F may be satisfied or not.

Construction 4.17 (the Posets $R_1^*(F)$, $R_0^*(F)$)

1. Arbitrarily choose one clause c_i, which depends on the variables v_{p1}, v_{p2}, v_{p3}. Let $c_k = \neg c_i$ and $F' = (V, C \cup \{c_k\})$ be the formula that is necessarily unsatisfiable and consists of $k + 1$ clauses.
2. Construct the posets $R^*(F)$ and $R_1^*(F) = R^*(F')$, as previously described in Construction 17.
3. Translate $R_1^*(F)$ into a matrix $\mathbf{A}^1 = [a_{i,j}]$, $1 \leq ij \leq |R_1^*(F)|$ as described in the graph representation theorem. Let the elements $\boldsymbol{ecc}_{k,0}^*$ and $\boldsymbol{ecc}_{k,1}^*$ (which are associated with the newly added negated clause) be represented by two vectors $\boldsymbol{V}_{k1}, \boldsymbol{V}_{k2} \in M$, respectively. Note that the vectors are identical, except for four entries $a_{k1,r}, a_{k1,t} \in \boldsymbol{V}_{k1}$ and $a_{k2,r}, a_{k2,t} \in \boldsymbol{V}_{k2}$, for which $a_{k1,r} > a_{k2,r}$ and $a_{k2,t} > a_{k1,t}$ hold. Let $\mathbf{A}^0$ be the matrix produced by replacing the value of $a_{k1,r}$ from 1 to 0, and denote by $R_0^*(F)$ the poset it represents.

Since it is not possible to satisfy the formula F', it is also not possible to encode $R_1^*(F)$ using $L = \{\beta_t \mid 0 \leq t < 4n\}$ and α labels alone. However, by using an additional non-alpha label β_{4n}, we may encode the poset by ensuring that for any $\boldsymbol{e} \in \{\boldsymbol{el}_i^* \boldsymbol{f}, \boldsymbol{el}_i^* \boldsymbol{t}\}$, $\beta_{4n} \in \text{spare}(f(\boldsymbol{e}))$ holds, making it certain that a valid encoding is possible regardless of the satisfiability of F; thus, $\dim^*(R^*(F)) \leq \dim^*(R_1^*(F)) = 4n + 2$.

In order to show that the decision whether $\mathbf{A}^0$ is preferable to $\mathbf{A}^1$ depends solely on the satisfiability of F, we show that checking whether $\dim^*(R_0^*(F)) = \dim^*(R_1^*(F))$ depends on the satisfiability of F, and the evaluation of the additional criteria of the matrix abduction method is necessarily tied and therefore does not affect the decision.

Dimension It is easy to see that the poset $R_0^*(F)$ may be encoded using $L = \{\beta_t \mid 0 \leq t < 4n\}$ and α labels if and only if F may be satisfied. Consider the hierarchy of the elements $\boldsymbol{ec}_k^*, \boldsymbol{ecc}_{k,0}^*, \boldsymbol{ecc}_{k,1}^*$ as it is in $R_1^*(F)$, compared to the hierarchy of the respective elements in $R_0^*(f)$, as illustrated in Fig. 6. For $R_0^*(F)$, the element $\boldsymbol{ec}_k^*$ may be encoded using one less nonalpha label, since it has one predecessor instead of two; thus an α label copy may be modified instead of a nonalpha label. Therefore, $\dim^*(R_0^*(F)) = \dim^*(R^*(F))$ holds, and $\dim^*(R_0^*(F)) = \dim^*(R^*(F)) = \dim^*(R_1^*(F))$ holds if and only if F may not be satisfied.

Amount of elements in the poset. Due to $R^*(F)$ construction definition, $R_0^*(F)$ represents the same amount of elements as $R_1^*(F)$ does.

Index of connectivity. Similarly, due to $R^*(F)$ construction definition, every element is connected, and the index of connectivity of both $R_0^*(F)$ and $R_1^*(F)$ is 1.

Index of directional change. It is easy to see that the change in the hierarchy of the elements $E = \{\boldsymbol{ec}_k^*, \boldsymbol{ecc}_{k,0}^*, \boldsymbol{ecc}_{k,1}^*\}$ does not affect any of the longest sequences that goes through the minimal necessary direction changes, even if the sequence goes through them. Given such a sequence S, one of the following applies:

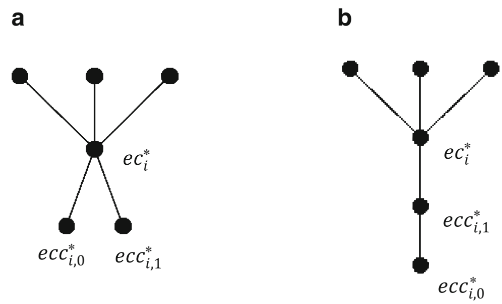


Fig. 6 CG_i^* hierarchy of $R_1^*(F)$ (**a**) and $R_0^*(F)$ (**b**)

- If S begins and ends in predecessors and successors of elements in E:
  In both $R_0^*(F)$ and $R_1^*(F)$, going through any subset of E will not require any directional changes. Moreover, any path that goes through a subset of these elements multiple times (i.e., after at least one direction change occurs) may necessarily be shortened.
- If S starts or ends at an element in E:
  Since no element in E is maximal or minimal, the travel may be prolonged into a predecessor/successor without increasing the amount of direction changes; thus this case is similar to the first.
- If S starts and ends at an element in E:
  This implies that the index of directional change is smaller than 2, which is necessarily false by the definition of $R_0^*(F)$ and $R_1^*(F)$. □

4.3 *Characterizing* **dim*** *in the Context of Matrix Abduction*

Consider we are given the abduction matrix **A** with a single missing entry, the complete matrices $\mathbf{A}^0$ and $\mathbf{A}^1$ produced by assigning either 0 or 1 in the missing entry of **A**, and the posets P_0 and P_1 that $\mathbf{A}^0$ and $\mathbf{A}^1$ represent, respectively. The only possible difference between P_0 and P_1 is limited to a single element, since the difference is limited to a single column in the matrix. We now show to what extent this difference affects $\dim^*(P_0)$ compared to $\dim^*(P_1)$, i.e., we delimit $|\dim^*(P_0) - \dim^*(P_1)|$. For this purpose, three separate cases should be considered:

1. $|P_0| \neq |P_1|$.
2. $P_0 = P_1$.
3. $|P_0| = |P_1| \wedge P_0 \neq P_1$.

We reach case 1 if the assigned value in one of the complete matrices makes for a column that overlaps the existing one; thus the matrix will represent a poset with one element removed. In this case, Proposition 21, which we show next, may be used.

Reaching case 2 means that either the assigned value produces an overlapping column in both complete matrices, or that the assigned value makes for a column representing an element which is positioned similarly in the hierarchy of the poset represented by both matrices. Clearly, handling this case is trivial.

We reach case 3 if the above conditions do not hold, and we regard this as *repositioning* of the element $\boldsymbol{e} \in P_0, P_1$. We show in Theorem 4.19 that in this case it holds that $|\dim^*(P_i) - \dim^*(P_{1-i})| \leq 1$.

One-point removal theorem for $\mathbf{dim_2}$ (Folklore) *Given a poset* $P = (X \cup \{\boldsymbol{e}\}, \leq_P)$, *it holds that* $\dim_2(P \setminus \{\boldsymbol{e}\}) \leq \dim_2(P) \leq \dim_2(P \setminus \{\boldsymbol{e}\}) + 2$.

Proof A proof for this theorem may be found in [15] as follows: The left part of the inequality holds due to monotony of $\dim_2$. Concerning the right part, we show that for any valid $\dim_2$ realization of $P \setminus \{\boldsymbol{e}\}$, by using two additional labels, we can encode P. Let (L, f) be a label-minimal $\dim_2$ realization of $P \setminus \{\boldsymbol{e}\}$. We define the valid $\dim_2$ realization $(L \cup \{\beta_1, \beta_2\}, f')$ of P (for $\beta_1, \beta_2 \notin L$) by:

$$f'(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}), & \boldsymbol{x} <_P \boldsymbol{e}, \\ f(\boldsymbol{x}) \cup \{\beta_1, \beta_2\}, & \boldsymbol{e} <_P \boldsymbol{x}, \\ f(\boldsymbol{x}) \cup \{\beta_1\}, & \boldsymbol{e} \nleq_P \boldsymbol{x} \wedge \boldsymbol{e} \ngeq_P \boldsymbol{x}, \end{cases}$$

and $f'(\boldsymbol{e}) = \{\beta_2\} \cup (\bigcup_{\boldsymbol{x} <_P \boldsymbol{e}} f'(\boldsymbol{x}))$. □

Proposition 4.18 Given the poset P, it holds that $\dim^*(P \setminus \{\boldsymbol{e}\}) \leq \dim^*(P) \leq \dim^*(P \setminus \{\boldsymbol{e}\}) + 2$, and this inequality is tight in the general case (proof is given in the appendix).

Theorem 4.19 *Given an abduction matrix* $\mathbf{A}$ *and two posets* P_0 *and* P_1 *produced by assigning* 0 *and* 1, *respectively*, *in the missing entry of* $\mathbf{A}$, *then*:

1. *If* $|P_i| < |P_{1-i}|$, *then* $\dim^*(P_i) \leq \dim^*(P_{1-i}) \leq \dim^*(P_i) + 2$.
2. *If* $|P_i| = |P_{1-i}|$ *and* $P_i \neq P_{1-i}$, *then* $|\dim^*(P_i) - \dim^*(P_{1-i})| \leq 1$.

Proof Let (L_0, f_0) be a $\dim^*$ realization of P_0. We show that if $|P_0| = |P_1|$ holds, we may produce $\dim^*$ realization $(L_0 \cup \{\beta\}, f_1)$ of P_1, $\beta \notin L_0$. Let $B_0, T_0, B_1, T_1 \subseteq P_0 \cap P_1$ such that $B_0 <_{P_0} \{\boldsymbol{e}\} <_{P_0} T_0$ and $B_1 <_{P_1} \boldsymbol{e} <_{P_1} T_1$. Let $\boldsymbol{E}_0$ and $\boldsymbol{E}_1$ be the columns in $\mathbf{A}^0$ and $\mathbf{A}^1$, respectively, that correspond to the repositioned element $\boldsymbol{e}$. The column $\boldsymbol{E}_1$ may be produced by altering a single value of $\boldsymbol{E}_0$ from 0 to 1. Therefore, it holds that $B_0 \subseteq B_1$ and $T_1 \subseteq T_0$, since $\boldsymbol{E}_1$ makes for a poset in which $\boldsymbol{e}$ may only become a successor of additional elements, and may no longer be a predecessor of some elements.

Lemma 4.20 $B_1 \cap T_0 = \emptyset$, *and for any* $\boldsymbol{t} \in T_0 \setminus T_1$, $\boldsymbol{b} \in B_1 \setminus B_0$, $\boldsymbol{t}$ *and* $\boldsymbol{b}$ *are incomparable.*

Proof Let $\boldsymbol{t} \in T_0$, $\boldsymbol{b} \in B_1 \setminus B_0$ and let their corresponding columns be $\boldsymbol{T}, \boldsymbol{B} \neq \boldsymbol{E}$, respectively. We denote by $a_{i,j}$ the entry of the value that differs between $\boldsymbol{E}_0$ and $\boldsymbol{E}_1$, and denote by $a_{i,k} \in \boldsymbol{T}$, $a_{i,m} \in \boldsymbol{B}$ the values that are compared with $a_{i,j}$ for the purpose of translating the matrix into a poset. The effect on the hierarchy between $\boldsymbol{e}$ and $\boldsymbol{t}$ that is caused by altering the value of $a_{i,j}$ from to 1 falls into one of the following two cases:

1. $\boldsymbol{e}$ and $\boldsymbol{t}$ become incomparable, since $1 = a_{i,j} > a_{i,k} = 0$ and for some other row n, it holds that $a_{n,k} > a_{n,j}$.
2. $\boldsymbol{e} <_{P_1} \boldsymbol{t}$ still holds, since $a_{i,k} = a_{i,j} = 1$ and for some other row n it holds that $a_{n,k} > a_{n,j}$.

(We disregard the option of $\boldsymbol{E}$ and $\boldsymbol{T}$ becoming identical, since the case of $|P_0| \neq |P_1|$ was discussed previously.) In none of the cases a successor of $\boldsymbol{e}$ in P_0 turns into a predecessor of $\boldsymbol{e}$ in P_1, and therefore $B_1 \cap T_0 = \emptyset$.

As stated in case 1, if $\boldsymbol{t} \in T_0 \setminus T_1$, then $a_{i,k} = 0$ and for some other row n it holds that $a_{n,k} > a_{n,j}$. In an opposite manner, for $b \in B_1 \setminus B_0$, it holds that $a_{i,m} = 1$, and for the same row n it holds that $1 = a_{n,k} > a_{n,j} = a_{n,m} = 0$; therefore, any such $\boldsymbol{b}$ and $\boldsymbol{t}$ are incomparable. □

We are now ready to define f_1.

1. If $B_1 \setminus B_0 = \emptyset$, define:

$$f_1(\boldsymbol{x}) = \begin{cases} f_0(\boldsymbol{x}), & \boldsymbol{x} \not\geq_{P_1} \boldsymbol{e}, \\ f_0(\boldsymbol{x}) \cup \{\beta\}, & \boldsymbol{x} \geq_{P_1} \boldsymbol{e}. \end{cases}$$

Since in this case the only difference between P_0 and P_1 is that $\boldsymbol{e}$ no longer precedes the elements in T_0/T_1, we may use the additional label $\beta \notin L_0$ to enforce this, and the encoding is valid.

2. If $B_1 \setminus B_0 \neq \emptyset$, define:

$$f_1(\boldsymbol{x}) = \begin{cases} L_B \cup f_0(\boldsymbol{e}) \setminus \{(f_\alpha(L_B))\alpha\}, & \boldsymbol{x} = \boldsymbol{e}, \\ f_0(\boldsymbol{x}), & \boldsymbol{x} \in B_1, \\ \{\beta\} \cup f_0(x), & \boldsymbol{x} \in P_1 \setminus (\{\boldsymbol{e}\} \cup B_1), \end{cases}$$

where $L_B = \bigcup_{\boldsymbol{b} \in B_1 \setminus B_0}(f_0(\boldsymbol{b}))$, and $\beta \notin L_0$. Note that for every $\boldsymbol{t} \in T_1$, it holds that $f_1(\boldsymbol{e}) \subset f_1(\boldsymbol{t})$ since if $\boldsymbol{e} <_{P_1} \boldsymbol{t}$, then for any $\boldsymbol{b} \in B_1$ it also holds that $\boldsymbol{b} <_{P_1} \boldsymbol{t}$. Also note that for any $\boldsymbol{t} \in T_0 \setminus T_1$, $f_1(\boldsymbol{e})$ and $f_1(\boldsymbol{t})$ are incomparable since $\beta \in f_1(\boldsymbol{t}) \wedge \beta \notin f_1(\boldsymbol{e})$ holds, and due to Lemma 23, there also exists an element $\boldsymbol{b} \in B_1 \setminus B_0$ such that $f_1(\boldsymbol{e}) \supseteq f_0(\boldsymbol{b}) \wedge f_0(\boldsymbol{b}) \not\subseteq (\{\beta\} \cup f_0(\boldsymbol{t})) = f_1(\boldsymbol{t})$; therefore, the encoding of all elements is valid.

Even though the above refers to the limitations derived from turning $a_{i,j}$ from 0 to 1, considering the duality properties of dim*, it is easy to see that the it is also true for turning $a_{i,j}$ from 1 to 0, and Theorem 4.19 holds in the general case. □

Remark 4.21 The matrix abduction method may be extended in order to complete multiple missing values of a matrix, and be used as a part of an ongoing debate. Theorem 2 may be particularly useful in such cases.

Example 1 Consider a debate over a certain undecided entry, where the losing side wants to overturn the current result of the matrix abduction. Suppose it has the ability to either introduce a new column to the matrix (which translates into a new element in the poset) or insist instead on altering a previously established value. We now know that unless altering the value would turn a column into a duplication of an existing column, this option is discouraged since it may, at best, lead to a tie, but never overturn the result of the abduction.

Example 2 Consider a debate over n entries at once in the same matrix. In this case, not only that the evaluation of $\dim^*$ for each complete matrix is hard, there may also be many value combinations to test. Fortunately, we now know that the maximal difference of $\dim^*$ between any two complete matrices depends directly on the amount of altered values; therefore, if we can identify two combinations for which the difference of $\dim^*$ is maximal, according to the theorem, then the need for testing additional combinations may be spared.

4.4 Specific Cases that Are Easy to Evaluate

Before devising an effective algorithm for evaluating the result of the matrix abduction method, we would like to identify cases in which it may be possible to easily evaluate the result of the abduction, ideally sparing the need for using an algorithm altogether. The next two propositions generalize some of these cases.

Proposition 4.22 *Given a poset P, and $\boldsymbol{e} \in P$ the minimum or the maximum of P, then* $\dim^*(P) = \dim^*(P \setminus \{e\})$.

Proof Let $P_1 = (X, \leq_P)$, P_0 be the two posets, where $\{\boldsymbol{e}\} \in P_1$ is the minimum of P_1, P_0 is the order induced by P_1 on $X \setminus \{\boldsymbol{e}\}$, and let (L, f) be a $\dim^*$ realization of P_0. Then, we may produce the valid $\dim^*$ realization (L', f') of P_1 as follows:

$$f'(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}) \cup \{\alpha\}, & \boldsymbol{x} \neq \boldsymbol{e}, \\ \emptyset, & \boldsymbol{x} = \boldsymbol{e}. \end{cases}$$

For any poset P such that $|P| > 1$, there exists a label-minimal $\dim^*$ realization in which α label is already used, and therefore $L' = L$, i.e., $\dim^*(P_0) = \dim^*(P_1)$. Considering the duality properties of $\dim^*$, we may derive that this principal holds in the case that $\boldsymbol{e}$ is the maximum of P_i, too. □

Proposition 4.23 *Given a poset P, a thin chain C of P and $\boldsymbol{e} \in C$, then* $\dim^*(P) = \dim^*(P \setminus \{\boldsymbol{e}\})$.

Proof Let $P_1 = (X, \leq_{P_1})$, P_0 be two posets, C a thin chain of P_1, $\boldsymbol{e}_1, \boldsymbol{e}_2 \in C$ such that $\boldsymbol{e}_1 \prec_{P_1} \boldsymbol{e}_2$, P_0 is the order induced by P_1 on $X \setminus \boldsymbol{e}_1$, and (L, f) dim* realization of P_0. We may produce the valid dim* realization (L', f') of P_1 as follows:

$$f'(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{e}_2), & \boldsymbol{x} = \boldsymbol{e}_1, \\ f(\boldsymbol{x}), & \boldsymbol{x} <_{P_1} \boldsymbol{e}_1, \\ f(\boldsymbol{x}) \cup \{\alpha\}, & \boldsymbol{x} \not\leq_{P_1} \boldsymbol{e}_1. \end{cases}$$

Similar to the previous proposition, we assume that $\alpha \in L$ and therefore $L' = L$. The poset $P_1 \setminus \{\boldsymbol{e}_1\}$ is equivalent to $P_1 \setminus \{\boldsymbol{e}_2\}$, and we get $\dim^*(P_1) = \dim^*(P_1 \setminus \{\boldsymbol{e}_1\}) = \dim^*(P_1 \setminus \{\boldsymbol{e}_2\})$. □

It is noteworthy that these two propositions fit well with the concept behind the matrix abduction. That is, inserting a maximum/minimum element or placing an element within a thin chain is equivalent to introducing information that already abides by the existing values, and the value of dim* reflects this. Moreover, note that using α labels, which allows giving realizations a kind of direction, is necessary for these propositions to hold (unlike $\dim_2$, for which they clearly do not).

5 Greedy Algorithms for Evaluating dim*

At this point, it is still hard to characterize the way labels are assigned in any good approximation of minimal dim* realizations, and it is still unknown what is the best approximation ratio that is feasible to achieve using a deterministic tractable algorithm (although such results exist for the case of $\dim_2$ [6] and dim [7]). In this section, we introduce an improvement of the existing greedy algorithm that was originally presented in [1]. Even though the new algorithm still does not provide a satisfying solution for all posets, it helps in identifying useful heuristics.

5.1 *Improved Algorithm for Finding* dim* *Realizations*

The new algorithm separates a given poset into levels and then processes the elements of each level separately, similar to the existing algorithm, and utilizes a set of heuristics in order to choose at each step the preferred assignment of labels. We assign each processed element with the minimal set of labels that contains the encodings of its predecessors, then insert additional labels as needed in order to solve every conflict in realization that this assignment has produced. In order to keep the implementation simple, we restrict ourselves to never remove any of the previously assigned labels, and instead we assign additional labels up to the point the encoding implies the correct order. Two questions remain open and are answered by heuristic functions:

- Which labels should be assigned in order to solve a conflict, given that multiple assignment combinations are available, considering the labels allocated up to that point?
- In what order the elements of the same level should be processed? (Since the amount of assignment combinations dwindles as more elements are processed.)

The last consideration is the amount of preallocated labels when the process begins. The assignment combinations available for each element depends on the amount of the already allocated labels; therefore, the initial amount of preallocated labels may significantly affect the encoding given to elements in the first few levels and consequently affect the complete realization. A lower bound of $\dim^*$ may serve as a heuristic for deciding how many labels to preallocate, but in many cases it would be better to retry generating the realization several times, each time preallocating a different amount of labels (and allocating additional labels as needed).

The algorithm may be described with two procedures, as defined in Fig. 7. The procedure `produce_realization`, given poset P, tries at each iteration to find a realization f, after preallocating d nonalpha labels. The heuristic h_1 chooses an assignment combination out of a given list, and the heuristic h_2 chooses the next element to process.

The procedure `process_poset` iterates through every element $\boldsymbol{x} \in P$ and assigns $f(\boldsymbol{x}) \subseteq L$. If the labels in L do not suffice in order to assign a valid encoding, new labels are inserted into L as needed.

The three *resolve* procedures are difficult to outline, but simple to describe. Each of the procedures generates every minimal assignment combination that may solve the conflict, under the principal that the encoding of each element only needs either more α label copies than the other, or exclusively contain any nonalpha label. The difference between the three is the additional constraints that affect it. For example, the procedure `resolve_containing_incomparable` takes advantage of the fact that the element $\boldsymbol{a}$ was not yet processed, and therefore may be assigned with additional labels freely. On the other hand, `resolve_contained_incomparable` generally has a more limited set of options, as changing the encoding of a previously processed element will also affect its successors and more likely to produce new conflicts. In order to avoid this, the algorithms also manages a list of constraints which must not be undone (also omitted from the presented outline). Out of the valid assignment combinations, every assignment and its effect on successors is evaluated using h_1, and the preferred one is applied by updating L and f accordingly.

5.2 *Heuristics Implementations for the Algorithm*

In this section, we discuss heuristics that may be used in the new algorithm – one implementation for h_2 and several implementations for h_1.

Generally speaking, heuristics that implement h_2 are given the elements of a level before it is processed, and decide the priority of how carefully each element should be encoded compared to others in the same level. Heuristics that implement h_1 are given a list of valid assignments; the effect of applying each of the assignments is simulated,

```
produce_realisation(P, h1, h2, f)
d ← 0
f ← ∅
while f = ∅
  L ← {(|P|)α, β1, β2 ... βd}
  f ← process_poset(P, h1, h2, f, L)
  d ← d + 1
  return f

process_poset(P, h1, h2, f, L)
  for each level 1 < n ≤ h(P) + 1
    let A be the set of all elements of level n, and let D = ∪_{1≤i<n} A_i
    for each a ∈ A :
      f(a) ← ϕ
      for each p ∈ pred_P(a)
        f(a) ← f(a) ∪ f(p)
      for each p ∈ D :
        if p <_P a ∧ f(p) ⊄ f(a) then
          f(a) ← f(a) ∪ {α}
        if a and p are incomparable, but f(a) ⊆ f(p) then
          (f, L) ← resolve_containing_incomparable(P, a, p, f, L, h1)
 (At this point, f(p) necessarily contains the encoding of every element
  it needs to, and is not contained by the encoding of any element in D)
      for each p ∈ D:
        if a and p are incomparable, but f(p) ⊆ f(a) then
          (f, L) ← resolve_contained_incomparable(P, a, p, f, L, h1)
while ≠ ϕ :
  a ← h2(P, f, A)
  A ← A\{a}
for each x ∈ A :
    if f(a) ⊂ f(x) ∨ f(x) ⊂ f(a) then
      (f, L) ← resolve_incomparables(P, a, x, f, L, h1)
```

Fig. 7 Procedures `produce_realization` and `process_poset`

and the heuristic chooses the preferable simulated state. Note that these heuristics may be given weights and combined in order to achieve a better evaluation.

Heuristic 1 (Preserve Encoding Size, as h_2) Given the antichain A and its incomplete realization f', we define the heuristic by $h_2(A) = \operatorname{argmin}_{a \in A} |f'(a)|$. That is, we assume that if an element is currently encoded using less nonalpha labels than others in the same antichain, then it should be processed earlier, with the underlying assumption that it is highly preferable to keep encoding the element by using more α label copies instead of introducing additional nonalpha labels (and therefore this option must be available). While this heuristic may be unintuitive, it proved to be the most effective of the ones tested, since it allows us to preserve preferred conflict resolutions from previous levels, which implicitly indicated that assigning the element additional nonalpha labels is undesired. Interestingly, this method also resembles the minimal encoding of a single antichain, where the element with the most α label copies is encoded by using nonalpha labels, while the element with least α label copies may be encoded by every available label.

Heuristic 2 (Balance Distribution of Non-alpha Labels, as $\boldsymbol{h_1}$) Given the poset P, the incomplete realization f', and the set of labels it uses (excluding α labels) denoted L, we define:

$$k(f, l, \boldsymbol{x}) = \begin{cases} 1, & l \in f(\boldsymbol{x}), \\ 0, & l \notin f(\boldsymbol{x}), \end{cases}$$

$$h_1(L, f', P) = -\mathrm{Var}\Big(\sum_{l \in L}\Big(\sum_{\boldsymbol{x} \in P} k(f', l, \boldsymbol{x})\Big)\Big),$$

i.e., each assignment is evaluated by minus the variance of the presence of every non-alpha label in the encoding. These criteria encourages encodings that utilize as many combinations as possible using the allocated labels. To see this, consider again the optimal encoding of an antichain, in which the variance is maximized.

Heuristic 3 (Minimize Label References, as $\boldsymbol{h_1}$) In this heuristic, we simply minimize the sum of nonalpha label assignments. While this may seem as a second attempt to achieve balance in the amount of label references in the encoding, it encourages instead assigning nonalpha labels to shorter paths. Unfortunately, this also encourages the avoidance of inevitable assignments equally to truly undesirable assignments since some elements simply must be encoded by many nonalpha labels. In order to fix this, an additional prioritization mechanism is used. Given the poset P, the realization f' and the estimated weight of each element w, we define the heuristic by:

$$h_1(f', P, w) = -\sum_{\boldsymbol{x} \in P} \frac{|f'(\boldsymbol{x})|}{w(\boldsymbol{x})}.$$

In our implementation, we have calculated the size of the maximal antichain below every element, then derived a rough estimation of the minimal amount of nonalpha labels which must be assigned to it as the weight of the element. The width of the poset may be evaluated using bipartite graph matching, for which exist efficient algorithms such as Hopcroft–Karp algorithm [5, 9]. Another method for bounding $\dim^*$ of a suborder, albeit much slower, is using a dynamic programming algorithm which attempts to find the minimal encoding of the poset that stems from each element separately.

It is noteworthy that when considering the weights, this heuristic encourages the assignment of nonalpha labels to shorter paths which consist of elements with higher weight, which highly resembles the use of maximal thinnest path in the first algorithm, and may be considered as its generalization.

Heuristic 4 (Mimic $\mathbf{dim_2}$, as $\boldsymbol{h_1}$) In this heuristic, we demonstrate the option of designing a heuristic that relies on existing algorithms. The evaluation of $\dim_2$ appears to be a good candidate, since it is simple to approximate and shares a lot of properties with $\dim^*$ (consider antichain and ordered antichains encodings, one-point removal theorem, etc.). One issue that must be addressed specifically is the extreme case of encoding chains, which demonstrates the lower bound of $\dim^*$ (as mentioned before), but also demonstrates the upper bound of $\dim_2$, since for a chain C of height $n-1$ it holds that $\dim_2(C)=n$. While thin chains within a poset may be eliminated, as previously mentioned, the difference makes it hard to estimate the similarity between $\dim^*$ and $\dim_2$ for more complex posets.

There are many known algorithms that evaluate $\dim_2$, where some target-specific poset classes (see [4, 6] for two of the most well-known results) and some offer a solution for the general case (see [3, 12]). Most algorithms for the general case rely, as described in [12], on a different kind of realization called *simple encoding*. That is, given a poset, the algorithms transform it into a graph by selecting certain critical elements, then approximate the simple encoding of the graph using graph coloring algorithms, and finally translate the simple encoding into a $\dim_2$ realization for the original poset. Unfortunately, the simple encoding, similar to $\dim_2$ realizations, relies on a single type of labels and apparently there is no sensible way of reusing it for $\dim^*$. Considering this limitation, we attempt to mimic the result of the existing algorithms indirectly. Given the poset P, the incomplete $\dim^*$ realization f' and the complete $\dim_2$ realization g, we define

$$h_1=\sum_{\substack{\boldsymbol{x},\boldsymbol{y}\in P\\ \boldsymbol{x}>_P\boldsymbol{y}}}\left(\frac{|f'(y)|}{|f'(x)|}-\frac{|g(y)|}{|g(x)|}\right),$$

i.e., the heuristic encourages assignments with similar length ratio in both realizations, between every two comparable elements (empty encodings are neglected). To see the reason to exclude the ratio between incomparable elements, consider the optimal encoding of an antichain. For an antichain A of width $n=2^b$, in the case of $\dim^*$ half of the elements are encoded using $\dim^*(A)/2$ nonalpha labels, while in the case of $\dim_2$, every element in the optimal encoding is encoded by using $\dim_2(A)/2$ labels. Therefore, in cases such as this, the heuristic provides a poor evaluation. Additionally, while the heuristic hints which elements should be assigned by more nonalpha labels than others, it gives no indication which specific nonalpha label should be assigned. However, unlike the heuristics presented previously, in this method we may identify and transfer difficult combinatorial decisions which allow minimizing the amount of needed labels in both realizations. In conclusion, we believe that this heuristic might be useful when paired with another heuristic, and divert otherwise neutral decisions (note that this heuristic was yet to be implemented and tested).

5.3 Result Comparison

Table 3 gives a comparison between the existing algorithms (the new algorithm, the previous one and exhaustive scan). The values presented are the average $\dim^*$ as evaluated by the algorithm as a function of the poset size. Note that for the exhaustive scan, several

Table 3
Result comparison

Poset size	Old	New	Exhaustive
5	2.76	2.40	2.40
6	3.64	3.17	3.00
7	4.39	3.61	3.33
8	5.20	3.92	3.63
9	6.28	4.58	3.90
10	6.82	4.92	4.12
11	7.48	5.33	4.42
12	8.42	5.88	4.68
13	9.18	6.20	4.70

posets were omitted from the average since the completion of the procedure took too long (therefore the data presented is slightly biased). In the new algorithm, heuristics 1, 2, and 3 were used (2 and 3 were given the same weight).

6 Conclusion and Further Work

Even though the new algorithm provides significantly better results than the previous one, we believe that there is still much room for improvement, both in evaluating $\dim^*$ and in the evaluation of matrix abduction in general. One straightforward direction is to examine and implement additional heuristics which may be used by the new algorithm, in order to further characterize the problem. Another possible direction would be to find additional ways of reusing existing algorithms for evaluating $\dim_2$, by efficiently separating the assignment of α labels and nonalpha labels (e.g., by first assigning α labels to every element in the poset using a local-search algorithm, then adapting the input for an existing $\dim_2$ evaluation algorithm). Additionally, methods of evaluating the matrix abduction method without explicitly evaluating $\dim^*$ should also be considered (e.g., by identifying hints telling which of two given posets is preferable for the abduction). Finally, the approximation ratio of the existing algorithms, and what approximation ratio is possible to achieve at all, is left open.

It is noteworthy that aside from its uses in matrix abduction, $\dim^*$ may be used as an alternative practical way for encoding hierarchies. Using $\dim_2$ realizations (bit-vector encodings) in order to encode hierarchical structures such as class objects (for object oriented programming environments) is one of the established methods, since it typically produces compact encodings and allows implementing very quick class inclusion checks. Unfortunately, as previously mentioned, $\dim_2$ realizations encode chains very poorly. On the other hand, $\dim^*$ realizations encode chains much more efficiently, and for antichains, the amount of bits needed in both realizations is almost identical. That is, for an antichain of size n, bit-vector encoding uses [6]: $\log_2(n) + \log_2\log_2(n)/2$, while $\dim^*$ realizations use $\log_2(n)$ nonalpha labels (a bit per label) and α label counting up to $\log_2(n)$, i.e., additional $\log_2\log_2(n)$ bits, which is almost on par. Therefore, finding an efficient algorithm that provides a good approximation for $\dim^*$ may be useful even by itself.

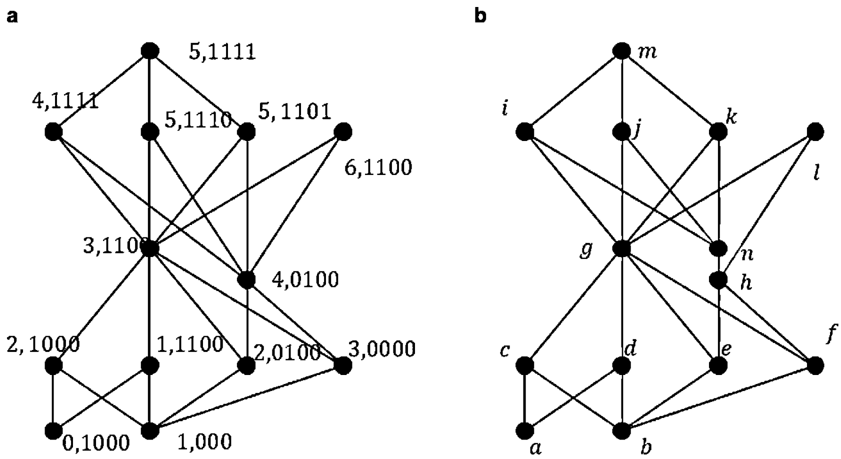


Fig. 8 A poset with dim* of 5 (**a**) and a poset with dim* of 7 (**b**)

7 Appendix

Proof of Proposition 21. It is easy to see that the proof used for the one-point removal theorem for $\dim_2$ may be also used as is for the case of $\dim^*$ in order to derive a similar bound. In order to see that no preferable method does exist for the case of $\dim^*$ in the general case, consider the example poset $E \backslash \{\boldsymbol{n}\} = (\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \ldots, \boldsymbol{m}\}, \leq_E)$, as illustrated in Fig. 8a. We show that for the insertion of a certain new element $\boldsymbol{n}$ it holds that $\dim^*(E) = \dim^*(E \setminus \{\boldsymbol{n}\}) + 2$. As illustrated in the same figure, there exists a realization of $E \setminus \{\boldsymbol{n}\}$ which uses multiple α labels and additional four nonalpha labels. Therefore, $\dim^*(E \setminus \{\boldsymbol{n}\}) \leq 5$ holds and we only need to prove that $\dim^*(E) \geq 7$. Our implementation of the exhaustive scan algorithm may not be used to prove this since it is impractically slow; instead, we give a detailed explanation which depends on two constraints that are proved below.

Constraint 1 For the poset E and any encoding of λ of E it holds that $|\lambda(\boldsymbol{g}) \cup \lambda(\boldsymbol{h})| \geq 4$.

Proof Since $\boldsymbol{g} \geq_E \{\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{f}\}$, then if $|\lambda(\boldsymbol{c}) \cup \lambda(\boldsymbol{d}) \cup \lambda(\boldsymbol{e}) \cup \lambda(\boldsymbol{f})| \geq 4$, it also holds that $|\lambda(\boldsymbol{g}) \cup \lambda(\boldsymbol{h})| \geq 4$. Otherwise, since $\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{f}$ make an antichain of size 4, then $\boldsymbol{g}$ must be encoded using at least two nonalpha labels, i.e., $|\lambda(\boldsymbol{g})| \geq 3$. Let $\lambda(\boldsymbol{g}) \supseteq \{(x_1)\alpha, \beta_1, \beta_2\}$. The encoding combinations that may be used for $\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{f}$ are:

- $\{(y_1)\alpha, \beta_1, \beta_2\}$,
- $\{(y_2)\alpha, \beta_1\}$,
- $\{(y_3)\alpha, \beta_2\}$,
- $\{(y_4)\alpha\}$,

where $y_1 < y_2, y_3 < y_4$. Note that either $\boldsymbol{c}$ or $\boldsymbol{d}$ is encoded by $\{(y_1)\alpha, \beta_1, \beta_2\}$, and either $\boldsymbol{e}$ or $\boldsymbol{f}$ is encoded by $\{(y_4)\alpha\}$, since both $\boldsymbol{c}$ and $\boldsymbol{d}$ are the successors of the antichain

$\{\boldsymbol{a}, \boldsymbol{b}\}$. Clearly, the antichain must be encoded using at least one nonalpha label γ, and consequently $\gamma \in \lambda(\boldsymbol{c}) \cap \lambda(\boldsymbol{d})$, which may be possible only if it holds that $\lambda(\boldsymbol{c}) \neq \{(y_4)\alpha\}$, $\lambda(\boldsymbol{d}) \neq \{(y_4)\alpha\}$ and either $\lambda(\boldsymbol{c}) = \{(y_1)\alpha, \beta_1, \beta_2\}$ or $\lambda(\boldsymbol{d}) = \{(y_1)\alpha, \beta_1, \beta_2\}$ holds. As a result it also holds that $\lambda(\boldsymbol{c}) = \{(y_4)\alpha\}$ or $\lambda(\boldsymbol{d}) = \{(y_4)\alpha\}$. If we assume that $|\lambda(\boldsymbol{g}) \cup \lambda(\boldsymbol{h})| < 4$, then since $\boldsymbol{g}$ and $\boldsymbol{h}$ are incomparable, and $|\lambda(\boldsymbol{g})| \geq 3$, it must hold that $\lambda_\alpha(\boldsymbol{h}) > \lambda_\alpha(\boldsymbol{g})$ and consequently, $\lambda(\boldsymbol{n}) \supseteq (\lambda(\boldsymbol{c}) \cup \lambda(\boldsymbol{d}) \cup \lambda(\boldsymbol{h})) \supseteq \{(\lambda_\alpha(\boldsymbol{h}))\alpha, \beta_1, \beta_2\} \supseteq \lambda(\boldsymbol{g})$, but this is not possible, since $\boldsymbol{n}$ and $\boldsymbol{g}$ are incomparable. □

Constraint 2 For the poset E and any encoding λ of E, if $|\lambda(\boldsymbol{m}) \cup \lambda(\boldsymbol{l})| < 7$, then $\lambda_\alpha(\boldsymbol{l}) > \lambda_\alpha(\boldsymbol{m})$ holds.

Proof Let $L_1 = \lambda(\boldsymbol{m}) \cup \lambda(\boldsymbol{l})$ and $L_2 = \lambda(\boldsymbol{g}) \cup \lambda(\boldsymbol{h}) \supseteq \{\beta_1, \beta_2, \beta_3\}$. Since $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}, \boldsymbol{l}$ make an antichain of size 4, and $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}, \boldsymbol{l}\} \geq_E \{\boldsymbol{g}, \boldsymbol{h}\}$, then the nonalpha labels $\beta_4, \beta_5 \in L_1 \setminus L_2$ must be modified to encode $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}, \boldsymbol{l}$, and the encoding combinations that may be used are:

- $L_2 \cup \{(y_5)\alpha, \beta_4, \beta_5\}$,
- $L_2 \cup \{(y_6)\alpha, \beta_4\}$,
- $L_2 \cup \{(y_7)\alpha, \beta_5\}$,
- $L_2 \cup \{(y_8)\alpha\}$,

where $y_5 < y_6, y_7 < y_8$. Consequently, it holds that $\lambda(\boldsymbol{m}) \supseteq (\lambda(\boldsymbol{i}) \cup \lambda(\boldsymbol{j}) \cup \lambda(\boldsymbol{k})) \supseteq (L_2 \cup \{\beta_4, \beta_5\})$, and since $\boldsymbol{m}$ and $\boldsymbol{l}$ are incomparable, then $\lambda_\alpha(\boldsymbol{l}) > \lambda_\alpha(\boldsymbol{m})$. It is easy to see that using additional nonalpha labels is not possible if $|\lambda(\boldsymbol{m}) \cup \lambda(\boldsymbol{l})| < 7$, and so we are done. □

Let λ be a label minimal encoding of E. According to the constraints above, if $|L_1| < 7$ then it holds that $\lambda_\alpha(\boldsymbol{l}) > \lambda_\alpha(\boldsymbol{m}) > \lambda_\alpha(\boldsymbol{n})$. Since $\boldsymbol{n}$ and $\boldsymbol{l}$ are incomparable, there must be a nonalpha label $\gamma \in (\lambda(\boldsymbol{n}) \setminus \lambda(\boldsymbol{l})) \subseteq (L_1 \setminus L_2) = \{\beta_4, \beta_5\}$, but since $(\lambda(\boldsymbol{i}) \cap \lambda(\boldsymbol{j}) \cap \lambda(\boldsymbol{k})) \supseteq \lambda(\boldsymbol{n})$, and $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ make an antichain of size 3, then the labels β_4, β_5 may not be modified in order to encode the elements. Therefore, an additional nonalpha label must be used, and for any valid encoding it holds that $|L_1| \geq 7$. □

References

1. Abraham, M., Gabbay, D.M., Schild, U.: Analysis of the Talmudic argumentum a fortiori inference rule (Kal Vachomer) using matrix abduction. Stud. Log. **92**(3), 281–364 (2009)
2. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. Found. Comput. Math. **9**(6), 717–772 (2009)
3. Caseau, Y., Habib, M., Nourine, L., Raynaud, O.: Encoding of multiple inheritance hierarchies and partial orders. Comput. Intell. **15**(1), 50–62 (1999)
4. Filman, R.E.: Polychotomic encoding: a better quasi-optimal bit-vector encoding of tree hierarchies. In ECOOP 2002 – Object-oriented Programming, pp. 545–561. Springer, Berlin, Heidelberg (2002)
5. Fulkerson, D.R.: Note on Dilworth's decomposition theorem for partially ordered sets. Proc. Am. Math. Soc. **7**, 701–702 (1956)
6. Habib, M., Nourine, L., Raynaud, O., Thierry, E.: Computational aspects of the two dimension of partially ordered sets. Theor. Comput. Sci. **312**(2), 401–431 (2004)

7. Hegde, R., Jain, K.: The hardness of approximating poset dimension. Electron. Notes Discret. Math. **29**, 435–443 (2007)
8. Hill, A.: Reasoning by analogy in inductive logic. Dissertation, University of Manchester (2013)
9. Hopcroft, J.E., Karp, R.M.: An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs. SIAM J. Comput. **2**(4), 225–231 (1973)
10. Johnson, C.R.: Matrix completion problems: a survey. In: Proceedings of Symposia in Applied Mathematics, vol. 40, pp. 171–198. American Mathematical Society (1990)
11. Karp, R.M.: Reducibility among combinatorial problems. In: Complexity of Computer Computations, pp. 85–103. Springer, New York (1972)
12. Krall, A., Vitek, J., Horspool, R.N.: Near optimal hierarchical encoding of types. In: ECOOP'97 – Object-oriented Programming, pp. 128–145. Springer, Berlin, Heidelberg (1997)
13. Laufer, A., Gabbay, D.M.: Topological aspects of matrix abduction 1. In: Koslow, A., Buchsbaum, A. (eds.) The Road to Universal Logic – Festschrift for the 50th Birthday of Jean-Yves Beziau, Volume II. Birkhäuser, Basel (2015)
14. Prade, H., Richard, G.: Homogeneous logical proportions: their uniqueness and their role in similarity-based prediction. Thirteenth International Conference on the Principles of Knowledge Representation and Reasoning, 2012
15. Raynaud, O., Thierry, E.: The complexity of embedding orders into small products of chains. Order **27**(3), 365–381 (2010)
16. Ribeiro, H.J. (ed.): Systematic Approaches to Argument by Analogy. Springer, Berlin (2014)
17. Schlechta, K.: Karl's remarks on Kal-algorithm. Personal communication (2011)
18. Trotter, W.T.: Combinatorics and Partially Ordered Sets. The John Hopkins University Press (1992)

M. Ben Adar Bessos (✉)
Bar-Ilan University, Ramat Gan, Israel
e-mail: mai.bessos@gmail.com

D. M. Gabbay
Department of Computer Science, King's College, London, UK
Bar-Ilan University, Ramat Gan, Israel

Deciding Theoremhood in Fibred Logics Without Shared Connectives

Sérgio Marcelino, Carlos Caleiro, and Pedro Baltazar

Abstract Fibring is a powerful mechanism for combining logics, and an essential tool for designing and understanding complex logical systems. Abstract results about the semantics and proof theory of fibered logics have been extensively developed, including general soundness and completeness preservation results. Decidability, however, a key ingredient for the automated support of the fibered logic, has not deserved similar attention.

In this chapter, we address the problem of deciding theoremhood in fibered logics without shared connectives. Namely, under this assumption, we provide a full characterization of the mixed patterns of reasoning that leads to theorems in the fibered logic, and uses it to prove a general decidability preservation result. The complexity of the decision procedure we obtain is also analyzed.

Keywords Combined logics · Fibring, theorem · Decidability · Complexity

Mathematics Subject Classification (2000) 03B22: Abstract deductive systems · 03B25: Decidability of theories and sets of sentences · 03B62: Combined logics

1 Introduction

Fibring is a powerful and appealing mechanism for combining logics, a valuable tool for the construction and analysis of complex logics, and thus a key ingredient of the general theory of universal logic [2, 3]. As first proposed by Dov Gabbay in [11, 12], given two logics $\mathcal{L}_1$ and $\mathcal{L}_2$, fibring should combine $\mathcal{L}_1$ and $\mathcal{L}_2$ into the smallest logical system for the combined language which is a conservative extension of both $\mathcal{L}_1$ and $\mathcal{L}_2$. However, it is not hard to see that a conservative extension of the given logics may not always exist. Still, this circumstance does not necessarily imply that the construction is meaningless, as one can then aim at being "as conservative as possible." This idea has led to the study of the fibring operation as yielding the smallest logic that extends $\mathcal{L}_1$ and $\mathcal{L}_2$, without worrying about conservativity [8]. It is worth mentioning that the characterization of conservativity in this context is a problem that remains highly unexplored.

Despite the depth of the track of work on fibered logics that led to a substantial understanding of their semantics and proof theory, including very general soundness and completeness preservation results (see [1, 5, 9, 14–16, 18], *inter alia*), the question of

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_18

decidability has not been satisfactorily addressed. The only general result related to, but markedly distinct from, decidability for fibered logics is [10], where the preservation by fibring of the semantic notion of *finite model property* is studied.

In this chapter, we focus on the decision problem for theorems of fibered logics, in the case when the logics being combined do not share any connectives. We manage to give a full characterization of the mixed patterns of reasoning that lead to the proofs of theorems in the fibered logic and, as a result, we obtain a general decidability preservation result for theoremhood, the first of this kind. We also analyze the complexity of the decision procedure obtained.

In Sect. 2 we recall the notions and results needed throughout the chapter, namely about fibered logics, and introduce some useful notation. In Sect. 3 we illustrate the difficulties involved in deciding theoremhood in fibered logics without shared connectives, and provide a thorough analysis of the mixed patterns of reasoning that may occur in the combined logic. Finally, in Sect. 4 we prove our main result: a decidability preservation result for theoremhood in fibered logics with no shared connectives, and a characterization of the complexity of the decision procedure obtained. We conclude, in Sect. 5, with an assessment of the results obtained and paths to pursue in future work.

2 Definitions

In this section, we recall the essential concepts that we are dealing with in this chapter, namely fibring, and introduce some useful notions and notations.

2.1 (Trans)finite Sequences

Along the chapter, we will need to deal with (not necessarily finite) sequences of objects. Let A be a set (of objects). Given an ordinal η, we use $\overline{a} = \langle a_\kappa \rangle_{\kappa<\eta}$ to denote an η-long sequence of elements of A, or simply an η-sequence, understood as a function from $\{\kappa : \kappa < \eta\}$ to A. The η-sequence $\overline{a}$ is said to be *injective* precisely when it is injective as a function, that is, when $a_i \neq a_j$ for all $i, j < \eta$ with $i \neq j$. As usual, if $\tau \leq \eta$, the sequence $\langle a_\kappa \rangle_{\kappa<\tau}$ is dubbed a *prefix* of $\overline{a}$.

Note that when η is a limit ordinal, an η-sequence does not have a last element. On the contrary, if η is a successor ordinal, and in particular a finite ordinal, then an η-sequence $\overline{a}$ can be understood as $a_0, a_1, \ldots, a_{\eta-1}$, and may also be represented by $\langle a_\kappa \rangle_{\kappa\leq\eta-1}$. The 0-sequence (*empty* sequence) is simply not represented.

2.2 Syntax

A *signature* is an $\mathbb{N}_0$-indexed family $\Sigma = \{\Sigma_n\}_{n\in\mathbb{N}_0}$ of sets. The elements of Σ_n are dubbed n-place *connectives*. Being indexed families of sets, the usual set-theoretic notions can be smoothly extended to signatures. We will sometimes abuse notation, and confuse Σ with the set $(\biguplus_{n\in\mathbb{N}_0} \Sigma_n)$ of all its connectives, and write $c \in \Sigma$ when c is some n-place connective and $c \in \Sigma_n$. For this reason, the *empty signature*, with no connectives at all, is simply denoted by $\emptyset$.

Let Σ, Σ' be two signatures. We say that Σ is a *subsignature* of Σ', and write $\Sigma \subseteq \Sigma'$, whenever $\Sigma_n \subseteq \Sigma'_n$ for every $n \in \mathbb{N}_0$. Expectedly, we can also define the *intersection* $\Sigma \cap \Sigma' = \{\Sigma_n \cap \Sigma'_n\}_{n\in\mathbb{N}_0}$, *union* $\Sigma \cup \Sigma' = \{\Sigma_n \cup \Sigma'_n\}_{n\in\mathbb{N}_0}$, and *difference* $\Sigma' \setminus \Sigma = \{\Sigma'_n \setminus \Sigma_n\}_{n\in\mathbb{N}_0}$ of signatures. Clearly, $\Sigma \cap \Sigma'$ is the largest subsignature of both Σ and Σ', and contains the connectives *shared* by Σ_1 and Σ_2. When there are no shared connectives, we have that $\Sigma \cap \Sigma' = \emptyset$. Analogously, $\Sigma \cup \Sigma'$ is the smallest signature that has both Σ and Σ' as subsignatures, and features all the connectives from both Σ and Σ' in a *combined signature*. Furthermore, $\Sigma' \setminus \Sigma$ is the largest subsignature of Σ' which does not share any connectives with Σ.

Given a signature Σ and a set P of *variables*, the generated set of *formulas* is the carrier set $L_\Sigma(P)$ of the free Σ-algebra generated by P.

As usual, we define the *size* of a formula as given by the function size such that $\mathsf{size}(p) = 1$ if $p \in P$, and $\mathsf{size}(c(\varphi_1, \dots, \varphi_n)) = 1 + \sum_{i=1}^{n} \mathsf{size}(\varphi_i)$. For $\Gamma \subseteq L_\Sigma(P)$, we further define $\mathsf{size}(\Gamma) = \sum_{\varphi\in\Gamma} \mathsf{size}(\varphi)$.

If $\varphi \in L_\Sigma(P)$ then we define the *head of* φ to be either $\mathsf{head}(\varphi) = p$ when $\varphi = p \in P$, or $\mathsf{head}(\varphi) = c$ when $\varphi = c(\varphi_1, \dots, \varphi_n)$ for formulas $\varphi_1, \dots, \varphi_n \in L_\Sigma(P)$ and $c \in \Sigma_n$. Clearly, if $\Sigma \subseteq \Sigma'$ and $P \subseteq P'$ then $L_\Sigma(P) \subseteq L_{\Sigma'}(P')$. Of course, given $\psi \in L_{\Sigma'}(P')$, $\mathsf{head}(\psi)$ may not be in Σ nor P.

We also define the *set of variables occurring in* φ to be either $\mathsf{var}(\varphi) = \{p\}$ when $\varphi = p \in P$, or $\mathsf{var}(\varphi) = \bigcup_{i=1}^{n} \mathsf{var}(\varphi_n)$ when $\varphi = c(\varphi_1, \dots, \varphi_n)$ for formulas $\varphi_1, \dots, \varphi_n \in L_\Sigma(P)$ and $c \in \Sigma_n$.

In the following, we shall assume that signatures are countable and sets of variables are denumerable. We assume a denumerable set P of variables fixed. If Σ is a countable signature then $L_\Sigma(P)$ is clearly denumerable.

Let $\Sigma \subseteq \Sigma'$ be signatures. We shall call a Σ*-monolith of* $\psi \in L_{\Sigma'}(P)$ to any outermost subformula of ψ whose head is in $\Sigma' \setminus \Sigma$. The set $\mathsf{Mon}_\Sigma(\psi)$ of all Σ-monoliths of ψ is defined as follows:

$$\mathsf{Mon}_\Sigma(\psi) = \begin{cases} \emptyset & \text{if } \psi \in P, \\ \bigcup_{i=1}^{n} \mathsf{Mon}_\Sigma(\psi_i) & \text{if } \psi = c(\psi_1, \dots, \psi_n) \text{ and } c \in \Sigma_n, \\ \{\psi\}, & \text{otherwise.} \end{cases}$$

We extend the notation also to sets of formulas, using $\mathsf{Mon}_\Sigma(\Delta)$ to denote $\bigcup_{\psi\in\Delta} \mathsf{Mon}_\Sigma(\psi)$, given $\Delta \subseteq L_{\Sigma'}(P)$. Clearly, if $\Gamma \subseteq L_\Sigma(P)$ then $\mathsf{Mon}_\Sigma(\Gamma) = \emptyset$.

We shall now consider a reasonable way of defining the perspective, from the point of view of Σ, that one may have of a formula in a given context $\Delta \subseteq L_{\Sigma'}(P)$ (which will most often be left implicit). For the purpose, we shall consider a (bijective) enumeration $e : \{n \in \mathbb{N} : 1 \leq n \leq |\mathsf{Mon}_\Sigma(\Delta)|\} \to \mathsf{Mon}_\Sigma(\Delta)$ of the relevant Σ-monoliths, and use a denumerable set of additional propositional variables $X = \{x_n : n \in \mathbb{N}_0\}$, disjoint from P. We define the function $\mathsf{skel}_\Sigma : \Delta \to L_\Sigma(P \cup X)$ as follows:

$$\mathsf{skel}_\Sigma(\psi) = \begin{cases} \psi & \text{if } \psi \in P, \\ c(\mathsf{skel}_\Sigma(\psi_1), \ldots, \mathsf{skel}_\Sigma(\psi_n)) & \text{if } \psi = c(\psi_1, \ldots, \psi_n) \text{ and } c \in \Sigma_n, \\ x_{e^{-1}(\psi)}, & \text{otherwise.} \end{cases}$$

We call $\mathsf{skel}_\Sigma(\psi)$ the *Σ-skeleton of ψ*. Clearly, $\mathsf{skel}_\Sigma(\psi)$ is obtained from ψ by substituting each of its Σ-monoliths ψ' by the variable x_n such that $e(n) = \psi'$.

Let $\Sigma \subseteq \Sigma'$ and $P \subseteq P'$. A *Σ'-substitution* is a function $\sigma : P \to L_{\Sigma'}(P')$, which extends freely to a function $\sigma : L_\Sigma(P) \to L_{\Sigma'}(P')$. Given a formula $\varphi \in L_\Sigma(P)$, $\sigma(\varphi)$ is the *instance* of φ by σ, sometimes denoted simply by φ^σ, and is the result of uniformly replacing each variable $p \in P$ occurring in φ by $\sigma(p)$. When $\Gamma \subseteq L_\Sigma(P)$ we use Γ^σ to denote $\sigma[\Gamma] = \{\varphi^\sigma : \varphi \in \Gamma\}$.

Given a context $\Delta \subseteq L_{\Sigma'}(P)$ and an enumeration e of $\mathsf{Mon}_\Sigma(\Delta)$, as well as $\varphi \in L_\Sigma(P \cup X)$ and a η-sequence $\overline{\alpha}$ of $L_{\Sigma'}(P)$ formulas, we will often write $\varphi[\overline{\alpha}]_\Sigma$ to denote the formula $\varphi^{\sigma(\overline{\alpha})}$ where $\sigma(\overline{\alpha}) : P \cup X \to L_{\Sigma'}(P)$ is such that $\sigma(\overline{\alpha})(p) = p$ if $p \in P$, $\sigma(\overline{\alpha})(x_n) = \alpha_{n-1}$ if $0 \le n-1 < \eta$, and $\sigma(\overline{\alpha})(x_n) = x_n$ otherwise. Hence, we will write $\mathsf{skel}_\Sigma(\psi)[\overline{e}]_\Sigma$ instead of ψ if we want to highlight the Σ-monoliths in the structure of ψ, where $\overline{e}$ is any sufficiently long prefix of the sequence $\langle e(\kappa)\rangle_{\kappa<|\mathsf{Mon}_\Sigma(\Delta)|}$.

Taking advantage of the notation, given two η-sequences $\overline{\alpha}$ and $\overline{\beta}$ of $L_{\Sigma'}(P)$ formulas, with $\overline{\alpha}$ injective, we will also write $\psi[\overline{\alpha}/\overline{\beta}]_\Sigma$ to denote the formula obtained by replacing each occurrence of α_i as a Σ-monolith of ψ by β_i, for all $i < \eta$. It is not difficult to check that $\psi[\overline{\alpha}/\overline{\beta}]_\Sigma = \mathsf{skel}_\Sigma(\psi)[\overline{\gamma}]_\Sigma$ where $\overline{\gamma}$'s length must be bigger than $e^{-1}(\alpha_\kappa)$ for all $\kappa < \eta$, with $\gamma_n = \beta_\kappa$ if $e(n+1) = \alpha_\kappa$, and $\gamma_n = e(n+1)$ if $e(n+1)$ does not occur in $\overline{\alpha}$.

These square bracket notations will be extended to sets of formulas in the obvious manner.

Example 2.1 Let Σ be the signature with exactly two connectives, a 0-place connective c and a 2-place connective g, that is, $\Sigma_0 = \{c\}$, $\Sigma_2 = \{g\}$ and $\Sigma_n = \emptyset$ for all $n \in \mathbb{N}_0 \setminus \{0, 2\}$. Let Σ' extend Σ with an additional 1-place connective f, that is, $\Sigma'_0 = \{c\}$, $\Sigma'_1 = \{f\}$, $\Sigma'_2 = \{g\}$ and $\Sigma'_n = \emptyset$ for all $n \in \mathbb{N}_0 \setminus \{0, 1, 2\}$.

Taking the $L'_\Sigma(P)$ formula $\psi = g(f(p), g(c, f(g(f(c), f(p)))))$ we have that

$$\mathsf{Mon}_\Sigma(\psi) = \{f(p), f(g(f(c), f(p)))\}.$$

Note, in particular, that the subformula $f(c)$ is not a Σ-monolith of ψ because if occurs inside the (outermost) monolith $f(g(f(c), f(p)))$. For the same reason, $f(p)$ is only a Σ-monolith of ψ because it also occurs outside $f(g(f(c), f(p)))$.

Hence, we have (in the appropriate context) that $\mathsf{skel}_\Sigma(\psi) = g(x_1, g(c, x_2))$, and thus

$$\psi = g(x_1, g(c, x_2))[f(p), f(g(f(c), f(p)))]_\Sigma \,.$$

Moreover,

$$\psi[f(p)/\beta]_\Sigma = g(\beta, g(c, f(g(f(c), f(p))))) \,,$$

noting that β only replaces the leftmost occurrence of $f(p)$ in ψ, where it is a Σ-monolith, leaving the second untouched.

2.3 Logical Consequence

A *logic (over signature Σ)* is a tuple $\mathcal{L} = \langle \Sigma, \_^{\vdash} \rangle$, where $\_^{\vdash} : 2^{L_\Sigma(P)} \to 2^{L_\Sigma(P)}$ is a consequence operator (see [17], for instance), that is, it satisfies the following properties:

$$\begin{array}{ll} \Gamma \subseteq \Gamma^{\vdash} & (\textit{extensiveness}) \\ \Gamma^{\vdash} \subseteq (\Gamma \cup \Delta)^{\vdash} & (\textit{monotonicity}) \\ (\Gamma^{\vdash})^{\vdash} \subseteq \Gamma^{\vdash} & (\textit{idempotence}) \\ (\Gamma^{\vdash})^{\sigma} \subseteq (\Gamma^{\sigma})^{\vdash} & (\textit{structurality}) \end{array}$$

for every $\Gamma, \Delta \subseteq L_\Sigma(P)$ and $\sigma : P \to L_\Sigma(P)$. Note that we do not require, in general, that the logic is *finitary*, i.e., it may happen that $\Gamma^{\vdash}$ properly contains the union of all $\Gamma_0^{\vdash}$ for finite $\Gamma_0 \subseteq \Gamma$.

As usual, we shall confuse the consequence operator with its induced Tarskian consequence relation. Thus, given $\varphi \in L_\Sigma(P)$, we will write $\Gamma \vdash \varphi$ whenever $\varphi \in \Gamma^{\vdash}$. When $\Gamma = \{\varphi_1, \ldots, \varphi_n\}$ is finite, we write $\varphi_1, \ldots, \varphi_n \vdash \varphi$ instead of $\{\varphi_1, \ldots, \varphi_n\} \vdash \varphi$. Moreover, as usual, if $\Gamma = \emptyset$ we write $\vdash \varphi$ instead of $\emptyset \vdash \varphi$, and dub φ a *theorem* of $\mathcal{L}$. We also write $\Gamma, \Delta \vdash \varphi$ instead of $\Gamma \cup \Delta \vdash \varphi$.

We shall call any $\Gamma \subseteq L_\Sigma(P)$ such that $\Gamma = \Gamma^{\vdash}$ a *theory* of $\mathcal{L}$, and denote the set of all theories of $\mathcal{L}$ by $\mathbf{Th}(\mathcal{L})$. It is well known that $\mathbf{Th}(\mathcal{L})$ constitutes a complete lattice under the inclusion ordering (see [17], for instance). The top theory of the lattice is $L_\Sigma(P)$, which is also called the *inconsistent* theory.

A logic $\mathcal{L} = \langle \Sigma, \_^{\vdash} \rangle$ is said to be *consistent* if $\emptyset^{\vdash} \neq L_\Sigma(P)$. Clearly, $\mathcal{L}$ is *inconsistent* (not consistent) precisely when $\vdash p$ for some $p \in P$, or alternatively when its only theory is inconsistent, that is, $\mathbf{Th}(\mathcal{L}) = \{L_\Sigma(P)\}$. $\mathcal{L}$ is said to be *trivial*, if for all nonempty $\Gamma \subseteq L_\Sigma(P)$ we have $\Gamma^{\vdash} = L_\Sigma(P)$. Equivalently, $\mathcal{L}$ is trivial if there exist distinct variables $p, q \in P$ such that $p \vdash q$. Another equivalent characterization is that $\mathcal{L}$ is trivial if $\mathbf{Th}(\mathcal{L}) \subseteq \{\emptyset, L_\Sigma(P)\}$. Of course, all inconsistent logics are trivial.

We say that a logic $\mathcal{L}' = \langle \Sigma', \_^{\vdash'} \rangle$ *extends* $\mathcal{L} = \langle \Sigma, \_^{\vdash} \rangle$ if $\Sigma \subseteq \Sigma'$, and $\vdash \subseteq \vdash'$, in the sense that $\Gamma^{\vdash} \subseteq \Gamma^{\vdash'}$ for every $\Gamma \subseteq L_\Sigma(P)$. We say that the extension of $\mathcal{L}$ by $\mathcal{L}'$ is *conservative* if for all $\Gamma \subseteq L_\Sigma(P)$, $\Gamma^{\vdash} = \Gamma^{\vdash'} \cap L_\Sigma(P)$. We also say that the extension of $\mathcal{L}$ by $\mathcal{L}'$ is *weakly conservative* if $\emptyset^{\vdash} = \emptyset^{\vdash'} \cap L_\Sigma(P)$. It is perhaps more common to express these properties in terms of the induced consequence relations. Clearly, $\mathcal{L}'$ extends $\mathcal{L}$ when $\Gamma \vdash \varphi$ implies $\Gamma \vdash' \varphi$ for all $\Gamma \cup \{\varphi\} \subseteq L_\Sigma(P)$. Furthermore, the extension is conservative precisely when $\Gamma \vdash \varphi$ if and only if $\Gamma \vdash' \varphi$, and weakly conservative when $\vdash \varphi$ if and only if $\vdash' \varphi$.

Given a signature Σ, it is well known that the set of all logics over Σ, $\mathbf{Log}(\Sigma)$, constitutes a complete lattice under the extension ordering defined above (see [17], for instance).

2.4 Hilbert Calculi

A *Hilbert calculus* is a pair $\mathcal{H} = \langle \Sigma, R\rangle$ where Σ is a signature, and $R \subseteq 2^{L_\Sigma(P)} \times L_\Sigma(P)$ is a set of *inference rules*. Given $\langle \Delta, \psi\rangle \in R$, we refer to Δ as the set of *premises* and to ψ as the *conclusion* of the rule. When the set of premises is empty, ψ is dubbed as an *axiom*. A rule is said to be *finitary* if it has a finite set of premises, and $\mathcal{H}$ is said to be *finitary* if all the rules in R are finitary. An inference rule $\langle \Delta, \psi\rangle \in R$ is often denoted by $\frac{\Delta}{\psi}$, or simply by $\frac{\psi_1 \ \dots \ \psi_n}{\psi}$ if $\Delta = \{\psi_1, \dots, \psi_n\}$ is finite, or even by $\frac{}{\psi}$ if $\Delta = \emptyset$.

Given $\Sigma \subseteq \Sigma'$ and $P \subseteq P'$, a Hilbert calculus $\mathcal{H} = \langle \Sigma, R\rangle$ induces a consequence operator $\_^{\vdash_\mathcal{H}}$ on $L_{\Sigma'}(P')$ such that, for each $\Gamma \subseteq L_{\Sigma'}(P')$, $\Gamma^{\vdash_\mathcal{H}}$ is the least set that contains Γ and is closed for all applications of instances of the inference rules in R, that is, if $\frac{\Delta}{\psi} \in R$ and $\sigma : P \to L_{\Sigma'}(P')$ is such that $\Delta^\sigma \subseteq \Gamma^{\vdash_\mathcal{H}}$ then $\psi^\sigma \in \Gamma^{\vdash_\mathcal{H}}$. Of course, this definition induces a logic $\mathcal{L}_\mathcal{H} = \langle \Sigma, \_^{\vdash_\mathcal{H}}\rangle$.

The definition of $\mathcal{L}_\mathcal{H}$ above is arguably too abstract, as it does not highlight the sequence of rule applications that leads one to conclude that $\Gamma \vdash_\mathcal{H} \varphi$, when that is the case. Let us be more detailed. Given $\Sigma \subseteq \Sigma'$, $P \subseteq P'$, and $\Gamma \subseteq L_{\Sigma'}(P')$, a $\mathcal{H}$-*derivation from* Γ is a sequence $\overline{\varphi} = \langle \varphi_\kappa\rangle_{\kappa<\eta}$ of formulas in $L_{\Sigma'}(P')$, for some ordinal η, such that, for each $\kappa < \eta$, either:

- $\varphi_\kappa \in \Gamma$, or
- there is $\frac{\Delta}{\psi} \in R$ and $\sigma : P \to L_{\Sigma'}(P')$ with $\psi^\sigma = \varphi_\kappa$ and $\Delta^\sigma \subseteq \{\varphi_\tau : \tau < \kappa\}$.

The fact that $\overline{\varphi}$ is a $\mathcal{H}$-derivation from Γ is denoted by $\Gamma \vdash_\mathcal{H} \overline{\varphi}$. We say that such a derivation is a $\mathcal{H}$-*proof from* Γ of each of its formulas, as it is clear that any prefix of a $\mathcal{H}$-derivation from Γ is also a $\mathcal{H}$-derivation from Γ.

Clearly, $\Gamma \vdash_\mathcal{H} \varphi$ precisely if φ has a $\mathcal{H}$-proof from Γ, that is, there exists some $\mathcal{H}$-derivation $\langle \varphi_\kappa\rangle_{\kappa<\eta}$ from Γ such that $\varphi = \varphi_\kappa$ for some $\kappa < \eta$, in which case $\langle \varphi_\iota\rangle_{\iota<\kappa+1}$ is a $\mathcal{H}$-proof of φ from Γ ending in φ.

Example 2.2 Along the chapter, in order to illustrate the problems at hand and the results obtained we will use the following collection of examples:

- $\mathcal{H}_{\mathsf{inc}(\Sigma)} = \langle \Sigma, R_{\mathsf{inc}}\rangle$, for each signature Σ, where R_{inc} has the unique rule

$$\frac{}{p}\,.$$

- $\mathcal{H}_{\mathsf{tonk}} = \langle \Sigma_{\mathsf{tonk}}, R_{\mathsf{tonk}}\rangle$, where Σ_{tonk} has a unique 2-place connective tonk, and R_{tonk} has the rules

$$\frac{p}{\mathsf{tonk}(p, q)} \qquad \frac{\mathsf{tonk}(p, q)}{q}\,.$$

- $\mathcal{H}_{\mathsf{cls}} = \langle \Sigma_{\mathsf{cls}}, R_{\mathsf{cls}}\rangle$, where Σ_{cls} has a unique 2-place connective $\Rightarrow$, and R_{cls} has the rules

$$\frac{}{(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))}$$

$$\frac{}{p \Rightarrow (q \Rightarrow p)} \qquad \frac{}{((p \Rightarrow q) \Rightarrow p) \Rightarrow p} \qquad \frac{p \qquad p \Rightarrow q}{q}\,.$$

- $\mathcal{H}_{\mathsf{int}} = \langle \Sigma_{\mathsf{int}}, R_{\mathsf{int}} \rangle$, where Σ_{int} has a unique 2-place connective $\to$, and R_{int} has the rules

$$\frac{}{(p \to (q \to r)) \to ((p \to q) \to (p \to r))}$$
$$\frac{}{p \to (q \to p)} \qquad \frac{p \quad p \to q}{q}.$$

- $\mathcal{H}_{\mathsf{neg}} = \langle \Sigma_{\mathsf{neg}}, R_{\mathsf{neg}} \rangle$, where Σ_{neg} has a unique 1-place connective $\neg$, and R_{neg} has the rules

$$\frac{p}{\neg\neg p} \qquad \frac{\neg\neg p}{p} \qquad \frac{p \quad \neg p}{q}.$$

- $\mathcal{H}_{\mathsf{cnj}} = \langle \Sigma_{\mathsf{cnj}}, R_{\mathsf{cnj}} \rangle$, where $\Sigma_{\wedge}$ has a unique 1-place connective $\wedge$, and R_{cnj} has the rules

$$\frac{p \wedge q}{p} \qquad \frac{p \wedge q}{q} \qquad \frac{p \quad q}{p \wedge q}.$$

Clearly, each $\mathcal{H}_{\mathsf{inc}(\Sigma)}$ induces an inconsistent logic, whereas $\mathcal{H}_{\mathsf{tonk}}$ is Prior's infamous tonk system and induces a consistent but trivial logic. The calculi $\mathcal{H}_{\mathsf{cls}}$ and $\mathcal{H}_{\mathsf{int}}$ induce the logics of *classical implication* and *intuitionistic implication*, respectively. Finally, $\mathcal{H}_{\mathsf{neg}}$ induces the logic of (classical or intuitionistic) *negation*, and $\mathcal{H}_{\mathsf{cnj}}$ the logic of (classical or intuitionistic) *conjunction*. Note that with the possible exception of the $\mathcal{H}_{\mathsf{inc}(\Sigma)}$ calculi, all other examples have very simple signatures with one single connective. This is a deliberate choice meant to keep the focus of attention on the relevant problems ahead, and not on the relative complexity of the syntax. △

Given a logic $\mathcal{L} = \langle \Sigma, \underline{\vdash} \rangle$, we can easily associate with it a Hilbert calculus $\mathcal{H}_{\mathcal{L}} = \langle \Sigma, \vdash \rangle$, where the consequence operator $\underline{\vdash}$ in the former is replaced by the induced consequence relation $\vdash$ in the latter. It is easy to check that $\mathcal{L}_{\mathcal{H}_{\mathcal{L}}} = \mathcal{L}$ (see [17], for instance).

For simplicity, we will use $\mathcal{L}_{\mathsf{name}}$ to denote the logic $\mathcal{L}_{\mathcal{H}_{\mathsf{name}}}$ for each of the calculi named in Example 2.2.

2.5 Fibring

Let $\mathcal{L}_1 = \langle \Sigma_1, \underline{\vdash}_1 \rangle$ and $\mathcal{L}_2 = \langle \Sigma_2, \underline{\vdash}_2 \rangle$ be two logics. The *fibring* of $\mathcal{L}_1$ and $\mathcal{L}_2$ is the smallest logic $\mathcal{L}_1 \bullet \mathcal{L}_2$ over the joint signature $\Sigma_{12} = \Sigma_1 \cup \Sigma_2$ that extends both $\mathcal{L}_1$ and $\mathcal{L}_2$. A direct characterization of this fibered logic can be most easily given by first defining the fibring of Hilbert calculi.

Given Hilbert calculi $\mathcal{H}_1 = \langle \Sigma_1, R_1 \rangle$ and $\mathcal{H}_2 = \langle \Sigma_2, R_2 \rangle$, let their *fibring* be the Hilbert calculus

$$\mathcal{H}_1 \bullet \mathcal{H}_2 = \langle \Sigma_{12}, R_1 \cup R_2 \rangle.$$

Clearly, besides joining their signatures, the fibring of the two calculi consists in simply putting together their rules.

We can now give a simple characterization (see [4]) of the fibring of two logics $\mathcal{L}_1$ and $\mathcal{L}_2$:

$$\mathcal{L}_1 \bullet \mathcal{L}_2 = \mathcal{L}_{\mathcal{H}_{\mathcal{L}_1} \bullet \mathcal{H}_{\mathcal{L}_1}} .$$

This means that if $\mathcal{L}_1 \bullet \mathcal{L}_2 = \langle \Sigma_{12}, \_^{\vdash_{12}} \rangle$ then, given $\Gamma \subseteq L_{\Sigma_{12}}(P)$, $\Gamma^{\vdash_{12}}$ is obtained by a (possibly transfinite) sequence of alternate applications of $\vdash_1$ and $\vdash_2$ using substitutions $\sigma : P \to L_{\Sigma_{12}}(P)$.

Both for logics and Hilbert calculi, when there are no shared connectives, i.e. $\Sigma_1 \cap \Sigma_2 = \emptyset$, the fibring is usually said to be *unconstrained*.

3 Theoremhood and Mixed Reasoning

In this section, we will review the notion of theoremhood in fibered logics, illustrated by means of a series of examples, and then, with the focus on unconstrained fibring, we obtain a technical result about the way mixed reasoning can be controlled in the fibered logic.

3.1 Theoremhood in Fibred Logics

We start by proving a few simple results characterizing the nature of the theoremhood relation in fibered logics. Let $\mathcal{L}_1 = \langle \Sigma_1, \_^{\vdash_1} \rangle$ and $\mathcal{L}_2 = \langle \Sigma_2, \_^{\vdash_2} \rangle$ be two logics, and $\mathcal{L}_1 \bullet \mathcal{L}_2 = \langle \Sigma_{12}, \_^{\vdash_{12}} \rangle$ be their fibring.

Proposition 3.1 $\mathcal{L}_1 \bullet \mathcal{L}_2$ has theorems if and only if $\mathcal{L}_1$ has theorems or $\mathcal{L}_2$ has theorems.

Proof If $\vdash_{12} \psi$, for some $\psi \in L_{\Sigma_{12}}(P)$, then, by definition of fibring, there exists a $(\mathcal{H}_{\mathcal{L}_1} \bullet \mathcal{H}_{\mathcal{L}_2})$-proof $\overline{\psi} = \langle \psi_\kappa \rangle_{\kappa < \eta + 1}$ (from $\emptyset$) such that $\psi_\eta = \psi$. Thus, it is also the case that $\vdash_{12} \psi_0$. Hence, by definition of derivation, there must exist $\frac{\Delta}{\varphi} \in (\vdash_1 \cup \vdash_2)$ and $\sigma : P \to L_\Sigma(P)$ such that $\varphi^\sigma = \psi_0$ and $\Delta^\sigma \subseteq \{\psi_\tau : \tau < 0\} = \emptyset$. Therefore, $\Delta = \emptyset$, and we have that $\vdash_i \varphi$ provided that $\frac{\emptyset}{\varphi} \in \vdash_i$. We conclude that either $\mathcal{L}_1$ has theorems or $\mathcal{L}_2$ has theorems.

Reciprocally, let $i \in \{1, 2\}$ and assume that $\mathcal{L}_i$ has a theorem, that is, $\vdash_i \varphi$ for some $\varphi \in L_{\Sigma_i}(P)$. By definition of fibring, $\vdash_i \varphi$ implies $\vdash_{12} \varphi$, and we conclude that $\mathcal{L}_1 \bullet \mathcal{L}_2$ has theorems. $\square$

This result tells us that the theoremhood relation is empty in logics resulting from the fibring of two logics without theorems. Let us look at an example.

Example 3.2 Take the logics $\mathcal{L}_{\mathsf{cnj}}$ and $\mathcal{L}_{\mathsf{neg}}$ from Example 2.2. Their fibring $\mathcal{L}_{\mathsf{cnj}} \bullet \mathcal{L}_{\mathsf{neg}}$ does not have theorems, since neither $\mathcal{L}_{\mathsf{cnj}}$ nor $\mathcal{L}_{\mathsf{neg}}$ has theorems. Just note that the calculi defining the two logics do not have any axioms.

The logic $\mathcal{L}_{\mathsf{cnj}} \bullet \mathcal{L}_{\mathsf{neg}}$ is not to be confused with the conjunction and negation fragment of classical logic, which of course has theorems, e.g., $\neg(p \wedge \neg p)$. △

Even if one of the logics has theorems, the theoremhood relation in the fibered logic can still be quite uninteresting.

Proposition 3.3 If $\mathcal{L}_1$ is inconsistent or $\mathcal{L}_2$ is inconsistent then $\mathcal{L}_1 \bullet \mathcal{L}_2$ is inconsistent.

Proof Let $\mathcal{L}_i$ be inconsistent, for some $i \in \{1, 2\}$. Then $\vdash_i p$, which implies $\vdash_{12} p$, and we conclude that $\mathcal{L}_1 \bullet \mathcal{L}_2$ is inconsistent. □

This result shows that every formula is a theorem in logics resulting from fibrings involving a trivial logic. Let us see an example.

Example 3.4 Take the $\mathcal{L}_{\mathsf{inc}(\Sigma)}$ from Example 2.2, for some signature Σ, and consider any other logic $\mathcal{L}$. Their fibring $\mathcal{L}_{\mathsf{inc}(\Sigma)} \bullet \mathcal{L}$ is inconsistent. Namely, $p \in P$ is a theorem, and therefore, by structurality, every formula is a theorem. △

Still, a fibered logic can be inconsistent even when both the logics being combined are consistent.

Proposition 3.5 If $\mathcal{L}_1$ is consistent and trivial and $\mathcal{L}_2$ has theorems then $\mathcal{L}_1 \bullet \mathcal{L}_2$ is inconsistent.

Proof If $\mathcal{L}_1$ is trivial, then we know that $p \vdash_1 q$, where $p, q \in P$ and $p \neq q$. If $\mathcal{L}_2$ has theorems, let $\vdash_2 \varphi$ for some $\varphi \in L_{\Sigma_2}(P)$. Now, $p \vdash_1 q$ implies $p \vdash_{12} q$, and structurality, on its turn, implies that $\varphi \vdash_{12} q$. But $\vdash_2 \varphi$ implies $\vdash_{12} \varphi$, and thus $\vdash_{12} q$. We conclude that $\mathcal{L}_1 \bullet \mathcal{L}_2$ is inconsistent. □

An example follows.

Example 3.6 Take the logics $\mathcal{L}_{\mathsf{tonk}}$ and $\mathcal{L}_{\mathsf{cls}}$ from Example 2.2, where $\mathcal{L}_{\mathsf{tonk}}$ is consistent but trivial, as $p \vdash_{\mathsf{tonk}} q$, and $\mathcal{L}_{\mathsf{cls}}$ has theorems, namely $\vdash_{\mathsf{cls}} p \Rightarrow p$. Their fibring $\mathcal{L}_{\mathsf{tonk}} \bullet \mathcal{L}_{\mathsf{cls}} = \langle \Sigma_{\mathsf{tonk}} \cup \Sigma_{\mathsf{cls}}, \_\vdash\_ \rangle$ is inconsistent. Take any formula $\varphi \in L_{\Sigma_{\mathsf{tonk}} \cup \Sigma_{\mathsf{cls}}}(P)$. Easily, $p \vdash_{\mathsf{tonk}} q$ implies that $p \vdash q$, and structurality in the fibered logic implies that $p \Rightarrow p \vdash \varphi$. Therefore, as $\vdash_{\mathsf{cls}} p \Rightarrow p$ implies $\vdash p \Rightarrow p$, we have $\vdash \varphi$. △

Triviality is actually preserved by fibring, leading to uninteresting trivial fibered logics, where the theoremhood relation is therefore total (when the fibered logic becomes inconsistent) or empty (when the fibered logic remains consistent).

Proposition 3.7 If $\mathcal{L}_1$ is trivial or $\mathcal{L}_2$ is trivial then $\mathcal{L}_1 \bullet \mathcal{L}_2$ is trivial.

Proof Let $\mathcal{L}_i$ be trivial, for some $i \in \{1, 2\}$. Then $p \vdash_i q$ with $p \neq q$, which implies $p \vdash_{12} q$, and we conclude that $\mathcal{L}_1 \bullet \mathcal{L}_2$ is trivial. □

At this point, we know that the theoremhood relationship in $\mathcal{L}_1 \bullet \mathcal{L}_2$ is only interesting if $\mathcal{L}_1$ and $\mathcal{L}_2$ are both nontrivial and at least one of the two has theorems. In that case, how can we decide if a (mixed) formula is a theorem of $\mathcal{L}_1 \bullet \mathcal{L}_2$?

Example 3.8 Take the logics $\mathcal{L}_{\mathsf{cls}}$ and $\mathcal{L}_{\mathsf{int}}$ from Example 2.2. Both logics are nontrivial, and have theorems. Furthermore, it is known that $\mathcal{L}_{\mathsf{cls}}$ and $\mathcal{L}_{\mathsf{int}}$ are decidable. The fibring $\mathcal{L}_{\mathsf{cls}} \bullet \mathcal{L}_{\mathsf{int}}$ was studied in [6, 7], and shown to be a conservative extension of both $\mathcal{L}_{\mathsf{cls}}$ and $\mathcal{L}_{\mathsf{int}}$. We will show below that the theorems of $\mathcal{L}_{\mathsf{cls}} \bullet \mathcal{L}_{\mathsf{int}}$ are indeed decidable. △

3.2 Consequences of Nonmixed Formulas

Let us now have a more technical look at the patterns of mixed reasoning that occur in fibered logics. We start with a result about syntax.

Lemma 3.9 Let $\Sigma \subseteq \Sigma'$ and $\Gamma \subseteq L_\Sigma(P)$. Then, for every $\sigma : P \to L_{\Sigma'}(P)$, and every two η-sequences $\overline{\alpha}$ and $\overline{\beta}$ of formulas in $L_{\Sigma'}(P)$, with $\overline{\alpha}$ injective, there exists $\rho : P \to L_{\Sigma'}(P)$ such that

$$\Gamma^\rho = \Gamma^\sigma[\overline{\alpha}/\overline{\beta}]_\Sigma.$$

Proof One should observe, to start with, that $\mathsf{Mon}_\Sigma(\Gamma) = \emptyset$. Thus, if $\alpha_\kappa \in \mathsf{Mon}_\Sigma(\varphi^\sigma)$ for some $\varphi \in \Gamma$, then there must exist a variable $p \in P$ occurring in φ such that $\alpha_\kappa \in \mathsf{Mon}_\Sigma(\sigma(p))$. Hence, the substitution defined by $\rho(q) = \sigma(q)[\overline{\alpha}/\overline{\beta}]_\Sigma$ for every $q \in P$ satisfies the conditions of the lemma. □

The previous lemma reflects the fact that the occurrence of Σ-monoliths in instances of $L_\Sigma(P)$ formulas is only possible if they are brought about by the substitution.

Next, we prove a (quite) technical lemma, characterizing the irrelevance of certain monoliths in derivations from a set of variables in logics obtained by unconstrained fibring, motivated by the square bracket monolith substitutions introduced earlier. Note that the disjointness of the signatures is instrumental in proving this result.

Lemma 3.10 Let $\mathcal{H}_1 = \langle \Sigma_1, R_1 \rangle$ and $\mathcal{H}_2 = \langle \Sigma_2, R_2 \rangle$ be Hilbert calculi such that $\Sigma_1 \cap \Sigma_2 = \emptyset$, $V \subseteq P$ and $\overline{\psi} = \langle \psi_\kappa \rangle_{\kappa<\eta}$ a sequence of $L_{\Sigma_{12}}(P)$ formulas.

If $V \vdash_{\mathcal{H}_{12}} \overline{\psi}$ and $\alpha \in L_{\Sigma_{12}}(P)$ then, either

- $\alpha = \psi_\kappa$ for some $\kappa < \eta$, or
- $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\eta}$ for every $\beta \in L_{\Sigma_{12}}(P)$ and $i \in \{1, 2\}$.

Proof Let us assume that $\alpha \neq \psi_\kappa$ for every $\kappa < \eta$. The proof of the second condition follows by complete transfinite induction on the size η of the derivation. For each $\iota < \tau \leq \eta$, we assume, by induction hypothesis, that $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\iota}$, and show that it implies $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\tau}$.

If $\tau = 0$ the result is trivial, as the derivation is empty. If τ is a limit ordinal then the result is immediate, by definition of derivation. If τ is a successor ordinal, we have to consider two cases.

1. $\psi_{\tau-1} \in V$.
   Then, $\psi_{\tau-1} \in V \subseteq P$, and we have that $\alpha \notin \mathsf{Mon}_{\Sigma_i}(\psi_{\tau-1}) = \emptyset$. Thus, $\psi_{\tau-1}[\alpha/\beta]_{\Sigma_i} = \psi_{\tau-1} \in V$.
   By induction hypothesis, we have $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\tau-1}$ and so, by definition of derivation, we also have $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\tau}$.
2. $\psi_{\tau-1} = \varphi^\sigma$, $\frac{\Delta}{\varphi} \in R_1 \cup R_2$, and $\Delta^\sigma \subseteq \{\psi_\kappa : \kappa < \tau - 1\}$.
   Here we have two possibilities, given that $i \in \{1, 2\}$.
   (a) $\frac{\Delta}{\varphi} \in R_i$.
   Applying Lemma 3.9 to $\Delta \cup \{\varphi\} \subseteq L_{\Sigma_i}(P)$, σ, α and β, we know that there exists ρ such that
   $\varphi^\rho = \varphi^\sigma[\alpha/\beta]_{\Sigma_i} = \psi_{\tau-1}[\alpha/\beta]_{\Sigma_i}$, and also $\Delta^\rho = \Delta^\sigma[\alpha/\beta]_{\Sigma_i} \subseteq \{\psi_\kappa[\alpha/\beta]_{\Sigma_i} : \kappa < \tau - 1\}$.
   By induction hypothesis, we have $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\tau-1}$ and so, by definition of derivation, we also have $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\tau}$.
   (b) $\frac{\Delta}{\varphi} \notin R_i$.
   If $\delta \in \Delta$ and $\alpha \in \mathsf{Mon}_{\Sigma_i}(\delta^\sigma)$ then (since $\delta \in L_{\Sigma_{3-i}}(P)$) either $\alpha = \delta^\sigma$ or $\mathsf{head}(\delta^\sigma) \in \Sigma_i$. By assumption the former cannot be the case, therefore we must have $\mathsf{head}(\delta^\sigma) \in \Sigma_i$. Hence $\delta \in P$ and $\beta \in \mathsf{Mon}_{\Sigma_i}(\delta^\sigma)$. Consider the substitution defined by $\rho(q) = \sigma(q)[\alpha/\beta]_{\Sigma_i}$ for every $q \in P$. Clearly, as above, $\varphi^\rho = \varphi^\sigma[\alpha/\beta]_{\Sigma_i} = \psi_{\tau-1}[\alpha/\beta]_{\Sigma_i}$, and also $\Delta^\rho = \Delta^\sigma[\alpha/\beta]_{\Sigma_i} \subseteq \{\psi_\kappa[\alpha/\beta]_{\Sigma_i} : \kappa < \tau - 1\}$.
   By induction hypothesis we have that $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\tau-1}$ and so, by definition of derivation, we also have $V \vdash_{\mathcal{H}_{12}} \langle \psi_\kappa[\alpha/\beta]_{\Sigma_i} \rangle_{\kappa<\tau}$. □

In order to finally state and prove our main result about mixed reasoning, we need the following definition, where we are going to collect extra variables to represent some contextual relevant properties.

Definition 3.11 Let $\mathcal{H}_1 = \langle \Sigma_1, R_1 \rangle$ and $\mathcal{H}_2 = \langle \Sigma_2, R_2 \rangle$ be Hilbert calculi, $V \subseteq P$ $\psi \in L_{\Sigma_{12}}(P)$ and $i \in \{1, 2\}$. We define $X^i_V(\psi) \subseteq \{x_0\} \cup \{x_{e^{-1}(\psi')} : \psi' \in \mathsf{Mon}_{\Sigma_t}(\psi)\}$ to be the set such that:

- $x_0 \in X^i_V(\psi)$ whenever $V^{\vdash_{\mathcal{H}_{12}}} \neq \emptyset$, and
- $x_{e^{-1}(\psi')} \in X^i_V(\psi)$ whenever $V \vdash_{\mathcal{H}_{12}} \psi'$.

Example 3.12 If $\mathcal{H}_1 = \mathcal{H}_{\mathsf{cnj}}$ and $\mathcal{H}_2 = \mathcal{H}_{\mathsf{tonk}}$, then, for $i \in \{1, 2\}$, we have the following equalities.

$$X^i_\emptyset(q) = \emptyset$$
$$X^i_{\{p\}}(q) = \{x_0\}$$
$$X^1_\emptyset(p \wedge \mathsf{tonk}(p, q)) = \{x_{e^{-1}(\mathsf{tonk}(p,q))}\}$$
$$X^1_{\{p\}}(p \wedge \mathsf{tonk}(p, q)) = \{x_0, x_{e^{-1}(\mathsf{tonk}(p,q))}\}$$
$$X^2_\emptyset(p \wedge \mathsf{tonk}(p, q)) = \{x_{e^{-1}(p \wedge \mathsf{tonk}(p,q))}\}.$$

△

We can finally prove the following result, relating proofs from variables in the fibered logic with proofs from variables in the component logics, in the case of unconstrained fibring.

Proposition 3.13 Let $\mathcal{H}_1 = \langle \Sigma_1, R_1 \rangle$ and $\mathcal{H}_2 = \langle \Sigma_2, R_2 \rangle$ be Hilbert calculi such that $\Sigma_1 \cap \Sigma_2 = \emptyset$, $V \subsetneq P$, and $\psi \in L_{\Sigma_{12}}(P)$. Then, for $i, j \in \{1, 2\}$ with $i \neq j$, $V \vdash_{\mathcal{H}_{12}} \psi$ if and only if

$$V, X_V^i(\psi) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi) \text{ or } (V \cup (\{x_0\} \cap X_V^i(\psi)))^{\vdash_{\mathcal{H}_j}} = L_{\Sigma_{12}}(P \cup X).$$

Proof We start by proving the simpler implication from right to left. Let $p \in P \setminus V$. If $V^{\vdash_{\mathcal{H}_{12}}} \neq \emptyset$ fix $\gamma \in V^{\vdash_{\mathcal{H}_{12}}}$, and let $\sigma : P \cup X \to L_{\Sigma_{12}}(P)$ be such that $\sigma(p) = \psi$, $\sigma(q) = q$ if $q \in P$ with $q \neq p$, $\sigma(x_n) = e(n)$ for $n > 0$, and

$$\sigma(x_0) = \begin{cases} x_0 & \text{if } V^{\vdash_{\mathcal{H}_{12}}} = \emptyset, \\ \gamma & \text{if } V^{\vdash_{\mathcal{H}_{12}}} \neq \emptyset. \end{cases}$$

Now, knowing on one hand that $V, X_V^i(\psi) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi)$ and using structurality, we obtain $V^\sigma, (X_V^i(\psi))^\sigma \vdash_{\mathcal{H}_i} (\mathsf{skel}_{\Sigma_i}(\psi))^\sigma$. But clearly, $V^\sigma = V$, $(X_V^i(\psi))^\sigma \subseteq V^{\vdash_{\mathcal{H}_{12}}}$, $(\mathsf{skel}_{\Sigma_i}(\psi))^\sigma = \psi$, and we conclude that $V \vdash_{\mathcal{H}_{12}} \psi$.

If, on the other hand, we know that $(V \cup (\{x_0\} \cap X_V^i(\psi)))^{\vdash_{\mathcal{H}_j}} = L_{\Sigma_{12}}(P \cup X)$, then we have $V \cup (\{x_0\} \cap X_V^i(\psi)) \vdash_{\mathcal{H}_j} p$. But clearly, $V^\sigma = V$, $(\{x_0\} \cap X_V^i(\psi))^\sigma \subseteq V^{\vdash_{\mathcal{H}_{12}}}$, $p^\sigma = \psi$, and we conclude again that $V \vdash_{\mathcal{H}_{12}} \psi$.

Let us now consider the implication from left to right, and assume that we have $V \vdash_{\mathcal{H}_{12}} \psi$. Clearly, $V^{\vdash_{\mathcal{H}_{12}}} \neq \emptyset$ and so $x_0 \in X_V^t(\varphi)$ for every $\varphi \in L_{\Sigma_{12}}(P)$ and every $t \in \{1, 2\}$. If $(V \cup (\{x_0\} \cap X_V^i(\psi)))^{\vdash_{\mathcal{H}_j}} = (V \cup \{x_0\})^{\vdash_{\mathcal{H}_j}} = L_{\Sigma_{12}}(P \cup X)$ then the statement immediately follows, hence we proceed assuming that we have

$$(V \cup \{x_0\})^{\vdash_{\mathcal{H}_j}} \neq L_{\Sigma_{12}}(P \cup X).$$

The proof follows by complete transfinite induction on the length of $\mathcal{H}_{12}$-derivations. Given that $V \vdash_{\mathcal{H}_{12}} \psi$, there must exist a $\mathcal{H}_{12}$-derivation $\overline{\psi} = \langle \psi_\kappa \rangle_{\kappa < \eta + 1}$ from V such that $\psi_\eta = \psi$. We want to show that $V, X_V^i(\psi_\eta) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\eta)$. Thus, we will prove that $V, X_V^i(\psi_\tau) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\tau)$ for any $\tau \leq \eta$ assuming, by induction hypothesis, that the left to right implication holds for any $\mathcal{H}_{12}$-derivation with length lower than τ, and for both i and j.

Note that the case when $\mathsf{head}(\psi_\tau) \in \Sigma_j$ is trivial. Indeed, in that situation, we have that $\mathsf{Mon}_{\Sigma_i}(\psi_\tau) = \{\psi_\tau\}$, and thus $\mathsf{skel}_{\Sigma_i}(\psi_\tau) = x_{e^{-1}(\psi_\tau)} \in X_V^i(\psi_\tau)$. But then, clearly, $V, X_V^i(\psi_\tau) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\tau)$. We assume, henceforth, that either $\psi_\tau \in P$ or $\mathsf{head}(\psi_\tau) \in \Sigma_i$, which implies that $\mathsf{skel}_{\Sigma_j}(\psi_\tau) \in P \cup X$.

We have to consider two cases.

(1) $\psi_\tau \in V$.
We have that $\mathsf{Mon}_{\Sigma_i}(\psi_\tau) = \emptyset$. But then, clearly, $V, X_V^i(\psi_\tau) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\tau)$ as $\mathsf{skel}_{\Sigma_i}(\psi_\tau) = \psi_\tau \in V$.

(2) $\psi_\tau = \varphi^\sigma$, $\frac{\Delta}{\varphi} \in R_1 \cup R_2$, and $\Delta^\sigma \subseteq \{\psi_\kappa : \kappa < \tau\}$.
Here we have two possibilities.

(a) $\frac{\Delta}{\varphi} \in R_i$.
Let $M = \bigcup_{\kappa \leq \tau} \mathsf{Mon}_{\Sigma_i}(\psi_\kappa)$. Consider any injective sequence $\overline{\alpha}$ of formulas in M where each formula in M appears exactly once, and define $\overline{\beta}$ to be of the same length sequence such that each $\beta_\iota = x_{e^{-1}(\alpha_\iota)}$. Note that $\psi_\kappa[\overline{\alpha}/\overline{\beta}]_{\Sigma_i} = \mathsf{skel}_{\Sigma_i}(\psi_\kappa)$ for every $\kappa \leq \tau$.
Hence, applying Lemma 3.9 to $\Delta \cup \{\varphi\} \subseteq L_{\Sigma_i}(P)$, σ, $\overline{\alpha}$ and $\overline{\beta}$, we know that there exists ρ such that $\varphi^\rho = \varphi^\sigma[\overline{\alpha}/\overline{\beta}]_{\Sigma_i} = \psi_\tau[\overline{\alpha}/\overline{\beta}]_{\Sigma_i} = \mathsf{skel}_{\Sigma_i}(\psi_\tau)$, and $\Delta^\rho = \Delta^\sigma[\overline{\alpha}/\overline{\beta}]_{\Sigma_i} \subseteq \{\psi_\kappa[\overline{\alpha}/\overline{\beta}]_{\Sigma_i} : \kappa < \tau\} = \{\mathsf{skel}_{\Sigma_i}(\psi_\kappa) : \kappa < \tau\}$, and we can conclude that $\{\mathsf{skel}_{\Sigma_i}(\psi_\kappa) : \kappa < \tau\} \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\tau)$.
By induction hypothesis we have that $V, X^i_V(\psi_\kappa) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\kappa)$ for each $\kappa < \tau$, and therefore, we have that $V, \bigcup_{\kappa<\tau} X^i_V(\psi_\kappa) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\tau)$.
Consider the substitution $\mu : P \cup X \to L_{\Sigma_i}(P \cup X)$ such that

- $\mu(p) = p$ if $p \in P$
- $\mu(x_n) = x_0$ if $e(n) \notin \mathsf{Mon}_{\Sigma_i}(\psi_\tau)$,
- and $\mu(x_n) = x_n$ otherwise.

Clearly, we have $V^\mu = V$, $(X^i_V(\psi_\kappa))^\mu \subseteq X^i_V(\psi_\tau)$ for each $\kappa < \tau$, and $(\mathsf{skel}_{\Sigma_i}(\psi_\tau))^\mu = \mathsf{skel}_{\Sigma_i}(\psi_\tau)$. Thus, by structurality and monotonicity, V, $X^i_V(\psi_\tau) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\tau)$.

(b) $\frac{\Delta}{\varphi} \in R_j$.
If $\psi_\kappa = \psi_\tau$ for some $\kappa < \tau$, by induction hypothesis we have that $V, X^i_V(\psi_\kappa) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\kappa)$ and so $V, X^i_V(\psi_\tau) \vdash_{\mathcal{H}_i} \mathsf{skel}_{\Sigma_i}(\psi_\tau)$.
We shall finish the proof by showing that no other case is possible. That is, assuming that either $\psi_\tau \in P \setminus V$ (the case when $\psi_\tau \in V$ was covered in **(1)**) or $\mathsf{head}(\psi_\tau) \in \Sigma_i$, and also that $\psi_\kappa \neq \psi_\tau$ for every $\kappa < \tau$, we will derive a contradiction. Let $p \in P \setminus V$. We split in yet another two cases.

(i) $\psi_\tau \in P \setminus V$.
Let $\psi_\tau = p \in P \setminus V$, and $M = \bigcup_{\kappa \leq \tau} \mathsf{Mon}_{\Sigma_j}(\psi_\kappa)$. Consider any injective sequence $\overline{\alpha}$ of formulas in M where each formula in M appears exactly once, and define $\overline{\beta}$ to be the same length sequence such that each $\beta_\iota = x_{e^{-1}(\alpha_\iota)}$. Hence, applying Lemma 3.9 to $\Delta \cup \{\varphi\}$, σ, $\overline{\alpha}$ and $\overline{\beta}$, we know that there exists ρ such that $\varphi^\rho = \varphi^\sigma[\overline{\alpha}/\overline{\beta}]_{\Sigma_j} = \psi_\tau[\overline{\alpha}/\overline{\beta}]_{\Sigma_j} = \mathsf{skel}_{\Sigma_j}(\psi_\tau) = p$, and also $\Delta^\rho = \Delta^\sigma[\overline{\alpha}/\overline{\beta}]_{\Sigma_j} \subseteq \{\psi_\kappa[\overline{\alpha}/\overline{\beta}]_{\Sigma_j} : \kappa < \tau\} = \{\mathsf{skel}_{\Sigma_j}(\psi_\kappa) : \kappa < \tau\}$. We conclude that $\{\mathsf{skel}_{\Sigma_j}(\psi_\kappa) : \kappa < \tau\} \vdash_{\mathcal{H}_j} \mathsf{skel}_{\Sigma_j}(\psi_\tau) = p$.
As $V \vdash_{\mathcal{H}_{12}} \langle\psi_\kappa\rangle_{\kappa<\tau}$, by induction hypothesis, we get to know that $V, X^j_V(\psi_\kappa) \vdash_{\mathcal{H}_j} \mathsf{skel}_{\Sigma_j}(\psi_\kappa)$ for each $\kappa < \tau$. Thus, we also have V, $\bigcup_{\kappa<\tau} X^j_V(\psi_\kappa) \vdash_{\mathcal{H}_j} p$.
Consider the substitution ν such that $\nu(x_n) = x_0$ for all $n \in \mathbb{N}_0$, and $\nu(q) = q$ for all $q \in P$. Clearly, $V^\nu = V$, $(\bigcup_{\kappa<\tau} X^j_V(\psi'_\kappa))^\nu \subseteq \{x_0\}$ and $p^\nu = p$.

Therefore, by structurality and using ν, we obtain $V^\nu, \bigcup_{\kappa<\tau} X_V^j(\psi'_\kappa)^\nu \vdash_{\mathcal{H}_j} p^\nu$, which implies $(V \cup \{x_0\})^{\vdash_{\mathcal{H}_j}} = L_{\Sigma_{12}}(P \cup X)$, and we obtain a contradiction.

(ii) $\mathsf{head}(\psi_\tau) \in \Sigma_i$.
Let $p \in P \setminus V$ and define, for all $\kappa \le \tau$, let $\psi'_\kappa = \psi_\kappa[\psi_\tau/p]_{\Sigma_j}$. Clearly, $\psi'_\tau = \psi_\tau[\psi_\tau/p]_{\Sigma_j} = p$. Let $\mu : P \cup X \to L_{\Sigma_{12}}(P \cup X)$ defined by $\mu(x_{e^{-1}(\psi_\tau)}) = p$, $\mu(x_n) = x_n$ if $n = 0$ or $e(n) \neq \psi_\tau$, and $\mu(q) = q$ for each $q \in P$. Easily, we have that $\mathsf{skel}_{\Sigma_j}(\psi'_\kappa) = (\mathsf{skel}_{\Sigma_j}(\psi_\kappa))^\mu$ for all $\kappa \le \tau$. Arguing as in **(i)**, we know that there exists ρ such that $\varphi^\rho = \varphi^\sigma[\overline{\alpha}/\overline{\beta}]_{\Sigma_j} = \psi_\tau[\overline{\alpha}/\overline{\beta}]_{\Sigma_j} = \mathsf{skel}_{\Sigma_j}(\psi_\tau) = x_{e^{-1}(\psi_\tau)}$, and also $\Delta^\rho = \Delta^\sigma[\overline{\alpha}/\overline{\beta}]_{\Sigma_j} \subseteq \{\psi_\kappa[\overline{\alpha}/\overline{\beta}]_{\Sigma_j} : \kappa < \tau\} = \{\mathsf{skel}_{\Sigma_j}(\psi_\kappa) : \kappa < \tau\}$. We conclude that $\{\mathsf{skel}_{\Sigma_j}(\psi_\kappa) : \kappa < \tau\} \vdash_{\mathcal{H}_j} x_{e^{-1}(\psi_\tau)}$. Therefore, by structurality and using μ, it is the case that $(\{\mathsf{skel}_{\Sigma_j}(\psi_\kappa) : \kappa < \tau\})^\mu \vdash_{\mathcal{H}_j} (\mathsf{skel}_{\Sigma_j}(\psi_\tau))^\mu$, that is, $\{\mathsf{skel}_{\Sigma_j}(\psi'_\kappa) : \kappa < \tau\} \vdash_{\mathcal{H}_j} p$.
As $V \vdash_{\mathcal{H}_{12}} \langle\psi_\kappa\rangle_{\kappa<\tau}$ and we assumed that $\psi_\kappa \neq \psi_\tau$ for all $\kappa < \tau$, we can use Lemma 3.10 to conclude that $V \vdash_{\mathcal{H}_{12}} \langle\psi'_\kappa\rangle_{\kappa<\tau}$, and by induction hypothesis, we get that $V, X_V^j(\psi'_\kappa) \vdash_{\mathcal{H}_j} \mathsf{skel}_{\Sigma_j}(\psi'_\kappa)$ for each $\kappa < \tau$. Thus, we also have $V, \bigcup_{\kappa<\tau} X_V^j(\psi'_\kappa) \vdash_{\mathcal{H}_j} p$.
Using the substitution ν as defined in (i), and arguing in the same manner, we arrive at a contradiction. ☐

4 Decidability

Using the technical results about mixed proofs obtained in the last section, for unconstrained fibring, we can now state and prove our main results concerning theoremhood decidability.

The result of Proposition 3.13 is quite promising, if we want to analyze the decidability of logics obtained by unconstrained fibring. Indeed, under the appropriate circumstances, the result allows us to reduce the problem of checking a theorem (or even a derivation from a finite set of variables) in a fibered logic to the problem of checking a derivation from variables in one of the component logics. Let us have a look at an enlightening example.

Example 4.1 Consider the unconstrained fibring $\mathcal{L}_{\mathsf{cls}} \bullet \mathcal{L}_{\mathsf{int}} = \langle \Sigma_{\mathsf{cls}} \cup \Sigma_{\mathsf{int}}, \vdash\rangle$ that we have used in Example 3.8. As direct applications of Proposition 3.13, we can conclude that

- $\vdash (p \to q) \Rightarrow (p \to q)$, because $(p \to q) \Rightarrow (p \to q) = (x_1 \Rightarrow x_1)[p \to q]_{\Sigma_{\mathsf{cls}}}$, and we have

$$\nvdash_{\mathsf{int}} p \to q \quad \text{and} \quad x_0 \vdash_{\mathsf{cls}} x_1 \Rightarrow x_1;$$

- $\vdash p \Rightarrow (p \to p)$, because $p \Rightarrow (p \to p) = (p \Rightarrow x_1)[p \to p]_{\Sigma_{\mathsf{cls}}}$, and we have

$$\vdash_{\mathsf{int}} p \to p \quad \text{and} \quad x_0, x_1 \vdash_{\mathsf{cls}} p \Rightarrow x_1;$$

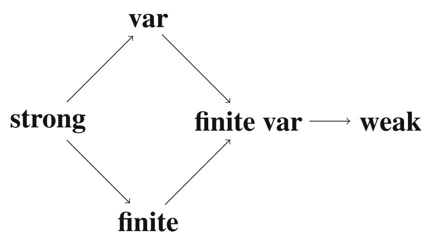


Fig. 1 The decidability spectrum for a logic $\mathcal{L}$

- $\not\vdash (p \to p) \Rightarrow p$, because $(p \to p) \Rightarrow p = (x_1 \Rightarrow p)[p \to p]_{\Sigma_{\mathsf{cls}}}$, and we have

$$\vdash_{\mathsf{int}} p \to p \quad \text{and} \quad x_0, x_1 \not\vdash_{\mathsf{cls}} x_1 \Rightarrow p.$$

Clearly, in this way, to decide theoremood in $\mathcal{L}_{\mathsf{cls}} \bullet \mathcal{L}_{\mathsf{int}}$, all we need is to have decision procedures for deciding derivation from variables in $\mathcal{L}_{\mathsf{cls}}$ and $\mathcal{L}_{\mathsf{int}}$. △

Recall that a set A is said to be *decidable* if there exists an algorithm D_A that always terminates such that

$$\mathsf{D}_A(w) = \begin{cases} \mathsf{yes} & \text{if } w \in A, \\ \mathsf{no} & \text{if } w \notin A. \end{cases}$$

There are several different flavors of *decidability* that make sense when applied to a logic $\mathcal{L} = \langle \Sigma, \_^{\vdash} \rangle$, some more standard than others. Assuming that Σ is decidable, we say that $\mathcal{L}$ is

- *strongly decidable* if there is an algorithm A such that, for each decidable $\Gamma \subseteq L_\Sigma(P)$, $\mathsf{A}(\mathsf{D}_\Gamma)$ terminates and outputs algorithm $\mathsf{D}_{\Gamma^\vdash}$ that decides $\Gamma^\vdash$;
- *finitely decidable* if there is an algorithm A such that, for each finite set $\Gamma \subseteq L_\Sigma(P)$, $\mathsf{A}(\Gamma)$ terminates and outputs an algorithm $\mathsf{D}_{\Gamma^\vdash}$ that decides $\Gamma^\vdash$;
- *var-decidable* if there is an algorithm A such that, for each decidable $V \subseteq P$, $\mathsf{A}(\mathsf{D}_V)$ terminates and outputs an algorithm $\mathsf{D}_{V^\vdash}$ that decides $V^\vdash$;
- *finitely var-decidable* if there is an algorithm A such that, for each finite set $V \subseteq P$, $\mathsf{A}(V)$ terminates and outputs an algorithm $\mathsf{D}_{V^\vdash}$ that decides $V^\vdash$;
- *weakly decidable* if $\emptyset^\vdash$ is decidable.

The decidability of a logic is often identified in the literature with weak decidability, precisely the version that we aimed at in this chapter, but one cannot deny the interest in strong or even finite decidability. Note that, due to the shape of Proposition 3.13, we will actually have to deal with finite var-decidability instead of weak decidability.

Of course, these different notions of decidability are related with each other in a simple and clear way, as illustrated in Fig. 1. As we know, one can smoothly replace strong decidability with finite decidability when dealing with a finitary logic. Moreover, in the presence of a finitary logic that has an implication-like connective enjoying some form of the *deduction theorem*, weak decidability is equivalent to finite decidability, and also to finite var-decidability.

Proposition 4.2 If $\mathcal{L}_1 = \langle \Sigma_1, \_^{\vdash_1} \rangle$ and $\mathcal{L}_2 = \langle \Sigma_2, \_^{\vdash_2} \rangle$ are both finite var-decidable, and $\Sigma_1 \cap \Sigma_2 = \emptyset$, then $\mathcal{L}_1 \bullet \mathcal{L}_2$ is finite var-decidable.

Proof As a result of Propositions 3.1, 3.3, 3.5 and 3.7, we can assume that both $\mathcal{L}_1$ and $\mathcal{L}_2$ are nontrivial (otherwise the fibring is trivial, and either every formula is a theorem, or no formula is a theorem), at least one of the logics has theorems (or the fibring would no have theorems). Let $\mathcal{L}_1$ be finite var-decidable with algorithm A^1, and $\mathcal{L}_2$ be finite var-decidable with algorithm A^2. Consider the following algorithm:

$$\begin{array}{ll} \mathsf{A}: & \mathsf{input} V \subseteq P \\ & \mathsf{output}\ \mathsf{D}: \begin{array}[t]{l} \mathsf{input}\ \varphi \in L_{\Sigma_1 \cup \Sigma_2}(P) \\ \mathsf{if}\ \mathsf{head}(\varphi) \notin \Sigma_2\ \mathsf{then}\ i \leftarrow 1\ \mathsf{else}\ i \leftarrow 2 \\ \mathsf{let}\ V_i = V \cup \{x_0\}, M = \mathsf{Mon}_{\Sigma_i}(\varphi), \psi = \mathsf{skel}_{\Sigma_i}(\varphi) \\ \mathsf{output}\ \mathsf{A}^i(V_i \cup \{x_{e^{-1}(\varphi')} : \varphi' \in M, \mathsf{A}(V)(\varphi') = \mathsf{yes}\})(\psi). \end{array} \end{array}$$

On input V, A outputs the algorithm $\mathsf{A}(V) = \mathsf{D}$, which uses V to compute $\mathsf{D}(\varphi)$ on input φ. Note that D does some syntactic analysis on the input formula, such as computing its head, monoliths, and skeleton. All these constructions are computable, given their definitions, and the assumption that Σ_1 and Σ_2 are decidable. For this reason, one can test membership in a signature, and compute a full enumeration $e : \mathbb{N} \to L_{\Sigma_1 \cup \Sigma_2}(P)$, and its inverse.

We claim that $\mathcal{L}_1 \bullet \mathcal{L}_2$ is finite var-decidable with algorithm A. We show, by induction on the structure of $\varphi \in L_{\Sigma_1 \cup \Sigma_2}(P)$ that, for all finite $V \subseteq P$, $\mathsf{A}(V)(\varphi)$ terminates, and

$$\mathsf{A}(V)(\varphi) = \begin{cases} \mathsf{yes} & \text{if } V \vdash_{12} \varphi\,, \\ \mathsf{no} & \text{if } V \nvdash_{12} \varphi\,. \end{cases}$$

For the induction base, let $\varphi = p \in P$. Obviously $\mathsf{head}(p) = p \notin \Sigma_2$ so $i = 1$, and $V_1 = V \cup \{x_0\}$. Moreover, $M = \mathsf{Mon}_{\Sigma_1}(p) = \emptyset$ and $\psi = \mathsf{skel}_{\Sigma_1}(p) = \varphi$. Easily, $V_1 \cup \{x_{e^{-1}(\varphi')} : \varphi' \in M, \mathsf{A}(V)(\varphi') = \mathsf{yes}\} = V_1 = V \cup X^1_V(p)$, as set in Definition 3.11, because the fibered logic has theorems and thus $X^1_V(p) = \{x_0\}$. Hence, in the last line, $\mathsf{A}^1(V_1)(p)$ tests precisely whether $V, X^1_V(p) \vdash_1 p$. Since $\mathcal{L}_1$ is finite var-decidable with A^1, the algorithm terminates. As $(V \cup \{x_0\})^{\vdash_2} \neq L_{\Sigma_1 \cup \Sigma_2}(P \cup X)$, or $\mathcal{L}_2$ would be trivial, Proposition 3.13 guarantees that the answer is yes if $V \vdash_{12} \varphi$, and no if $V \nvdash_{12} \varphi$.

For the induction step, let $\varphi = c(\varphi_1, \ldots, \varphi_n)$ for some n-place connective $c = \mathsf{head}(\varphi) \in \Sigma_1 \cup \Sigma_2$. Clearly $i = 1$ if $c \in \Sigma_1$, and $i = 2$ if $c \in \Sigma_2$, and $V_i = V \cup \{x_0\}$. Moreover, $M = \mathsf{Mon}_{\Sigma_i}(\varphi) = \bigcup_{j=1}^n \mathsf{Mon}_{\Sigma_i}(\varphi_j)$ and $\psi = \mathsf{skel}_{\Sigma_i}(\varphi)$. By induction hypothesis, for each $\varphi' \in M$, $\mathsf{A}(V)(\varphi')$ terminates and $\mathsf{A}(V)(\varphi') = \mathsf{yes}$ precisely if $V \vdash_{12} \varphi'$ and so, easily, $V_i \cup \{x_{e^{-1}(\varphi')} : \varphi' \in M, \mathsf{A}(V)(\varphi') = \mathsf{yes}\} = V \cup X^i_V(\varphi)$, as set in Definition 3.11, because the fibered logic has theorems and $x_0 \in X^i_V(\varphi)$. Hence, in the last line, $\mathsf{A}^i(V \cup X^i_V(\varphi))(\mathsf{skel}_{\Sigma_i}(\varphi))$ tests whether $V, X^i_V(\varphi) \vdash_i \mathsf{skel}_{\Sigma_i}(\varphi)$. Since $\mathcal{L}_i$ is finite var-decidable with A^i, the algorithm terminates. As $(V \cup \{x_0\})^{\vdash_{3-i}} \neq L_{\Sigma_1 \cup \Sigma_2}(P \cup X)$, or $\mathcal{L}_{3-i}$ would be trivial, Proposition 3.13 guarantees that the answer is yes if $V \vdash_{12} \varphi$, and no if $V \nvdash_{12} \varphi$. □

Let us now have a quick look at the complexity of the decision procedure.

Proposition 4.3 Let A be the decision algorithm for $\mathcal{L}_1 \bullet \mathcal{L}_2$ defined in the proof of Proposition 4.2, using algorithms A^1 and A^2 for deciding $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively. The running time of $\mathsf{A}(V)(\varphi)$ for inputs $V \subseteq P$ and $\varphi \in L_{\Sigma_1 \cup \Sigma_2}(P)$ is given by a function f such that

$$f(m,n) \leq n \times \max(f^1(m+n,n), f^2(m+n,n))$$

where $m = \mathsf{size}(V)$, $n = \mathsf{size}(\varphi)$, and f^1, f^2 are the running time functions corresponding to algorithms $\mathsf{A}^1, \mathsf{A}^2$.

Proof The proof, for a given finite $V \subseteq P$ and $\varphi \in L_{\Sigma_1 \cup \Sigma_2}(P)$, is done by induction on the structure of φ.

For the induction base, let $\varphi = p \in P$. Note that the algorithm sets $i = 1$, $V_1 = V \cup \{x_0\}$, $M = \mathsf{Mon}_{\Sigma_1}(p) = \emptyset$, and $\psi = \mathsf{skel}_{\Sigma_1}(p) = p$. Hence, we have only to consider the running time of $\mathsf{A}^1(V_1)(p)$. If $m = \mathsf{size}(V)$, as $\mathsf{size}(p) = 1$, we have that $\mathsf{size}(V_1) \leq m + 1$, and the running time of the algorithm is

$$f(m,1) = f^1(m+1,1) \leq \max(f^1(m+1,1), f^2(m+1,1)).$$

For the induction step, let $\varphi = c(\varphi_1,\ldots,\varphi_k)$ for some k-place connective $c = \mathsf{head}(\varphi) \in \Sigma_1 \cup \Sigma_2$. Note that the algorithm sets, $i = 1$ if $c \in \Sigma_1$ and $i = 2$ if $c \in \Sigma_2$, and $V_i = V \cup \{x_0\}$. If $m = \mathsf{size}(V)$, and $n = \mathsf{size}(\varphi)$, note that $\mathsf{size}(V_i) \leq m + 1$, $\mathsf{size}(M) \leq n - 1$, and $\mathsf{size}(\mathsf{skel}_{\Sigma_i}(\varphi)) \leq n$. Note also that $V_i \cup \{x_{e^{-1}(\varphi')} : \varphi' \in M, \mathsf{A}(V)(\varphi') = \mathsf{yes}\} = V \cup X^i_V(\varphi)$, whose size is bounded by $m + 1 + (n-1) = m + n$. As we have to consider the running time of $\mathsf{A}^i(V \cup X^i_V(\varphi))(\mathsf{skel}_{\Sigma_i}(\varphi))$, and also of all $\mathsf{A}(V)(\varphi')$ for $\varphi' \in M$, using the induction hypothesis and the natural monotonicity of the running time functions, the total running time of the algorithm is given by

$$\begin{aligned}
&f(m,n) \leq \\
&f^i(m+n,n) + \textstyle\sum_{\varphi' \in M} f(m,n') = \\
&f^i(m+n,n) + \textstyle\sum_{\varphi' \in M} (n' \times \max(f^1(m+n',n'), f^2(m+n',n'))) \leq \\
&f^i(m+n,n) + \textstyle\sum_{\varphi' \in M} (n' \times \max(f^1(m+n,n), f^2(m+n,n))) = \\
&f^i(m+n,n) + \left(\textstyle\sum_{\varphi' \in M} n'\right) \times \max\left(f^1(m+n,n), f^2(m+n,n)\right) = \\
&f^i(m+n,n) + \mathsf{size}(M) \times \max\left(f^1(m+n,n), f^2(m+n,n)\right) \leq \\
&f^i(m+n,n) + (n-1) \times \max\left(f^1(m+n,n), f^2(m+n,n)\right) \leq \\
&n \times \max\left(f^1(m+n,n), f^2(m+n,n)\right),
\end{aligned}$$

where, for improved readability, we used n' as an abbreviation of $\mathsf{size}(\varphi')$. ☐

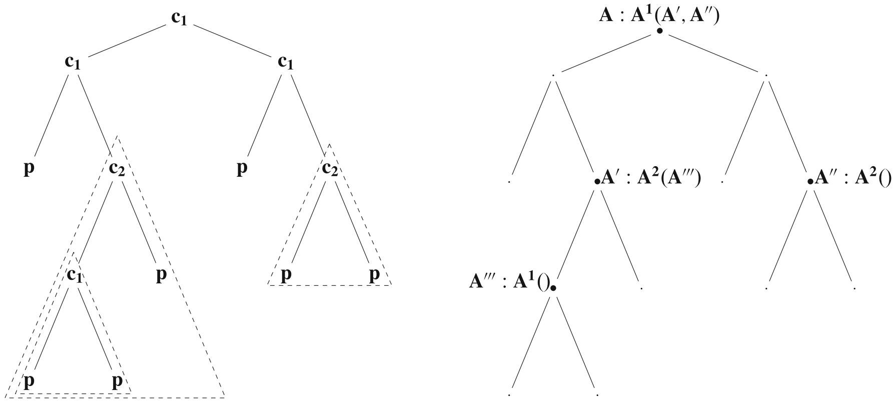


Fig. 2 Deciding a mixed formula

The upperbound for the running time of the decision procedure that we established in Proposition 4.3 is better understood if we look at the scheme shown in Fig. 2. We are assuming a simple scenario where c_1 is a 2-place connective of $\mathcal{L}_1$, and c_2 a 2-place connective of $\mathcal{L}_2$. On the left-hand side of the figure, we have the syntactic tree of the combined formula $c_1(c_1(p, c_2(c_1(p, p), p)), c_1(p, c_2(p, p)))$. The dashed triangles indicate the monolithical structure of the formula. Since the head of the formula is in the signature of logic $\mathcal{L}_1$, triangles appear as soon as one finds the c_2 connective when traveling down the tree. Inside each triangle, the same repeatedly applies. On the right-hand side of the figure we see how the decision algorithm stretches along the tree structure. At the root note, A is executed, recursively deploying other executions of the algorithm at the vertices of the outermost triangles, corresponding to its monoliths, denoted by A' and A''. The outputs of A' and A'' are then used at the root node to execute A^1, which explains the annotation $\mathsf{A} : \mathsf{A}^1(\mathsf{A}', \mathsf{A}'')$. The same structure is propagated downwards. Executing A' deploys A''' at the root of the inner triangle, and then uses its output on A^2. On its turn, A''' simply calls A^1, as there are no further monoliths to be analyzed. The same happens with execution A'', on the right side. This example shows well that the running time of the decision procedure is bound by the size of the formula (i.e., the number of nodes of the tree) times the running time of each call to either A^1 or A^2, as at most one such call is necessary at each node.

Clearly, this means that the decision problem for the fibered logic reduces polynomially to the decision problems of the logics given.

Example 4.4 Consider again the unconstrained fibring $\mathcal{L}_{\mathsf{cls}} \bullet \mathcal{L}_{\mathsf{int}} = \langle \Sigma_{\mathsf{cls}} \cup \Sigma_{\mathsf{int}}, \vdash \rangle$ that we have used in Example 4.1. It is well known that both logics are weakly decidable, and also finitary, which makes them finitely var-decidable as both enjoy the deduction theorem. The corresponding decision problems are known to be in coNP for $\mathcal{L}_{\mathsf{cls}}$, and PSPACE for $\mathcal{L}_{\mathsf{int}}$. As a direct application of Propositions 4.2 and 4.3, we can conclude that $\mathcal{L}_{\mathsf{cls}} \bullet \mathcal{L}_{\mathsf{int}} = \langle \Sigma_{\mathsf{cls}} \cup \Sigma_{\mathsf{int}}, \vdash \rangle$ is also finitely var-decidable, and that the corresponding decision problem is in PSPACE. △

5 Conclusion

We have studied in detail the theoremhood problem for logics obtained by unconstrained fibring, and shown that it is decidable provided that the component logics enjoy the slightly stronger notion of finite var-decidability. The result we obtained is the first of its kind, thus opening the way to formal tool support for fibered logics in a neat and modular way. We also showed that the complexity of the theoremhood decision problem for the fibered logic is essentially the same as the complexity of the finite var-decidability problem of the hardest component logic. Although distinct in nature, it is worth mentioning that our decision algorithm bears some similarities with the Nelson–Oppen approach to deciding joint equational theories [13], a connection that may be worth exploring.

Still, the result we obtained is not fully satisfactory, as it does not allow one to conclude much about the finite decidability of the fibered logic. Indeed, the key result used in the decidability proof, Proposition 3.13, is simply not strong enough. It seems to be possible to extend the proposition to deal with sets of nonmixed hypotheses instead of only sets of variables, but reasoning under arbitrary sets of hypotheses seems to be out of the reach of the techniques we have used. In any case, it should be remarked that extending Proposition 3.13 to deal not just with sets of variables as hypotheses but also with sets of nonmixed hypotheses would certainly help shedding new light over the fibring construction, with possible applications going beyond decidability. Namely, such a result may be fundamental for solving the long standing problem of conservativity of fibring, at least in the unconstrained case. We are working towards such a result.

Acknowledgment The authors acknowledge the support of EU FP7 Marie Curie PIRSES-GA-2012-318986 project GeTFun: Generalizing Truth-Functionality, and the FEDER/FCT project PEst-OE/EEI/LA0008/2013 and PTDC/EIA-CCO/113033/2009 ComFormCrypt of SQIG-IT. The first author also acknowledges the support of FCT scholarship SFRH/BPD/76513/2011.

References

1. Beckert, B., Gabbay, D.: Fibring semantic tableaux. In: Proceedings of the International Conference on Automated Reasoning with Analytic Tableaux and Related Methods, TABLEAUX'98, pp. 77–92. Springer, Berlin, Heidelberg (1998)
2. Béziau, J.-Y.: Universal Logic. In: Childers, T., Majers, O. (eds.) Proceedings of the VIII International Symposium, LOGICA'94, pp. 73–93. Czech Academy of Science, Prague, CZ (1994)
3. Béziau, J.-Y.: The challenge of combining logics. Log. J. IGPL **19**(4), 543 (2011)
4. Caleiro, C.: Combining Logics. PhD thesis, IST, Universidade Técnica de Lisboa, PT (2000). http://sqig.math.ist.utl.pt/pub/CaleiroC/00-C-PhDthesis.ps
5. Caleiro, C., Carnielli, W., Rasga, J., Sernadas, C.: Fibring of logics as a universal construction. In: Gabbay, D., Guenthner, F. (eds.) Handbook of Philosophical Logic, 2nd edn., vol. 13, pp. 123–187. Springer, Dordrecht (2005)
6. Caleiro, C. Ramos, J.: Combining classical and intuitionistic implications. In: Konev, B., Wolter, F. (eds.) Frontiers of Combining Systems 07, Lecture Notes in Artificial Intelligence, pp. 118–132. Springer, Berlin, Heidelberg (2007)
7. Caleiro, C., Ramos, J.: From fibring to cryptofibring: a solution to the collapsing problem. Log. Univers. **1**(1), 71–92 (2007)

8. Caleiro, C., Sernadas, A.: Fibring logics. In: Béziau, J.-Y. (ed.) Universal Logic: An Anthology (From Paul Hertz to Dov Gabbay), pp. 389–396. Birkhäuser, Basel (2012)
9. Carnielli, W., Coniglio, M., Gabbay, D., Gouveia, P., Sernadas, C.: Analysis and Synthesis of Logics: How To Cut And Paste Reasoning Systems. Applied Logic, vol. 35. Springer, Dordrecht (2008)
10. Coniglio, M., Sernadas, A., Sernadas, C.: Preservation by fibring of the finite model property. J. Log. Comput. **21**(2), 375–402 (2011)
11. Gabbay, D.: Fibred semantics and the weaving of logics, part 1: Modal and intuitionistic logics. J. Symb. Log. **61**(4), 1057–1120 (1996)
12. Gabbay, D.: Fibring Logics. Oxford Logic Guides, vol. 38. Clarendon Press, Oxford (1999)
13. Nelson, G., Oppen, D.: Simplification by cooperating decision procedures. ACM Trans. Program. Lang. Syst. **1**(2), 245–257 (1979)
14. Rasga, J., Sernadas, A., Sernadas, C., Viganò, L.: Fibring labelled deduction systems. J. Log. Comput. **12**(3), 443–473 (2002)
15. Sernadas, A., Sernadas, C., Caleiro, C.: Fibring of logics as a categorial construction. J. Log. Comput. **9**(2), 149–179 (1999)
16. Sernadas, C., Rasga, J., Carnielli, W.: Modulated fibring and the collapsing problem. J. Symbol. Log. **67**(4), 1541–1569 (2002)
17. Wójcicki, R.: Theory of Logical Calculi. Kluwer, Dordrecht (1988)
18. Zanardo, A., Sernadas, A., Sernadas, C.: Fibring: Completeness preservation. J. Symbol. Log. **66**(1), 414–439 (2001)

S. Marcelino (✉) · C. Caleiro · P. Baltazar
SQIG – Instituto de Telecomunicações
Dep. Matemática – Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal
e-mail: sergiortm@gmail.com

On Rules and Refereeing in Football

Amirouche Moktefi and Fabien Schang

Abstract Recent developments in international football governance seem to be progressively leading toward an increasing use of technological devices for refereeing purposes. Opponents to change are often portrayed as old-fashioned and conservative. Philosophy might be of some help to overcome the dispute. In this paper, we first explore several concepts that are central to the current debate on football refereeing. Then, we determine the business of referees in relation to rules. We assess different arguments displayed regarding the role of chance and skills in competitions. Finally, we argue for the idea of referees as full players in football games.

Keywords Rules · Rule-following · Football · Referee · Technology-assisted refereeing · Rationality in sports · Chance · Skills · Competition · Game · Disagreement

Mathematics Subject Classification 03A99

Clichés argue against a proper philosophy of football: its fans are seldom taken seriously by the most respectable customers of philosophical literature, although some attempts to develop a philosophy of football have recently been made [11]. This paper is meant as a step forward towards a philosophy of football. The idea is to consider the relations between rules and refereeing in football and carry on with some thoughts developed by Tarcisio Pequeno and Jean-Yves Béziau on the relationship between games and rules.

Pequeno and Béziau [8] offered a set of five necessary conditions to make a game, namely: *competition*, *fiction*, *chance*, *skill* and *fun*. Two of them (*fun* and *fiction*) might be disputed and will not be considered in our paper. Although one might expect to have fun by playing a game, we do not consider it to be necessary to make the game. Also, fiction does not seem to be a necessary criterion of games outside the restricted area of simulation games, and as such is not relevant for our purpose. Hence, we would like here to stress the concepts of *competition*, *chance* and *skill*, which we consider to be central in football, notably in regard to the issue of refereeing. A patient and careful clarification of these concepts might bring some light upon the philosophical problem of rule-following and refereeing in sports. We will particularly pay attention to the modern debates on technology-assisted refereeing in football. Indeed, high claims are regularly made for technological assistance to improve football refereeing and prevent errors. Opponents to

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_19

reform are often portrayed as old-fashioned and conservative. It is merely a slight exaggeration to affirm that the champions of technology in sports accuse their challengers of opposition to truth, and hence of favouring failure and vice against perfection and virtue. We would like to show that this accusation actually relies upon a misunderstanding of the very idea of football. For this purpose, we will explore the scope of several concepts that are central to the current dispute on football refereeing. As such we aim at providing a common language in order to organize these debates and create better conditions for a dialogue between the supporters and the opponents of technology-assisted refereeing. It is hoped that this work will contribute to the emergence of a serious philosophy of football.

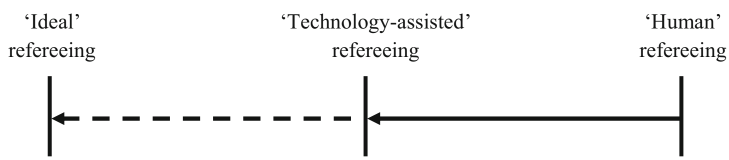
1 Toward Technology-Assisted Refereeing

Recent developments in international football governance seem to progressively lead toward an increasing use of technological devices for refereeing purposes. Sepp Blatter, current president of FIFA (the governing body of international football), openly expressed his desire to pursue this path if he were to win the next FIFA elections in 2015. Actually, several national football associations urged FIFA to accelerate refereeing reforms and offered to test refereeing technologies in their local football competitions. Change might be slow however. Indeed, it must be kept in mind that FIFA is not the body that determines football rules. Historical contingencies make that the game is ruled by the International Football Association Board (IFAB), where FIFA has only half of the votes, while the other half belongs to the four football associations of the United Kingdom. FIFA itself and IFAB are genuinely known as conservative bodies, and it is not granted that new rules on technology assistance will get the votes required for their adoption. Moreover, one should not underestimate resistance to reform among the football community. Michel Platini, President of the Union of European Football Associations (UEFA), openly opposes technology assistance and favours adding more 'human' referees, if needed.

Platini's suggestion was actually introduced in recent international competitions, where two additional referees were located behind the goals. This solution did not prevent the occurrence of 'ghost goals' (legitimate goals that were denied), however. For instance, Ukraine was refused a legitimate goal in its game against England during the 2012 UEFA European Championship, despite the action happening right under the eyes of the 'additional' referee. Ironically, 2 years earlier England was refused a legitimate goal against Germany and was eventually excluded from the 2010 FIFA World Cup tournament. These 'incidents' made the FIFA feel it necessary to introduce a new technological device during the recent World Cup that took place in Brazil (12 June–13 July, 2014) to test whether the ball crossed the goal line. This step forward might prove to be a turning point in FIFA policy by opening the way to further technology assistance despite the controversies that it might raise [5, 9].

Controversy is no stranger to football anyway. People argue, agree and disagree about it all around the world. Unsurprisingly, the fairness and objectivity of refereeing is one of the favourite objects of dispute among football fans. Although refereeing errors depend partly on the referee's technical and psychological skills (on awarding penalties, see [10]), the very occurrence of errors is inevitable, as some decisions require unusual optical and perceptual skills (on judging offside, see [1]). Does this imperfection of human refereeing entail the necessity of technology assistance? This question is certainly an interesting challenge to the philosopher. It offers an interesting switch in the perspective: the issue is not to be viewed as a mere technological problem with philosophical dimensions, but rather as a philosophical problem in its own with some technological dimensions. This switch is necessary to accurately address the problem of technology-assisted refereeing. It makes the question 'how to improve football refereeing' depend on 'what football and what refereeing we want to have' and that is what we believe is the appropriate order.

One rough way of expressing the controversy on technology-assisted refereeing is to say that some think that technology assistance will improve refereeing, while others think it will not. We should keep in mind that this is a simplification, however, because we actually have a wide range of positions depending on 'how much' technology one introduces and what use is made of it. Still, the rough formulation has the advantage of pointing out the main issue at dispute: does technology improve refereeing? During the Brazil World Cup, we regularly surveyed the arguments displayed on two major and rival French radio stations: RTL and RMC. Both stations devoted large slots to the tournament, including live broadcasting of the games and talk shows to debrief and analyse the games with football experts, practitioners, journalists and fans. The former station (RTL) team mostly defended arguments in favour of technology-assisted refereeing, while the latter (RMC) strongly and regularly expressed its opposition to technology assistance and reluctantly commented on refereeing decisions at all. A survey of the opposed arguments that were displayed provides the following picture of the dispute:



Let us postulate some ideal refereeing (where errors do not occur) achieved by an omniscient and objective referee. He is omniscient in the sense that he sees every action in the game and does not miss anything. He is objective in the sense that he bases his decisions only on the complete information that he has and any other referee who would have the same amount of information would make the same decision. There seems to be a kind of general agreement that such an ideal refereeing is impossible and that human refereeing is far from being infallible. It also seems to be a large consensus as to the fact that technology-assisted refereeing will not reach that ideal refereeing, although it still is closer to it than human refereeing. So, the disagreement does not concern as much the quality of technology-assisted refereeing as the opportunity of appealing to it. In a sense, opponents highlight the *impossibility* of filling the gap between technology-assisted refereeing and ideal refereeing, while supporters make high claims for the *possibility* of filling the gap between human refereeing and technology-assisted refereeing, thereby shortening the gap between achievable refereeing and ideal refereeing. So everyone agrees that refereeing will never be ideal (i.e. supposedly optimal and fair) and that technology-assisted refereeing is theoretically better (i.e. more optimal and fair) than human refereeing. The disagreement is about the opportunity of changing the refereeing practicalities for the sake of this incomplete improvement. One way to make both parts agree would be to define some 'good use' that could be made of technology in refereeing. There is no agreement so far but there is a *theoretical* possibility to make the two parts agree, even on a minimal scale.

Of course, in practice, it is more complicated. Still the disagreement is not a necessary one. There is hope of reaching some conciliation. One reason to keep hope is that actually both parts already seem to agree on the importance of not appealing to technology all along the game. This 'good use' restricts the appeal to technology to 'decisive moments' in the game where a team challenges the referee's decision. As such technology is viewed as a means to establish the facts and consequently eliminate the disagreements and reach consensus in litigious situations.

The process at work in the enhancement of refereeing by technology-aid might be called an *objectivation*. It is the idea that technology provides further information on the contentious action and as such increases the chances of settling the matter. Two referees who disagree at first might agree after consulting the technological aid (for instance, a TV screening of the action from different perspectives). Actually, the process of decision-making might even be enlarged to a wide audience by involving the viewers of the game, who would then act as witnesses [2]. In a sense, this objectivation establishes the 'fact' by removing its dependence on the perception of the referee alone. Such a 'nothing-to-lose' use of technology aims at reaching some objectivity and looking for consensual decisions when disagreements occur. The quality of the refereeing will then be assessed by the agreement that it settles between the opposite teams. A 'good refereeing' makes the teams agree and leads to no dispute. Such a 'good refereeing' is not optimal, however. Indeed, appealing to technology in instances where there was no disagreement might also reveal errors that would have not been detected in the 'good refereeing' mode. Agreement is not a guarantee of correctness. As such, the 'good use' of technology does not aim at an optimization of refereeing (reduction of errors) but just its *objectivation* (reduction of interpretation).

2 What Referees Do

An analogy can be made with court ruling, in this respect. Technology-assistance advocates tend to depict the referee as a judge whose task is to make the best decision in a conflict between two sides with opposed interests. No one stands against the use of scientific devices to unveil the truth in a crime and punish the culprit. Since mistaken decisions in court as well as in football are unfair, one would expect the same disposition to introduce technological assistance in football in order to offer the same support that scientific assistance does in court. This depiction makes settling disagreements the business of refereeing decision-making. It puts the referee in the position between the opposed teams, a place where it is common to find the referee on the field when disputes occur. This observation would be misleading, however. Indeed, the decisions of the referee are not the least concerned with the disagreement of the teams. The FIFA guide states that 'Each match is controlled by a referee who has full authority to enforce the laws of the game in connection with the match to which he has been appointed' [4, p. 24]. So the referee stands between the game and its laws, not between the opposed players. He might, for instance, very well make decisions against the will of both teams as long as those decisions are made in accordance with the rules of the game. Suppose that the ball went out of and came back onto the field. It is the duty of the referee to stop the game even if both teams want to continue playing.

One way of making the point is to apply Hintikka's distinction between definitory and strategic rules: 'Definitory [rules] define the basic moves in the game. They tell us what is and what is not admissible in a game. Following them [. . .] assures that you are playing the game correctly [. . .] Strategic rules [are] the rules which give insight into how to play a game successfully and well' [6, pp. 32–33]. The referee is concerned with definitory rules. The strategic rules of the teams to win the game do not affect his refereeing (for a history of football strategies, see [12]). The referee just assures that players play the game *accurately*, independently of how *successfully* they play. As such, refereeing is not about making the game better, but just about making it accurate, *independently* of its consensus potential. Pequeno and Béziau [8] introduced a slightly different typology of rules: framework rules (material conditions), deontic rules (permitted and forbidden actions), teleological rules (goal of the game) and strategic rules (technical modes of playing). The latter are broadly understood in the same sense as Hintikka's strategic rules.

Framework and deontic rules also broadly cover Hintikka's definitory rules. Teleological rules to determine the winner of the game are more difficult to work but do not affect our current discussion. Football's teleological rules are so 'analytical' (the team that scores the more, wins) that the referee does not need to care about them. Now the game has to be played correctly in order to be played successfully. Indeed, both teams have to be subject to the same definitory rules, otherwise that would infringe an essential condition of the game, i.e. competition, which grants equity among players [8]. This precondition is also necessary to secure the legitimacy of the winner which supposedly is the team whose skills have been better than its opponent's in the game. Hence, the business of the referee is not to settle disagreements between the teams. It rather is to secure that the definitory rules of the game are insured regardless of how the teams do react to his decisions.

This rule-formalist approach, assuming rules to suffice in order to rule a game, might be disputed on the ground that the referee might have to improvise in situations where the rules are not that clear as to what to do [7]. Actually, referees might even rule against the rules in case of 'minor infringements' in order to save the flow of the game. As such, the referee would privilege the 'spirit' rather than the 'letter' of the law [3]. Our objection to this reading is that it misses the 'big picture' by limiting the focus on the two teams playing. Generally, games take place in tournaments, where every team is not opposed to just one other team but rather to several others. Any game has consequences on the others, and the issue at stake is not just the winner of the game but a whole set of outputs, including card penalties, injuries, goal average, etc. As such, a permissive refereeing in a given game would be unfair to other teams who could get benefits from potential decisions. Of course, there might be general recommendations for the whole tournament, but then explicating them would make them belong to the letter rather than the spirit. Naturally, one still has to observe how this body of guidelines could be applied by the referee. For instance, if it is recommended at the beginning of a competition that injurers should be severely punished by the referee, one still has to figure out whether a concrete case lies within this recommendation.

3 On Chance and Skills

When refereeing errors occur, they might allow a team to be successful despite violating a definitory rule. That's what makes the game look unfair. However, the very possibility of the referee making errors seems to be involved in the definitory rules themselves, so that this possibility applies on both teams and as such might be disadvantageous (but not necessarily unfair) to one team over the other. The admission of errors in the game simply introduces an element of *chance* into the game. This should not be confused with unfairness, which applies only in games where skills alone determine the winner. Further light can be thrown on this issue by means of the three concepts alluded to earlier: *skill*, *chance* and *competition* [8].

Skill is developed in order to strengthen the probability of victory in a ruled competition. Without this general rationale in mind, it would barely make sense for football players to train hard for a competition where skills do not determine the winner. Disputes on football refereeing seem to originate in this blatant incompatibility between skill and chance. At first, one would expect skills to be the only element to decide between opposed teams. As such, football referees are not expected to affect the determination of the winner, since this decision has its own rule (the winner is the team that scored more than the other). The referee is just asked to secure that the competition takes place in accordance with the rules. The slightest suspicion as to the influence of the referee in the making of the score might spoil the acknowledgment of the winner among football fans.

For instance, West Germany's victory at the 1990 World Cup is often underestimated on the grounds that the team was awarded a 'generous' penalty in the final game against Argentina. Ironically, Maradona's 'Mano de Dios' in the 1986 World Cup usually does not affect the recognition of the Argentinian final victory among football fans. Of course, it is admitted that Maradona unduly made use of his hand to overcome the goalkeeper's fist and score a goal that would happen to be decisive during the quarter final game against England. Still, fans would say that Argentina would have won anyway and that they actually deserved to win for their offensive playing style and the genius of their players. Contrariwise, Greece is said not to deserve its victory at the 2004 UEFA European Championship because of its 'boring' defensive playing style. These concerns have nothing to do with rules, however, and could in no way be seriously considered in a general reflection about refereeing in football. As skilled as Argentina might have undoubtedly been, their genius cannot stand as a compensatory argument to justify infringements of rules. Moreover, as boring as Greece might have admittedly been, they played in accordance to the rules and as such deserved their victory. Further discussion on the matter is useless for our present concern.

In football games, the referee takes care of definitory rules. He makes decisions to secure that those rules are granted. Should a referee's decisions be challenged by unhappy players? In the present state, it might look inadequate since players do not contribute to the decision-making. The football referee stands alone as the supreme and unique authority in making decisions about rule-following. In this respect, football refereeing differs from other sports like rugby or tennis, where players can challenge decisions and ask for the referee to reconsider his decision (after screening a video replay or appealing to 'hawk-eye' technology, for instance). As true as this claim might be, it needs some adjustment. Firstly, it must be kept in mind that football refereeing practices are ruled too, and that teams do insure that referees do not infringe their own definitory rules. As such, they can challenge refereeing as long as it is not about the very judgement of the referee, but rather about the application of that judgement. Then, although rugby and tennis do appeal to technology assistance, final decisions are still held by referees, who continue to stand as the supreme guarantee of the correctness of rule-following. Technology assistance in rugby and tennis aims at reducing the element of chance in order to favour the skills of players, while chance is, in the present state, kept as an essential constituent of the very idea of football practice.

The acceptance of chance (together with competition and skills) as a desirable element of football has strong implications on refereeing policy. Indeed, it trivializes the referee's fallibility and makes the appeal to technology unnecessary. It might be objected that modern football already reduces the elements of chance through the design of the material conditions of its practice (stadiums, shoes, balls, etc.). The introduction of more human referees, championed by Platini and other technology opponents, might be viewed as a way to minimize chance in football refereeing as well. Hence, the introduction of technology assistance would only join this long list of reforms made in football history to make it better. Moreover, that would not definitely eliminate the element of chance, since there will always be uncontrollable elements, notably the natural conditions of the game. Playing under bad weather conditions might have a crucial influence on a game. For instance, the wind alters the path of the ball and makes it difficult to grasp for. The direction of the wind becomes crucial and might favour one team over the other. Such natural conditions interfere with the game and make chance an important element that the teams and the referee have to deal with, as long as football practice is possible. The acceptance of chance as an element of football does not entail its desirability, however, and it remains legitimate to wonder whether it is still suitable to reduce that element as much as possible to favour skills. Hence, the introduction of technology assistance in football refereeing is viewed as an improvement for those who object to the element of chance and view it as undesirable. All others will consider technology to actually alter the very idea of the game. As such, technology assistance would be a wish or a will rather than a need or a necessity.

4 The Referee as a Full Player

Football as it presently is, i.e. a game that involves both skill and chance, embodies the idea of a referee as a full player of the game. Although he does not compete with the teams in the same game, he still plays at his best to secure the rules. The teams themselves and football associations insure that the referee plays *correctly* (in accordance with his definitory rules). Finally, the referee himself devises and appeals to strategic rules to play the game *successfully*. The referee's *moves* in the game are independent of the agreement of the teams, since the correctness of the game is irrespective of the players' opinions. The best referees are asked by football associations to keep playing at the same 'stage' of the game and might even be promoted to the next stage. Worse referees are retrograded to the lower stage, since they did not prove to have the skills to play at the current stage. From this perspective, it appears that technology assistance would not be anything other than a dummy solution to what turns out not to be a problem in the first place. One might similarly imagine that players would wear some kind of tele-controlled shoes or electronic viewfinders to improve their ability to score goals! Neither solution improves the human skills of the players (including the referee). An obvious objection to this reading, already alluded to above, is that such improvements are already at work in football practice (stadiums, balls, etc.). The issue then becomes a matter of balance. How much change can we introduce in football rules without altering its very nature? This brings us again to our main point: the introduction of technology assistance is not a need, it is a will. This is not meant to reject the possibility of introducing technology for refereeing purposes. It simply states that if there are any reasons for such a reform, they are not to be found in the nature of the game. Football governance might need technology-assisted refereeing, football itself does not.

Acknowledgment This paper draws upon research supported by the Estonian Research Council PUT 267 'Diagrammatic mind: logical and communicative aspects of iconicity' (Principal investigator Professor Ahti-Veikko Pietarinen) and by the basic research program at the National Research University Higher School of Economics (Moscow, Russia). A previous version of this paper was presented at the Faculty Seminar on Philosophy in the Ragnar Nurkse School of Innovation and Governance (Tallinn University of Technology, 24 September, 2014) and at the 10th Annual Estonian Philosophy Conference (University of Tartu, 25–27 September, 2014). We would like to express our gratitude to the participants who attended our talks for their remarks and suggestions that contributed to the shaping of this paper.

References

1. Catteeuw, P. et al.: Offside decision making in the 2002 and 2006 FIFA world cups. J. Sports Sci. **28**(10), 1027–1032 (2010)
2. Collins, H.: The philosophy of umpiring and the introduction of decision-aid technology. J. Philos. Sport **37**, 135–146 (2010)
3. Colwell, S.: The 'letter' and the 'spirit': football laws and refereeing in the twenty-first century. Soccer Soc. **1**(1), 201–214 (2000)
4. FIFA: The Laws of the Game 2014/2015. FIFA, Zurich (2014)
5. Fleming, N.: Tell fans definitive calls are an impossible goal. Nature **497**, 537 (2013)
6. Hintikka, J., Bachman, J.: What If ... ? Toward Excellence in Reasoning. Mayfield Publishing, Mountain View, CA (1991)
7. McFee, G.: Fairness, epistemology, and rules: a prolegomenon to a philosophy of officiating. J. Philos. Sport **38**(2), 229–253 (2011)
8. Pequeno, T., Béziau, J.-Y.: Rules of the game. In: Béziau J.-Y., Coniglio, M.E. (eds.) Logic without Frontiers, pp. 121–134. College Publication, London (2011)
9. Pielke Jr., R.: How innovations such as goal-line technology make sport better. PoliticalScience blog (hosted by The Guardian), 5 June 2013

10. Plessner, H., Betsch, T.: Sequential effects in important referee decisions: the case of penalties in soccer. J. Sport Exerc. Psychol. **23**, 254–259 (2001)
11. Richards, T. (ed.): Soccer and Philosophy. Open Court, Chicago & La Salle, IL (2010)
12. Wilson, J.: Inverting the Pyramid: the History of Football Tactics. Orion, London (2014)

A. Moktefi (✉)
Chair of Philosophy, Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Tallinn, Estonia
e-mail: amirouche.moktefi@ttu.ee

F. Schang
Higher School of Economics, Department of Philosophy, National Research University, Moscow, Russian Federation
e-mail: schang.fabien@voila.fr

Arrow-Hexagons

Alessio Moretti

Abstract Oppositional geometry gives a mathematical model of oppositional phenomena through "oppositional structures" (logical squares, hexagons, cubes, ...). It's so far known formal entities, the backbone of which are the "oppositional bi-simplexes (and poly-simplexes) of dimension *m*", are distributed into three families (the α-, β- and γ-structures). However, some recent studies by different authors exhibit strange structures, notably *strange variations* of the notion of "oppositional hexagon" (or "logical hexagon"). In this paper we show that inside the oppositional tetrahexahedron, i.e. the $\beta 3$-structure (discovered in 1968 and rediscovered in 2008) – a 3-D solid made of a logical cube and 6 logical "strong hexagons", containing 14 vertices and 36 implication arrows – there are in fact $C^6_{14} = 30\,030$ strange hexagons, which we call "hybrid hexagons". In this paper, through a systematic study of those among them that have as invariant property a regular perimeter made of alternated arrows (henceforth "arrow-hexagons"), we show that they divide into a much smaller number of families, nine, each containing several isomorphic instances of the same oppositional pattern. An interesting result seems to be that when seen from the viewpoint of their mutual transformations (i.e. moving from one to another kind of arrow-hexagon, just by exchanging one of its 6 vertices with one among the remaining $14-6 = 8$ vertices of the tetrahexahedron), these arrow-hexagonal patterns taken as points can be displayed in a new kind of 3-D structure. The latter, by putting into order these points (each representing a family of arrow-hexagons), gives some kind of morphogenetic cartography of the arrow-hexagons of the $\beta 3$-structure. As we will argue, since several arrow-hexagons play the role of "attractors", there are reasons for thinking that such a cartography could be very meaningful in the future for modelling "oppositional dynamics", that is the systematic formal study of the situations where a given complex oppositional structure sees its shape change within time.

Keywords Oppositional geometry · Logical hexagon · Hybrid hexagon · Arrow-hexagon · Oppositional tetrahexahedron · Chromatic signature · Oppositional duality · Oppositional structure · Oppositional role · Oppositional morphodynamics · Oppositional dynamics

Mathematics Subject Classification (2010) Primary 51L99, 05C99 · Secondary 00A30, 03A05, 03B05, 03B10, 03B45

The main results of this paper were discovered by the author in 2011. They were presented at the Third World Congress on the Square of Opposition, Beirut (Lebanon), 26–30 June, 2012.

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic, DOI 10.1007/978-3-319-15368-1_20

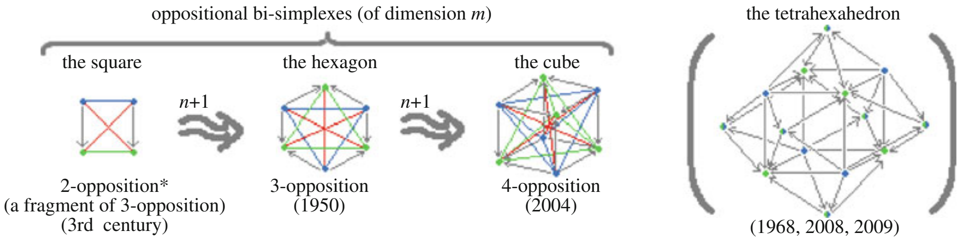


Fig. 1 The four basic "conceptual actors" of oppositional geometry

1 The Starting Point: Oppositional Geometry (in a Nutshell)

Oppositional geometry is, seemingly, a candidate for being a full-fledged young new branch of mathematics, centred on the concept of "opposition". The latter notion is based on Aristotle's discovery of two different forms of opposition: "contrariety" and "contradiction". The logical square expresses the way in which these two forms of opposition are distinct but correlated. Oppositional geometry, in turn, generalises three main oppositional objects: Aristotle-Apuleius' "square of opposition" (or "logical square"), Jacoby–Sesmat–Blanché's "logical hexagon", and Sauriol's "logical tetrahexahedron" (Fig. 1).[1]

Among the main families of oppositional structures, here one must mention the α-, the β- and the γ-structures. The α-structures (or "oppositional bi-simplexes of dimension m") express n-opposition (a bi-simplex is composed of a blue simplex of contrariety intertwined with its dual green simplex of subcontrariety).[2] They are represented by means of their arrows, neglecting visually the three other relations (Fig. 2).

The β-structures express the oppositional closures of the former: to each α-structure corresponds one (and one only) β-structure (its closure). However, each β-structure contains several α-structures, the one of which it is the closure and several instances of smaller ones. Giving all the possible n-oppositions (the maximal one and the smaller ones) yields the *complete* "conceptual panorama" of any oppositional situation (Fig. 3).

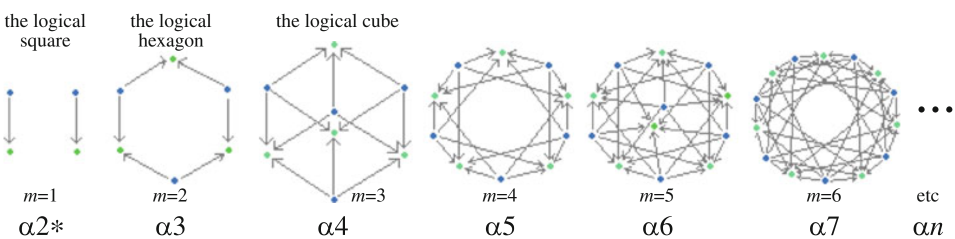


Fig. 2 The series of the oppositional αn-structures (the "bi-simplexes of dimension m"), expressing n-opposition

[1] For an overview, cf. [21].

[2] The first elements of this theory were presented in [17]; a mathematically more compact and powerful axiomatisation has been given in [25].

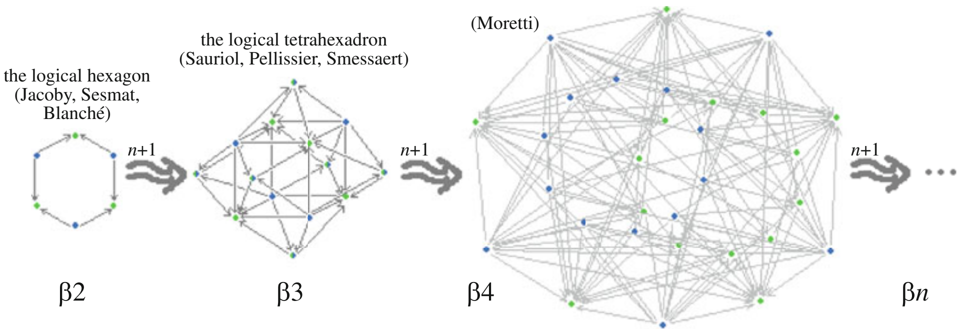


Fig. 3 The series of the oppositional βn-structures (the "oppositional closures")

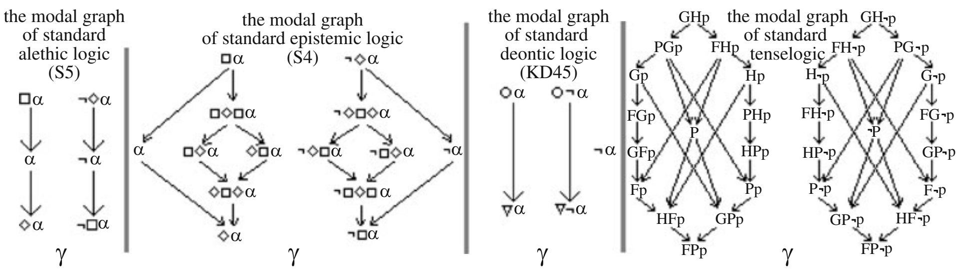


Fig. 4 Some oppositional γ-structures (or "oppositional generators") already present in, but not yet fully understood by, modal logic

The γ-structures are, so to say, the "generators" of the β-structures: (1) they model in a very compact format situations where "modalities" (i.e. opposed entities which can be modal-logical, but not only – they can be more broadly "concepts") are given a negation operator and (2) they can be associated to oppositional closures (that is, the γ-structures map, but not bi-univocally, into the β-structures) (Fig. 4).

Oppositional geometry allows, through some translation rules (i.e. Pellissier's "setting technique"), to switch from the γ- to the β-structures (and, therefore, to give access to entire bunches of α-structures, cf. [19]). As a consequence the β-structures generate an important new instance of equivalence relation, the "equivalence modulo βn", and, therefore, each βn-structure is an equivalence class. As such, oppositional geometry is a way of measuring "oppositional complexity": the βn-structures are increasingly complex oppositionally speaking.

The theory of the oppositional bi-simplexes generalises into several directions. One is the theory of the oppositional *poly*-simplexes.[3] This corresponds to the study of the oppositional geometry of formal systems which rely on some many-valued logic (instead than on classical 2-valued logic). In the poly-simplicial approach oppositional concepts become "diffracted": there are more than one "shade" of contradiction, of subalternation, of contrariety and of subcontrariety. For instance, by adding a pivotal black simplex (between

[3] The oppositional poly-simplexes were first proposed in [18] (starting from Chapter 18).

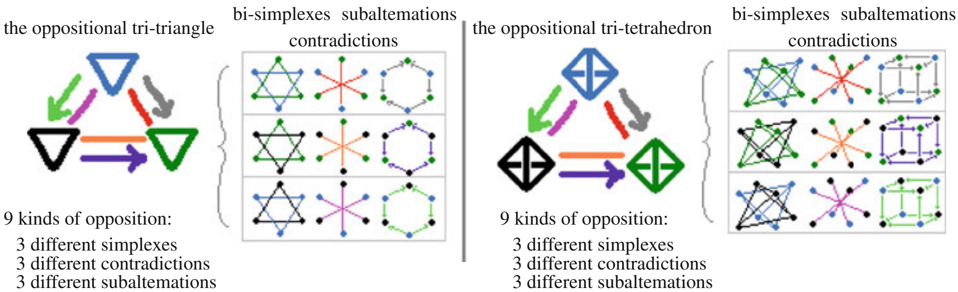


Fig. 5 In an oppositional tri-simplex (whatever its dimension) the classical oppositional relations are diffracted

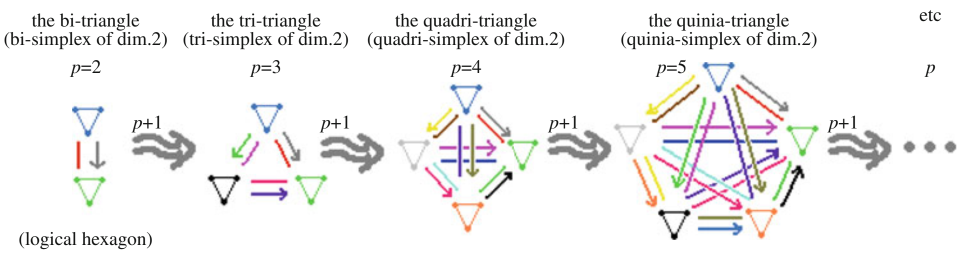


Fig. 6 The series of "oppositional poly-simplexes of dimension 2" (i.e. the "oppositional poly-triangles")

the classical blue and green ones) the tri-simplexes diffract contradiction and subalternation into three (each of the three pairs of simplexes of a tri-simplex, together forming a bi-simplex, has its own kind of contradiction and its own kind of subalternation) (Fig. 5).

As a result of this generalisation, each kind of simplex (line, triangle, tetrahedron, . . .) generates an infinite series. For instance, the simplex of dimension 2 (i.e. the triangle) generates an infinite series of oppositional poly-triangles (the oppositional bi-triangle being the logical hexagon) (Fig. 6).

The poly-simplexes require, in order to articulate the γ-, the β- and the α-structures, a suitable formal treatment involving sheaf theory (cf. [2, 3]). Moreover, they are difficult to represent graphically. However, this can still be done for one of the simplest ones, the oppositional tri-triangle (which is the tri-simplicial version of the logical hexagon) (Fig. 7).

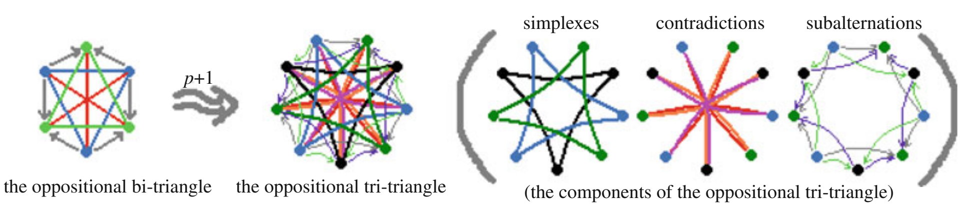


Fig. 7 The tri-simplicial version of the logical hexagon (which is a bi-simplex, the bi-triangle) can still be represented

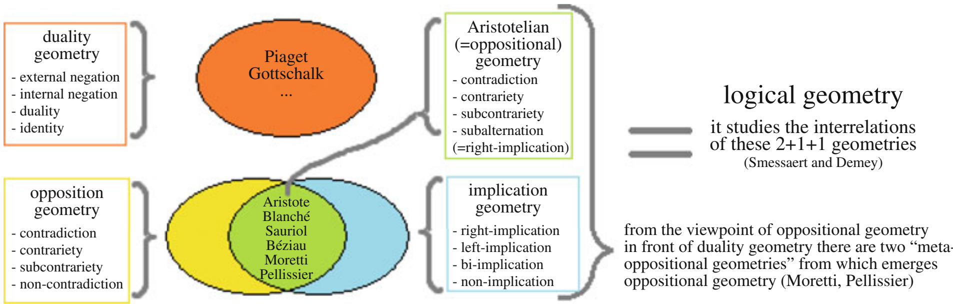


Fig. 8 Oppositional geometry can be seen as emerging from two "meta-oppositional" geometries

Another very important generalisation leads to what H. Smessaert and L. Demey propose calling "logical geometry" (or "logical geometries"). Here some more "geometrical components" are unveiled (the "implication geometry" and the "opposition geometry"), from the interplay of which oppositional geometry (as well as, possibly, some other new domains) can be seen to emerge (Fig. 8).

In the rest of this paper we remain focussed on the bi-simplicial viewpoint of oppositional geometry (which is currently the most powerful approach to "opposition"), leaving for other studies the exploration of these extra dimensions of the theory.

2 Some Recent Remarks on Some Strange Hybrid Hexagons

Oppositional geometry owes much to the logician and philosopher Jean-Yves Béziau. With a seminal paper in 2003 that dealt with a defence of the concept of "paraconsistent logic" (against H. Slater's attack of it in 1995), he re-launched the interest of the logical and philosophical communities for logical hexagons (and thereby, among other things, gave a start to our own discoveries, which he would later direct in our PhD, which was defended in 2009). He also built a remarkable institutional framework for geometrical-oppositional research by creating and keeping alive the now famous series of the "World Congress(es) on the Square of Opposition" from 2007 to now[4]. In more recent times (i.e. in 2012, cf. [6]) Béziau proposed, among several other things, a quite provocative triad of non-standard new oppositional figures, made only of red lines of contradiction and of grey (or black) arrows of implication (Fig. 9).

Now, from the point of view of oppositional geometry if we speak, for instance, of logical hexagons there are *stricto sensu* only two kinds of them: the "strong" and the "weak".[5] Moreover, the difference between these two, contrary to Béziau's proposal, does not touch the quality (expressed by colours) of the relations constituting these logical hexagons: these relations (i.e. the colours of the nine segments inside the hexagon, ex-

[4] On Béziau's reaction to Slater and on the origin of oppositional geometry, cf. [20].
[5] Weak hexagons were (re-)discovered (after Sauriol's first quick mention of them in 1968, cf. [26]) and put into light by Pellissier in 2008, cf. [25].

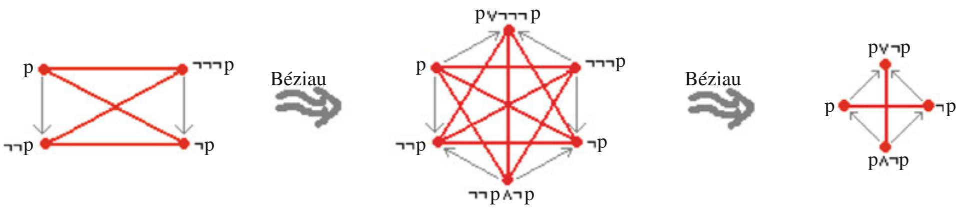


Fig. 9 Béziau's provocative proposal, in 2012, of three oppositional structures made almost only of red contradiction lines

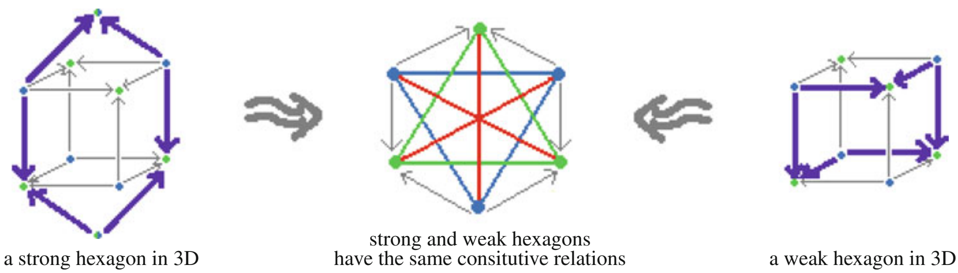


Fig. 10 Comparing "strong" and "weak" logical hexagons: geometry can help the eye

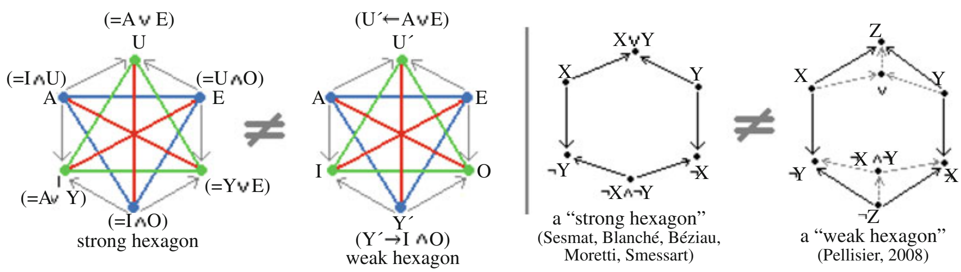


Fig. 11 The difference between "strong" and "week" logical hexagons concerns the interdefinability of the vertices (Pellissier)

cluding the perimeter made of six grey alternated arrows), are the same in the two kinds of hexagons. Rather, the difference between strong and weak logical hexagons touches two other things. Firstly, the geometrical dimensionality (inside the $\beta 3$-structure): strong hexagons are "planar" (i.e. 2-D), whereas weak hexagons are "broken" (i.e. 3-D) (Fig. 10).

Secondly, it touches the interdefinability of the vertices: (a) in a strong hexagon each green vertex is defined as being logically *equivalent* to the inclusive disjunction of its two adjacent blue vertices, whereas, dually, each blue vertex is defined as being logically *equivalent* to the conjunction of its two adjacent green vertices; (b) in a weak hexagon, Pellissier *docet*, things change: the logical equivalencies are weakened and become *implications* (in one or the other direction, alternatively) (Fig. 11).

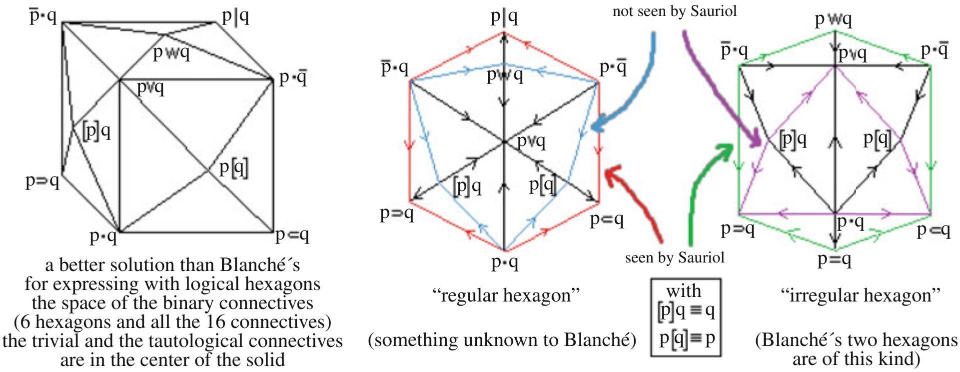


Fig. 12 Sauriol's "logical tetrahexahedron" (1968) and, inside it (hidden), two "arrow-hexagons" (here: in *blue* and *violet*)

So, how can we understand Béziau's provocative "red" hexagon, the inside of which is made of nine red segments of contradiction? Is it a new kind of logical hexagon? Is this possible? Does this contradict oppositional geometry?

As it happens, in the last years this standard view, according to which a logical hexagon (no difference whether strong or weak) is made of two logical simplexes (a blue triangle of contrariety and a green triangle of subcontrariety), intertwined so that central symmetry means contradictory negation (i.e. red diagonals) has been, if not challenged, at least puzzled, not only by Béziau, but also by at least three other groups of scholars having made independently some other strange comparable observations.

First, some strange hexagons (highlighted by us visually in Fig. 12) already appeared 47 years ago in a figure by Pierre Sauriol (1968) – the first known discoverer of the oppositional tetrahexahedron –, who however did not discuss it (in fact Sauriol tried to use his tetrahexahedron mainly in order to diminish Blanché's discovery and polemical use of the logical hexagon, and thereby in order to put a stop to Blanché's use of hexagons for building an anti-analytical philosophy of logic, cf. [11, 12]). Sauriol highlights two kinds of hexagons in the tetrahexahedron he has discovered (hexagons called by him "regular" and "irregular", put by us in red and in green in his figure), but he does not see (or at least he does not mention) that two more kinds of "arrow-hexagons" (i.e. two chains of six alternated arrows, put by us in blue and in violet in his figure) are quite visible in the 3-D figures he presents in his paper (cf. [26, 27]; to this we will come back) (Fig. 12).

Secondly, much more recently, D. Dubois and H. Prade discovered a strange and quite mysterious oppositional hexagon made of arrows not only in its perimeter but also in part of its inside (namely, its three diagonals are arrows, instead of being red segments of contradiction). They did this while working, with tools akin to those of oppositional geometry, on the relations holding between "possibility theory" and "formal concept analysis" (Fig. 13).[6]

[6] Cf. [14]; some critical remarks on the "suboptimal" character of their approach can be found in [22].

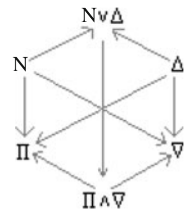


Fig. 13 A strange version of the logical hexagon (with some relations modified) proposed by Dubois and Prade (2012)

Thirdly, we have made other similar remarks while studying the possible combinatorics of oppositional structures, looking (as we are) for some kind of "oppositional product" and "sum" (or "coproduct"). This first appeared in our attempt to combine, by some kind of "oppositional distribution" (centred on the "$\wedge$" connective), an oppositional hexagon with an oppositional tetrahexahedron;[7] this operation generates strange new hexagons that keep a regular perimeter of alternated arrows but have blue (instead of red) diagonals and an orange (instead of green) lower triangle (orange means "mutual independence", cf. Sect. 3.6) (Fig. 14).

Let us call, from now on, "hybrid hexagons" those hexagons that see any change of colour whatsoever with respect to the standard colours (i.e. relations) of the logical hexagons, and let us call "arrow-hexagons" those hexagons (standard or hybrid) that keep intact the perimeter of alternated arrows (some hybrid hexagons are arrow-hexagons, but not all).

The same result, i.e. an instance of "arrow-hexagon" with blue diagonals and an orange lower triangle (instead of the standard green triangle), appeared in a study of the oppositional distribution (centred on the "$\wedge$" connective) of two logical hexagons, suggesting that this is a stable, meaningful behaviour (some kind of "signature" of the general operation of oppositional $\wedge$-distribution) (Fig. 15).

The study of the same operation (oppositional distribution), still over two hexagons, but this time seen from the viewpoint of another junctor (i.e. not the "$\wedge$", but the "$\vee$" connective), yields a similar (in fact: "oppositionally dual")[8] result: an "arrow-hexagon"

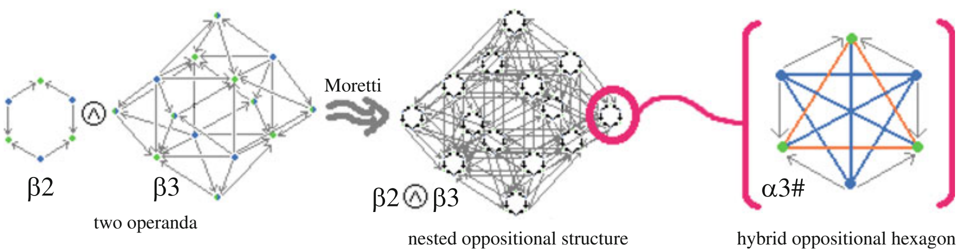


Fig. 14 Distributing conjunctively a logical hexagon over a tetrahexahedron produces a bundle of strange "*blueish*" hybrid hexagons

[7] By a "distribution" we mean the application of the first element (here a logical hexagon), taken as a whole, on each vertex of the second (here an oppositional tetrahexahedron).

[8] Two oppositional structures are said to be "oppositionally dual" if either of the two is obtained from the other by switching the direction of each arrow and exchanging every blue component (i.e. vertices or segments) with a green one and *vice versa* (blue–green vertices and orange segments remain unchanged).

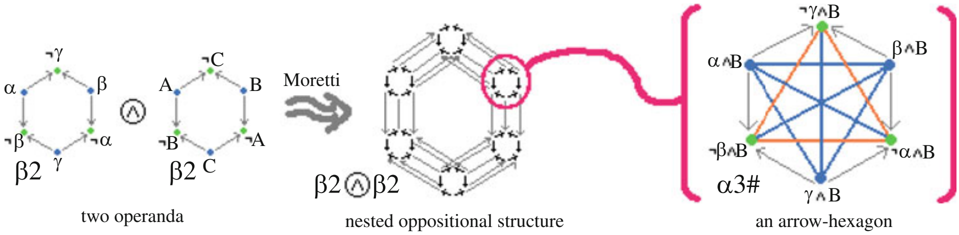


Fig. 15 Distributing conjunctively a logical hexagon over another one produces a bundle of strange "*blueish*" arrow-hexagons

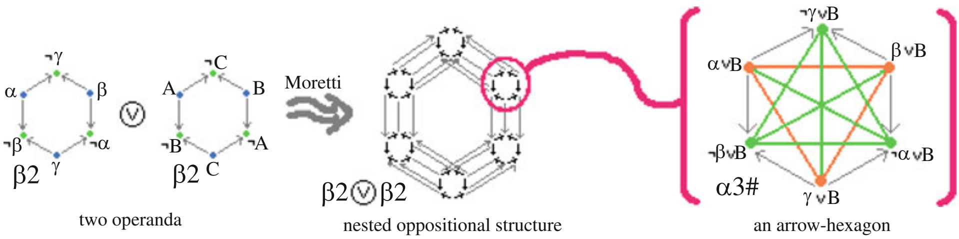


Fig. 16 Distributing disjunctively a logical hexagon over another one produces a bundle of strange "*greenish*" arrow-hexagons

with green (instead of red) diagonals and an orange upper triangle (instead of the standard blue one) (Fig. 16).

Moreover, while studying the oppositional geometry of partial order, we remark, again, that strange logical hexagons emerge that, at least in some cases, might be meaningful. This is precisely the case of the famous "hexagon of orders" (considered by many as the most elegant known instance of the logical hexagon) when it is "plunged" into partial order (i.e. when one considers the possibility of having, for a and b, not only the three possibilities $a > b$, $a < b$, or $a = b$, but also a fourth one: "a and b cannot be mutually ordered", i.e. "$a \parallel b$"). Then its inner relations (i.e. its oppositional inner colours – not the perimeter made of alternated arrows) have to change (because of the presence, now, of the order relation "$\parallel$") and as a result, inside the new oppositional structure of orders (which is no longer a $\beta 2$, but a $\beta 3$-structure, i.e. not an oppositional hexagon, but an oppositional tetrahexahedron) it becomes a new kind of strange hexagon, another instance of "arrow-hexagon", with two blue (instead of red) diagonals and a partly orange (instead of totally green) lower triangle (Fig. 17).

This new arrow-hexagon of partial order (which is a fragment of the "tetrahexahedron of partial order") is important for at least two reasons. Firstly, it clearly embodies an instance of "oppositional metamorphosis": the passage from the hexagon of orders (in the mathematical space of *total* order) to its hybrid counterpart (in the mathematical space of *partial* order) seems to be the first promising step towards a formalisation of "oppositionally dynamic phenomena". Secondly, this hybrid hexagon seems to show (as, in fact, do also Sauriol's aforementioned two strange hexagons) that such strange "hybrid hexagons"

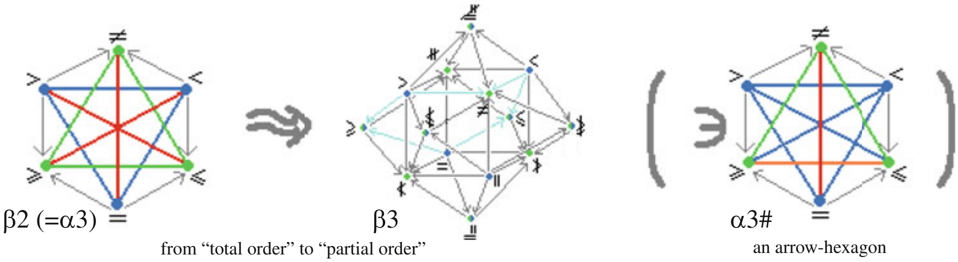


Fig. 17 The oppositional dynamics of the order relations: a $\beta 2$ becomes $\beta 3$ leaving an $\alpha 3\#$ “shadow” of itself inside $\beta 3$

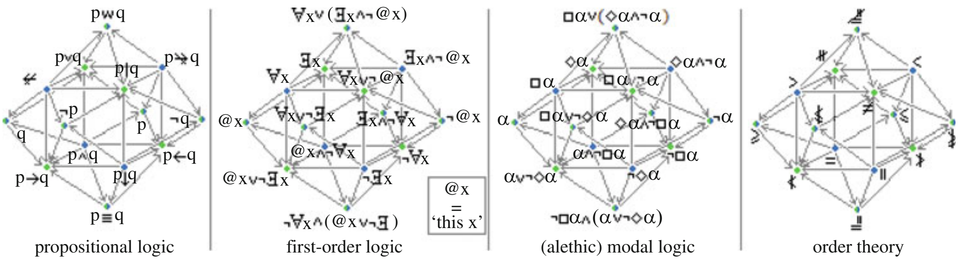


Fig. 18 The oppositional tetrahexahedron “pops up” in many different fields of fundamental research

(and more precisely these arrow-hexagons) are, in fact, very natural, since they are implicitly present (although generally unseen) in each $\beta 3$-structure, the latter being, notoriously, a very powerful “attractor” of oppositional geometry (Fig. 18).

Hence the question, possibly clarifying the previous remarks (made by Béziau and other scholars) mentioned in this section: how many such arrow-hexagons are there, in fact, inside the $\beta 3$-structure? Do they have interesting invariant properties? The rest of this paper will try to give some element of answer to these questions.

3 How Many “Arrow-Hexagons” Are There in the $\beta 3$-Structure?

The $\beta 3$-structure seems to be the first place to look in a systematic way for hybrid hexagons in general and for arrow-hexagons in particular, since it is a very well-known structure (among those of oppositional geometry) and has the advantage of being both very rich and still visually perceivable (differently from the general βn-structures, with $n \geq 4$, which are hyper-dimensional and, therefore, visually counterintuitive). However, a simple combinatorial calculation shows that in the tetrahexahedron there are exactly C^6_{14}, that is 30 030 possible “hexagons”, 30 020 of which are hybrid! In the rest of this study we shall study only arrow-hexagons. In order to have, in the midst of this thick combinatorial jungle, a way of naming the transformations that we are be interested in, we propose the following terminology for naming the tetrahexahedron’s vertices (“b” is

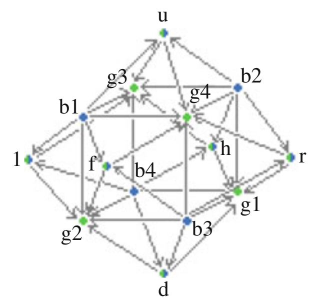


Fig. 19 The convention adopted throughout this study for naming without ambiguity any of the tetrahexahedron's 14 vertices

for blue, "g" is for green, "u", "d", "l", "r", "f" and "h" are, respectively, for up, down, left, right, front, behind) (Fig. 19).

So, we will start by considering the two standard instances of oppositional hexagon, the "strong" and the "weak", and then we will try to widen our scope with a suitable "transformation game".

3.1 Strong Oppositional (or Logical) Hexagons

Inside the β3-structure our starting point is given by the two kinds of oppositional hexagons that we know: the "strong" and the "weak" (according to Pellissier's terminology, cf. Pellissier (2008)). In this section let us focus on strong hexagons. As is known, there are six strong oppositional hexagons inside the β3-structure, two for each of the three "Cartesian axes" of the oppositional tetrahexahedron (Fig. 20).[9]

From now on, let us play a new kind of "transformation game" with these six strong hexagons.

To start with, let us ask ourselves what all the possible "1-move transformations" of a strong hexagon are, that is the hexagonal structures obtained by changing place of one of its six vertices in a way such that the resulting structure still has a perimeter made of six alternated arrows. We know that the six strong hexagons of any β3-structure have three kinds of vertices: either blue or green (if this vertex belongs to the "heart" of the β3-structure, i.e. the oppositional cube) or blue–green (this vertex is one of the six "spikes" or "Egyptian pyramids" of the β3-structure). As it turns out, the only points (i.e. vertices) of a strong hexagon leading to a possible (i.e. viable) move (i.e. one leading to another arrow-hexagon) are the blue–green points; moving a blue or a green point, conversely, generates hexagonal structures the perimeter of which is not uniquely made of alternated arrows, thus hybrid oppositional hexagons that are not arrow-hexagons, and we are not interested in such "broken arrow perimeter" hexagons in our present transformation game (Fig. 21).

[9] The first was discovered by Jacoby (1950), Sesmat (1951) and Blanché (1953) independently; the second and the third by Béziau (2003); the fourth by Moretti (2004) and Smessaert (2009) independently; the fifth and the sixth by Smessaert (2009). As it happens Sauriol had discovered all six (in 1968) but his paper was not noticed until 2012, cf. [21].

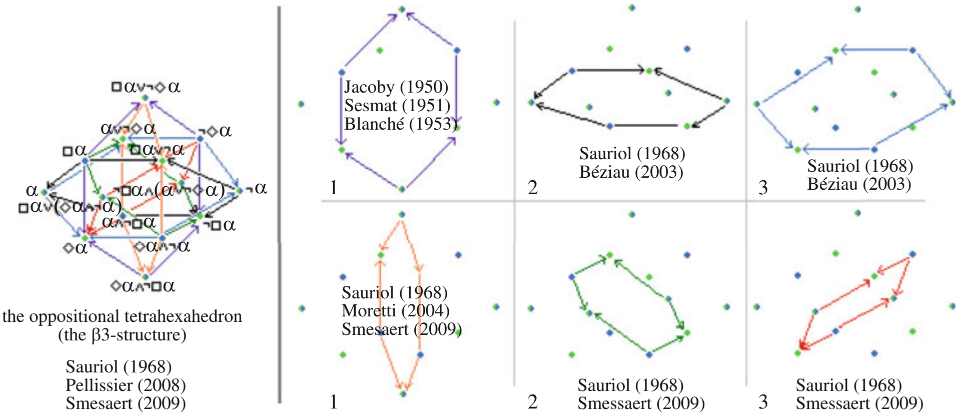


Fig. 20 The six strong hexagons contained in the oppositional tetrahexahedron (and their discoverers)

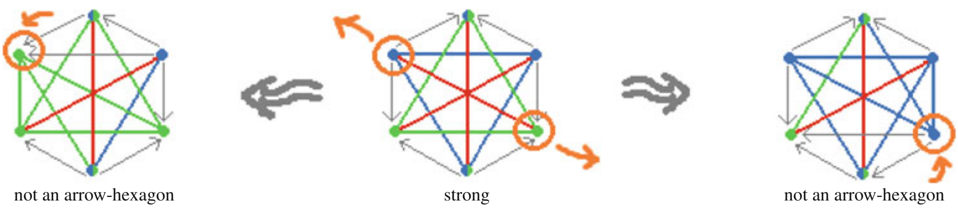


Fig. 21 No move of either a blue or a green vertex of a strong hexagon leads to an arrow-hexagon

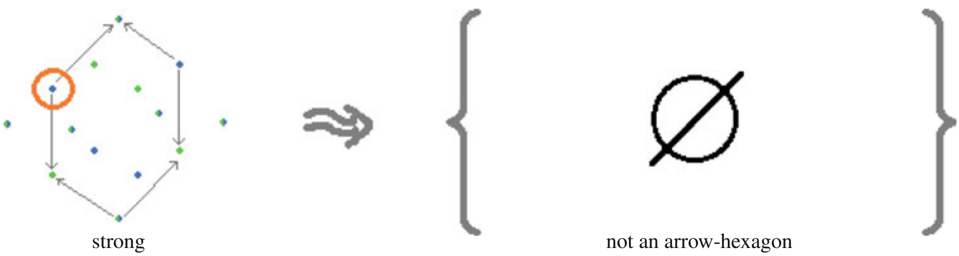


Fig. 22 Some vertices of strong hexagons (i.e. all those belonging to the cube) lead to no possible 1-move transformation

So, from now on we will express as follows, by a drawing, the fact that there are no solutions for a 1-move of a vertex (of an arrow-hexagon) such as the one highlighted by an orange circle (Fig. 22).

Conversely, changing the place of one of the two blue–green vertices of a strong hexagon (of an oppositional tetrahexahedron) happens to lead to two possible satisfactory outcomes, i.e. to two new instances of arrow-hexagons (we call them, momentarily, “X1”) (Fig. 23).

Since each of the two blue–green vertices can lead to two transformations (here we depict only the two “upper” ones), the outcome of which is isomorphic (a “X1” arrow-

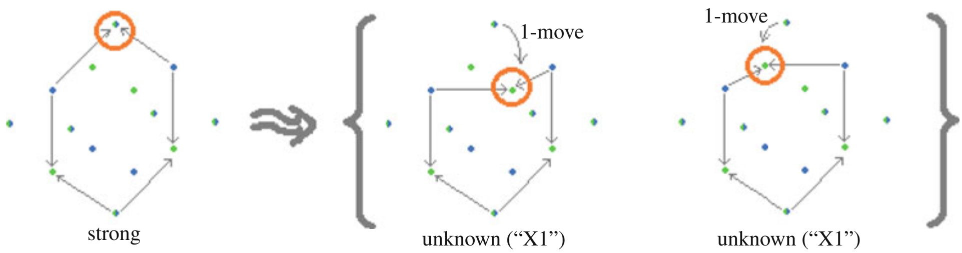


Fig. 23 All possible “1-move transformations” of any blue–green vertex of a strong hexagon

hexagon), we see that for each strong hexagon of a $\beta3$-structure there are globally four possible transformations into a new arrow-hexagon. Moving the lower, instead of the higher blue–green vertex of the strong hexagon gives an arrow-hexagon of type “X1 bis”; blue and green vertices are switched, with respect to those of the “X1”, and the order of each arrow is reversed (we will examine all this in Sect. 3.3, which is devoted to the special kind of arrow-hexagon “X1” or “X1 bis” thus obtained).

3.2 *Weak Oppositional (or Logical) Hexagons*

As we already recalled, Pellissier has discovered the presence of four instances of what he called “*weak* logical hexagons” inside the $\beta3$-structure. More precisely, they are contained inside the “heart” of the $\beta3$-structure, that is, inside the $\alpha4$-structure (i.e. the “oppositional cube”) (Fig. 24).

As Pellissier established (2008), these four hexagons have properties that are different with respect to the classical (i.e. the strong) ones. It can also be remarked (cf. [16]) that in certain contexts of modelling application the weak hexagons, with respect to the strong ones, are particularly meaningful and interesting in so far as they can play the role of “extreme choice”; the oppositional hexagon such that it “takes the best” (or “the worst!”) of two intertwined options cannot be a strong hexagon and has to be a weak hexagon (Fig. 25).

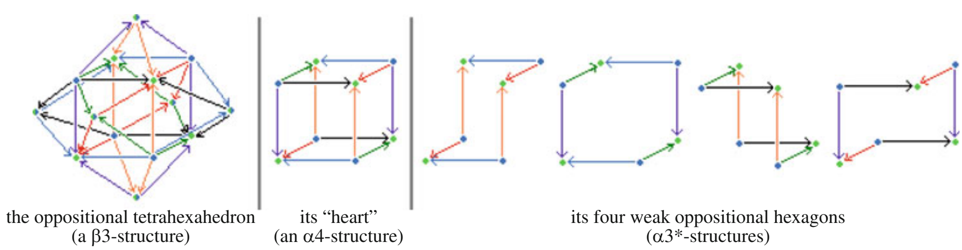


Fig. 24 The oppositional tetrahexahedron ($\beta3$) contains an “oppositional cube” ($\alpha4$), which contains four “weak hexagons” ($\alpha3*$)

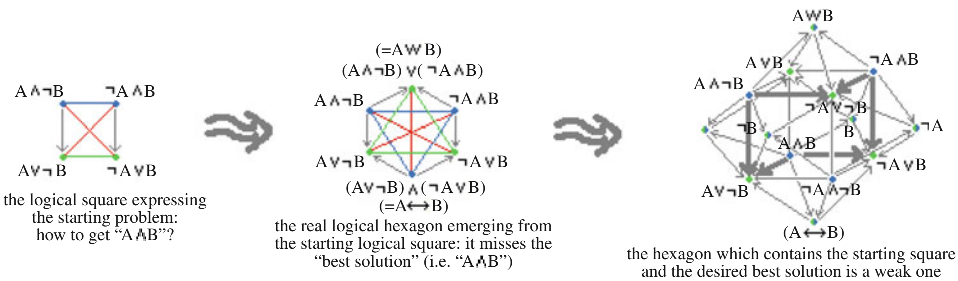


Fig. 25 An application of weak hexagons to modelling political theory; avoiding conflict through "oppositional crossbreeding"

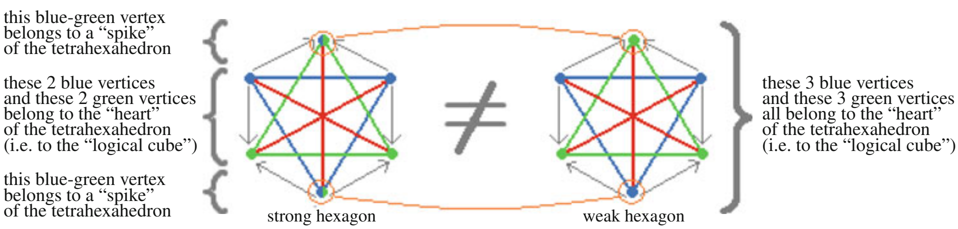


Fig. 26 The chromatic signature of strong and weak hexagons differ with respect to the colour of some of their vertices

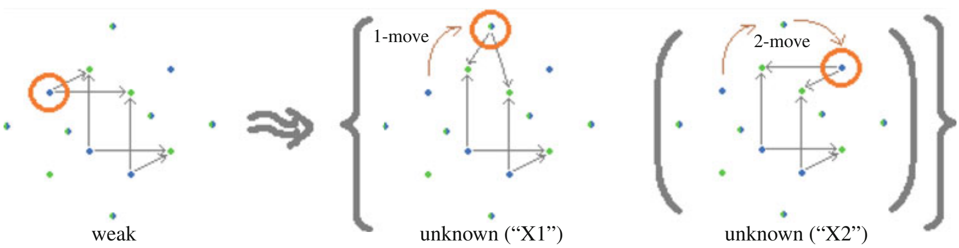


Fig. 27 Possible 1-move (and 2-moves) transformation(s) of any of the six vertices of a weak hexagon

It is worth stressing this point (which will become more meaningful later) by saying that weak hexagons have an "oppositional role".

Let us now introduce the idea that, truly speaking, there is a slightly different "chromatic signature" of the two kinds of oppositional hexagons (strong vs. weak): the two mixed blue–green points of the former (which represent those vertices that do not belong to the cube, the tetrahexahedron's "spikes") get lost in the latter (Fig. 26).

From the viewpoint of our transformation game the next natural combinatorial question is then: what are all the possible 1-move transformations of a *weak* hexagon (of a tetrahexahedron)? The answer is synthetically provided by the Fig. 27.

We can see that for any of the six vertices of this hexagonal structure (which is an instance of an arrow-hexagon) this time there are two different kinds of outcome: the

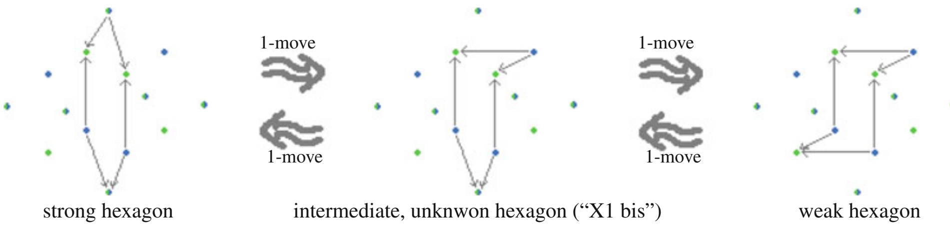


Fig. 28 Is there anything “morphodynamical” between strong and weak oppositional hexagons?

“X1” arrow-hexagon (already seen in Sect. 3.1) and the “X2” arrow-hexagon (which is new). The first transformation, the “1-move” transformation, leads to the same shape (the “X1” arrow-hexagon) as the one obtained starting from a strong hexagon in the previous section (this shape will be described in Sect. 3.3 below). The second transformation, the “2-moves” transformation shown in parentheses in the previous figure, corresponds to a shape (the “X2” arrow-hexagon) which will be studied in Sect. 3.4. However, for reasons to appear later, we will consider that the latter is obtained through *two* 1-moves (and not just *one*), so we will not consider it as a transformation directly reachable through a 1-move. The intuitive idea is that in our game one moves vertices by making them “slide” over the stellar surface of the tetrahexahedron, from vertex position to vertex position, and that from this viewpoint some moves are “shorter” than others (they take two moves instead of one).

Let us also remark that now each of the two classical kinds of hexagon (strong and weak) can be seen as a *transformed* counterpart of the other. In other words, let us remark that in order to transform a strong hexagon into a weak one (or *vice versa*) one only needs to displace two vertices of the starting one (be it strong or weak) in a suitable way (Fig. 28).

The rest of this paper will consist as much as possible in a systematic study of the progressive structure-preserving deformations of oppositional hexagons. More precisely, we are concerned with the property of being an “arrow-hexagon” (i.e. an hexagon the perimeter of which is made of regularly alternated arrows). It is this issue that we are going to explore in the next sections.

So at this stage, qualitatively speaking, we have four different kinds of hexagons (in the next two sections we shall study the two new ones that we just discovered) (Fig. 29).

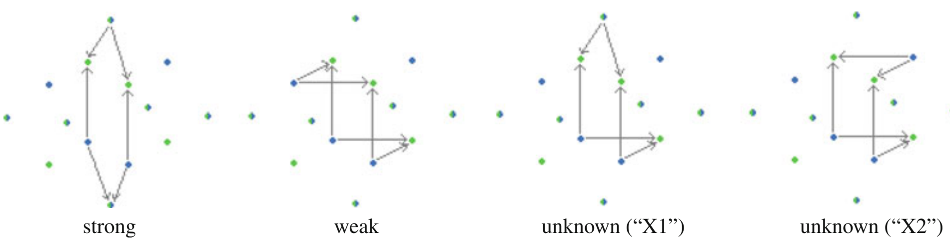


Fig. 29 Synoptic view of the four kinds of hexagons perceived so far inside the oppositional tetrahexahedron (or β3-structure)

The natural question emerging from this exploration will be that of reflecting on how to order them mutually.

3.3 The Unknown Cases "X1" (and "X1 bis") Are "Parrot" Hexagons

Let us focus on the new kind of arrow-hexagons we found and named "X1" (and, dually, "X1 bis"). As we saw, there are two ways of generating it: either by deforming a strong hexagon or by deforming a weak hexagon. Let us start with weak hexagons, naming them 1, 2, 3 and 4, and labelling them with the labels of our reference tetrahexahedron (cf. Sect. 3 above, Fig. 19) (Fig. 30).

One way of deforming each of the four weak hexagons consists in applying a 1-move transformation to one of its three blue vertices to consider a blue–green vertex instead. This generates the structure previously called "X1". For reasons soon to appear, let us call it a "green-beaked parrot" (i.e. a "parrot" whose "beak" is a green vertex). There are 12 of them, distributed (for easiness) in 3 quartets. The first such quartet of green-beaked parrots is obtained by a 1-move (in the weak hexagons 1, 2, 3 and 4, respectively) of the vertices b1, b4, b3 and b2, respectively (Fig. 31).

The green-beaked parrots 1 and 3 are said to be "twins" (because they differ only in one vertex, which makes them symmetric), as are parrots 2 and 4.

The second quartet of green-beaked parrots, represented with respect to the 14 vertices of the β3-structure is the one resulting from a 1-move (of the weak hexagons 1, 2, 3 and 4, respectively) of the vertices b4, b1, b2 and b3, respectively (Fig. 32).

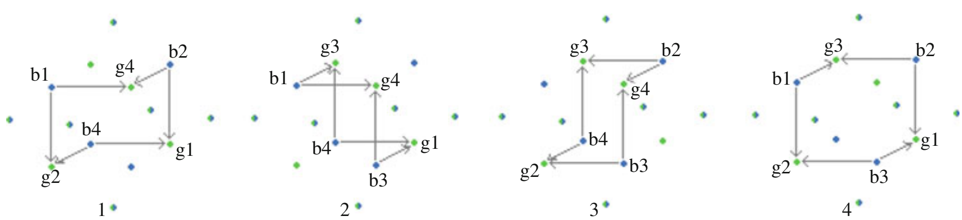


Fig. 30 Starting from the four weak hexagons and by a 1-move of some of their vertices we can obtain "parrot" arrow-hexagons

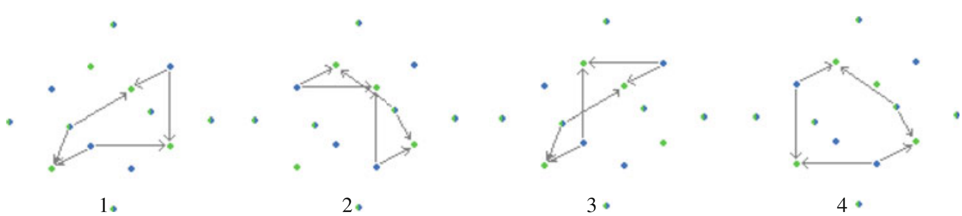


Fig. 31 The first quartet of "green-beaked parrots" (1–4)

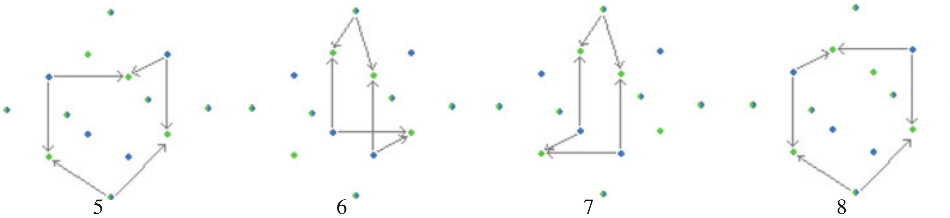


Fig. 32 The second quartet of the "green-beaked parrots" (5–8). The green-beak parrots 5 and 8 are twins, as well as are the 6 and 7

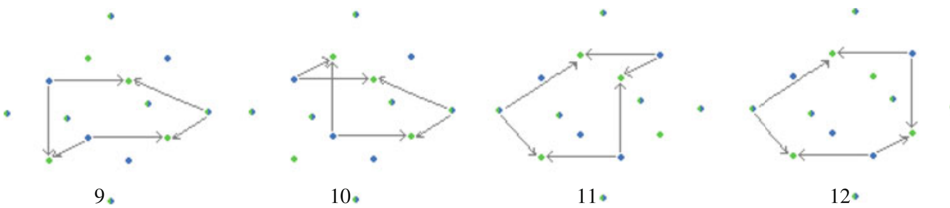


Fig. 33 The third quartet of the "green-beaked parrots" (9–12). The green-beaked parrots 9 and 10 are twins, as are parrots 11 and 12

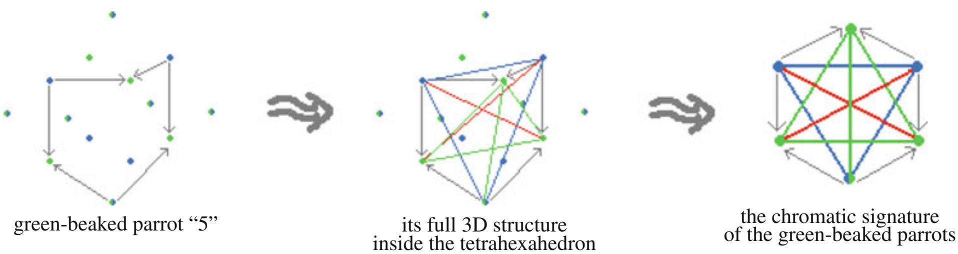


Fig. 34 The chromatic signature of the "green-beaked parrot" arrow-hexagons

Similarly, the third and last quartet of green-beaked parrots is the one resulting from a 1-move (in the weak hexagons 1, 2, 3 and 4, respectively) of the vertices b2, b3, b4 and b1, respectively (Fig. 33).

Now, if one considers not only the arrows (subalternation) but also the other three Aristotelian oppositional relations (contradiction, contrariety and subcontrariety) these arrow-hexagons have an inner structure that is dramatically different from that of strong and weak hexagons; one diagonal becomes blue instead of red (which means that contradiction along this diagonal has become a contrariety). This can be highlighted by the chromatic signature of the green-beaked parrots (here we highlight this feature on green-beak parrot "5" in our previous list) (Fig. 34).

The second possible 1-move transformation of a weak hexagon consists in moving not a blue, but any of the three *green* vertices of any of the weak hexagons. For reasons similar to the previous ones, this gives 12 arrow-hexagons, each isomorphic to the one previously

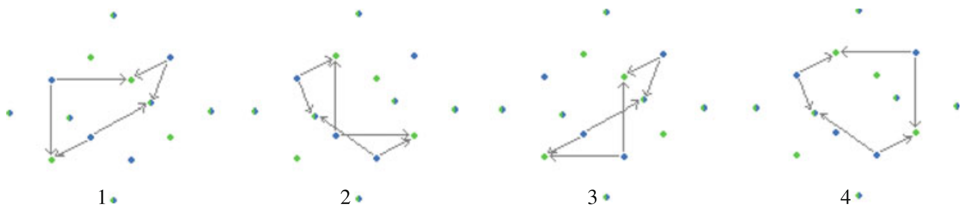


Fig. 35 The first quartet of “blue-beaked parrots” (1–4). As one can see, blue-beaked parrots 1 and 3 are twins, as are 2 and 4

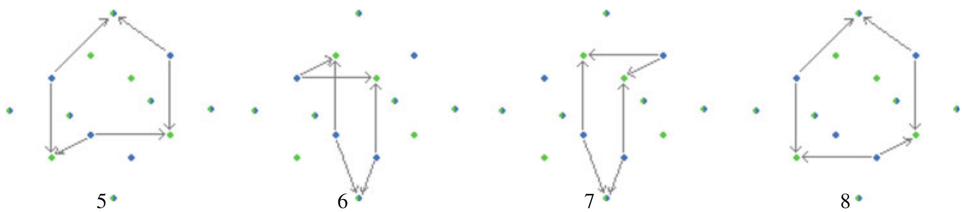


Fig. 36 The second quartet of “blue-beaked parrots” (5–8). As one can see, blue-beaked parrots 5 and 8 are twins, as parrots 6 and 7

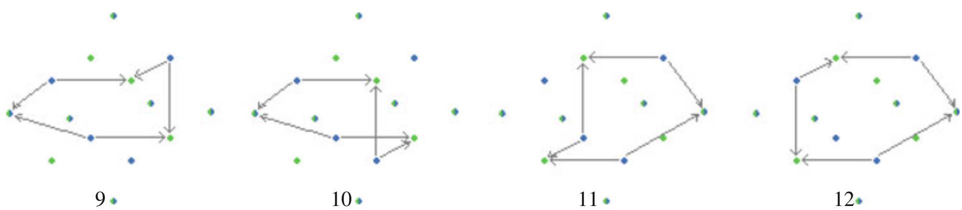


Fig. 37 The third quartet of “blue-beaked parrots” (9–12)

called “X1 bis”, which from now we shall call “blue-beaked parrots”. As previously, it turns out that there are 12 of these, which we distribute (for easiness) into three quartets.

Relying as previously on the list of the weak hexagons 1–4 (Fig. 30), the first quartet is the one resulting from a 1-move of the vertices g1, g4, g3 and g2, respectively (Fig. 35).

The second quartet of blue-beaked parrots is the one obtained by a 1-move of the vertices g4, g1, g2 and g3, respectively (Fig. 36).

The third and last quartet of blue-beaked parrots is that obtained by a 1-move of the vertices g2, g3, g4 and g1, respectively (Fig. 37).

One should remark that blue-beaked parrots 9 and 10 are twins, as are parrots 11 and 12.

The chromatic signature of the blue-beaked parrots is the following (here we highlight this feature on blue-beaked parrot “8” in our previous list) (Fig. 38).

Remark It is a meaningful feature that the green-beaked and the blue-beaked parrots are “oppositionally dual” (one can be obtained from the other by inversing the direction of

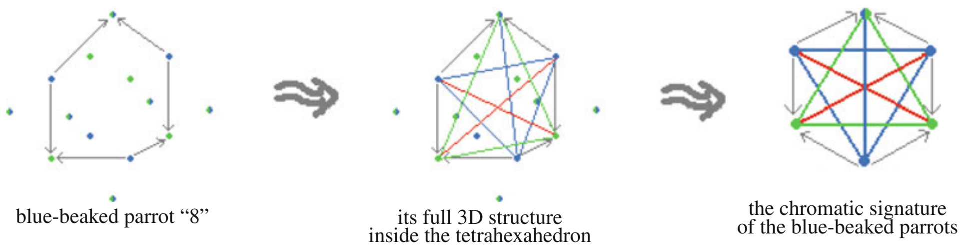


Fig. 38 The chromatic signature of the "blue-beaked parrot" arrow-hexagons

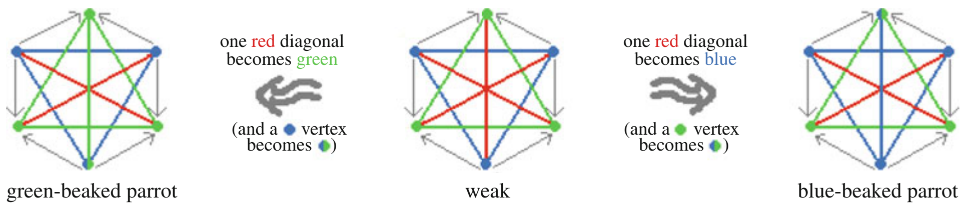
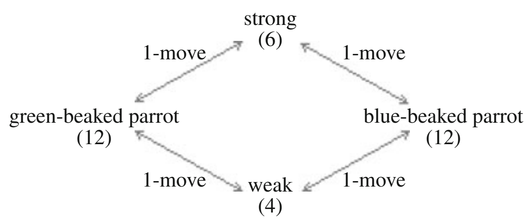


Fig. 39 "Weak" hexagons lie between "green-beaked parrot" hexagons and "blue-beaked parrot" hexagons

Fig. 40 A morphogenetic chart of the arrow-hexagons of the oppositional tetrahexahedron know so far



each arrow and by exchanging blue and green). As previously, this is highlighted by their chromatic signature (Fig. 39).

We are now able to put into some order at least three of the four kinds of arrow-hexagons that we have encountered so far. In fact, at a meta-level, we have the following "transformation path", between green or blue-beaked parrots, between strong and weak hexagons: a strong hexagon can become weak by undergoing a temporary 1-move transformation into a green- or a blue-beaked parrot (and *vice versa*); and a green-beaked parrot can become blue by undergoing a temporary 1-move transformation into a strong or a weak hexagon (and *vice versa*) (Fig. 40).

Note that in some sense we can order these arrow hexagons from the viewpoint of what we can call their "oppositional degeneration" (Fig. 41).

This gives a second kind of morphogenetic chart, namely one where the transitions between different kinds of arrow-hexagons follow the progression of the degeneration (Fig. 42).

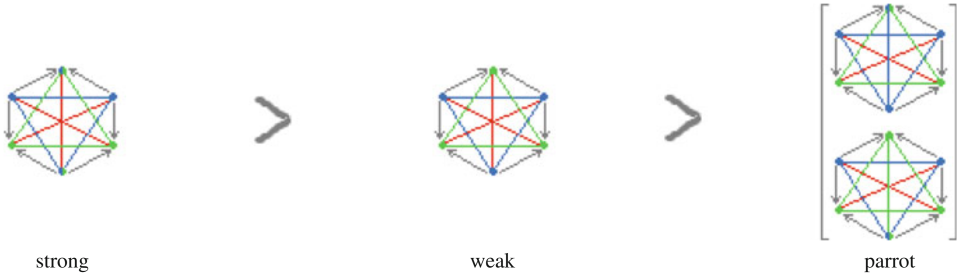
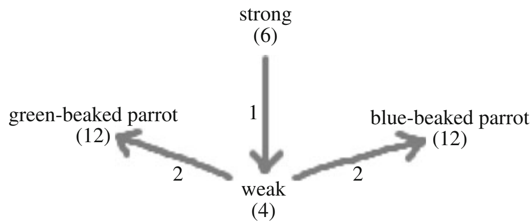


Fig. 41 Arrows-hexagons, so far, can be put in linear order of growing “oppositional degeneration”

Fig. 42 A morphogenetic degeneration chart: a possible way of modelling “oppositional entropy”



These are possibly the first elements of a wider study of how oppositional (hexagonal) structures can transform themselves into others through small transformations. So let us see what comes next.

What are all the possible 1-move transformations of a parrot? In order to answer that question, several distinct cases must be considered, for parrots lose several symmetries with respect to the strong and weak hexagons (although they still possess some). We restrict ourselves to the transformations of a green-beaked parrot (and take the one called “5” in our list). The possible transformations occasioned by a move of the parrot’s “beak” are the following two: one is a 1-move transformation, the other is a 2-moves transformation (Fig. 43).

This brings no novelty: the 1-move generates a strong hexagon while the 2-moves transformation generates the green-beaked parrot twin of the starting one.

The transformations occasioned by a move of the parrot’s “tail” are the following two, both of which are a 1-move (Fig. 44).

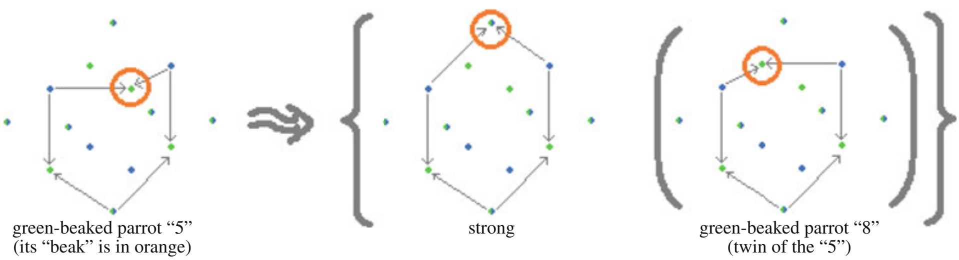


Fig. 43 The possible 1-move (and 2-moves) transformation(s) of a parrot’s “beak”

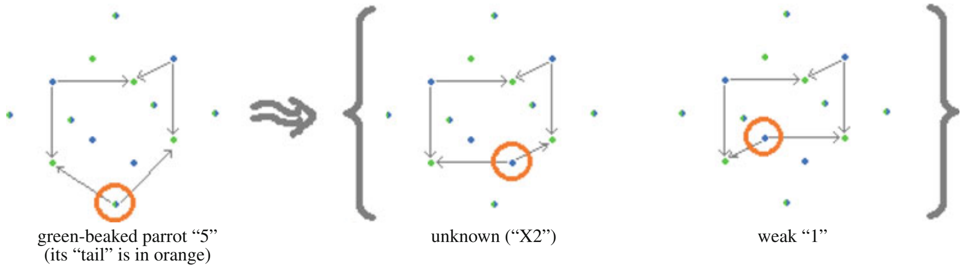


Fig. 44 The possible 1-move transformations of a parrot’s “tail”

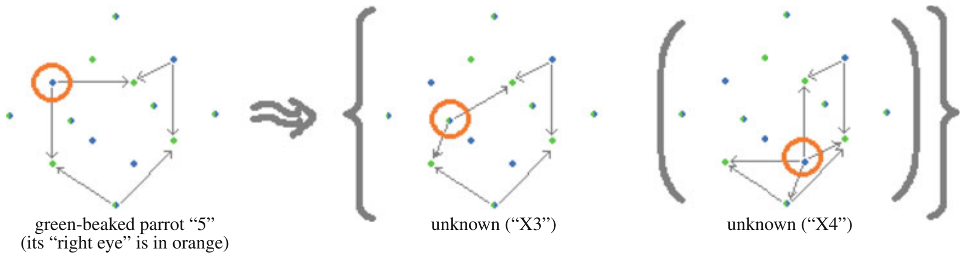


Fig. 45 The possible 1-move (and 2-moves) transformation(s) of a parrot’s “eye”

One can see that one of these two 1-move transformations leads to the new kind of arrow-hexagon called previously “X2” by us (we shall study it in the next section), while the other leads to a weak hexagon.

Next, the transformations occasioned by a move of one of the parrot’s “eyes” are, in turn, the following two, the first of which is a 1-move, while the second is a 2-moves transformation (Fig. 45).

One can see that the 1-move transformation leads to a new kind of arrow-hexagon, which we call momentarily “X3”, while the 2-moves transformation leads to a further new arrow-hexagon, which we call momentarily “X4” (both will be studied later in this chapter).

Finally, no move of any of the parrot’s two “legs” leads to a viable arrow-hexagon (Fig. 46).

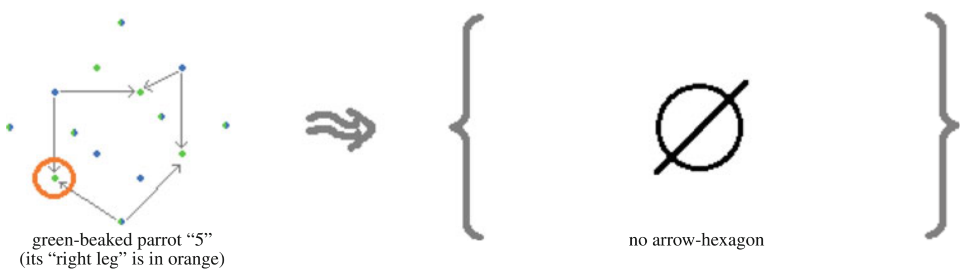


Fig. 46 A parrot’s “leg” admits neither a 1-move nor a 2-moves transformation

As a global result, the green-beaked parrot "5" studied here leads (as well as any other parrot), by a 1-move transformation of any of its vertices, to five possible arrow-hexagons.

Being done with parrots, let us turn to one new family of arrow-hexagons.

3.4 The Unknown Case "X2" Consists of "Book" Arrow-Hexagons

As it happens, with respect to the morphogenetic chart of the previous section, the blue and green parrots can be transformed into each other by a 1-move transformation in another way. This is done by passing through a new kind of arrow-hexagon, the one that we previously called "X2", which we now call, for a visually understandable reason, "book". There are 12 such book-shaped arrow-hexagons, which we distribute into three quartets. The book-hexagons can also be obtained, outside our game, starting from the oppositional cube (i.e. the $\alpha 4$-structure, the "heart" of the oppositional tetrahexahedron). By deleting, one by one, each of the four *front–rear* arrows of the cube (as well as those touching these), one obtains the first quartet of book-shaped hexagons (Fig. 47).

These four are also generated, respectively, by a 1-move of the tail of the green-beaked parrots 1, 2, 3 and 4 or (equivalently) by a 1-move of the tail of the blue-beaked parrots 3, 4, 1 and 2. As one can see, books 1 and 3 are "twins" (in a sense comparable to what justified us speaking about "twin parrots"), and books 2 and 4 are twins as well.

Similarly, by deleting, one by one, each of the four *vertical* arrows of the oppositional cube, one obtains the second quartet of book-shaped hexagons (labelled 5 to 8) (Fig. 48).

Again, these four are also generated, respectively, by a 1-move of the tail of the green-beaked parrots 6, 8, 7 and 5 or (equivalently) by a 1-move of the tail of the blue-beaked parrots 7, 5, 6 and 8. As one can see, books 5 and 7 are twins, as are books 6 and 8.

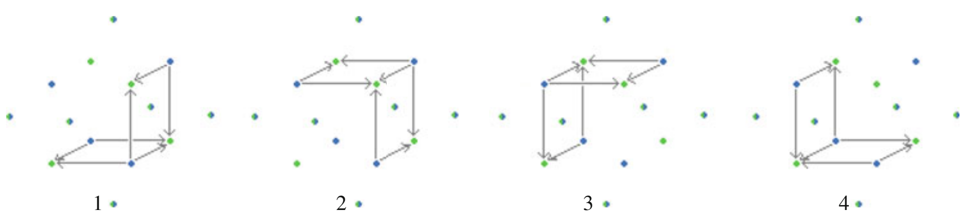


Fig. 47 The first quartet of "books" (1–4)

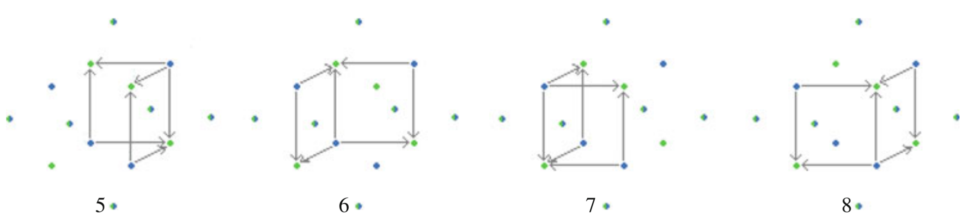


Fig. 48 The second quartet of the "books" (5–8)

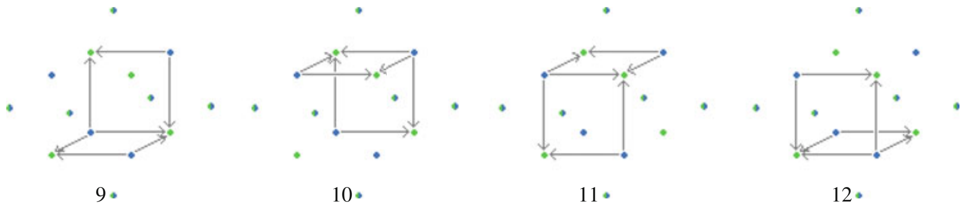


Fig. 49 The third quartet of the “books” (9–12)

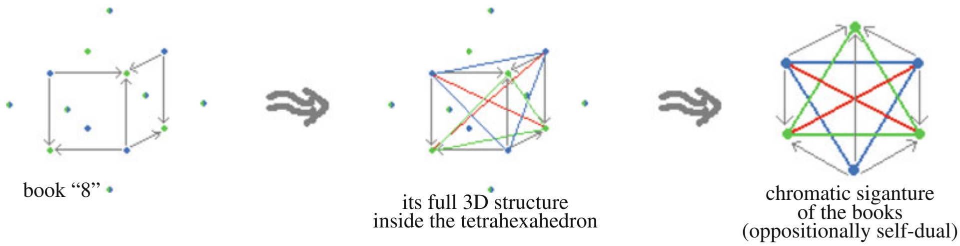


Fig. 50 The chromatic signature of the “book” arrow-hexagons

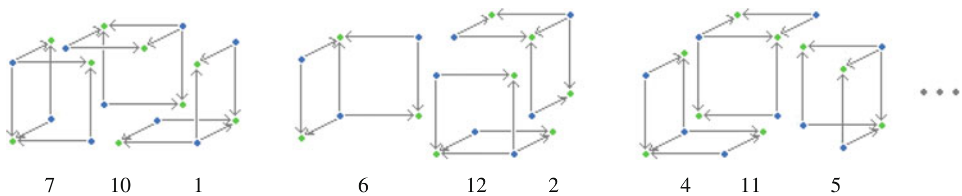


Fig. 51 Triples of books (one for each quartet) can pave the tetrahexahedron’s heart: the oppositional cube

In the same way, by deleting the four *horizontal* arrows of the cube, one by one, we obtain the third and last quartet of books (Fig. 49).

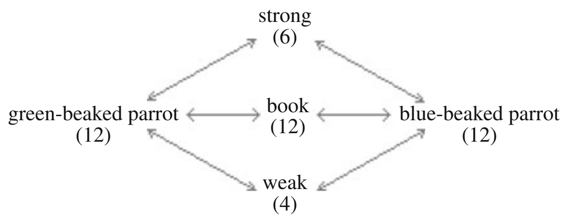
As previously, these four books can also be generated by a 1-move of the tail of the green-beaked parrots 12, 10, 11 and 9 or (equivalently) by a 1-move of the tail of the blue-beaked parrots 11, 9, 12 and 10. Clearly, books 9 and 11 are twins, as are books 10 and 12.

As it happens, each of these 12 books has the same internal structure. This can be put into evidence by drawing the chromatic signature of any of the book-shaped hexagons (we do it for book “8” in our list) and this signature is “oppositionally self-dual” (Fig. 50).

Note that triples of books (such that each triple contains one book of each quartet) can pave the tetrahexahedron’s heart (Fig. 51).

We can now express the morphodynamical relations holding between these five kinds of hexagons (strong, weak, books, blue-beaked parrots and green-beaked parrots) through the following morphogenetic chart (which refines the one of the previous section) (Fig. 52).

Fig. 52 Morphogenetic chart of the arrow-hexagons contained inside the oppositional tetrahexahedron known so far



Note that, in the middle line, we have a left–right symmetry with respect to being green or blue (whereas we have no top–bottom symmetry; the structure of the strong hexagons cannot be said to be symmetric with respect to that of the weak hexagons). The major gain thus obtained is that one can go, by minimal transformations (i.e. 1-moves), from one to the other of each of these five families of arrow-hexagons (we come back to this at the end of the paper). Note also that, so far, geometrically speaking, the "books" play a pivotal role. More generally the strong hexagons, the weak hexagons and the books are oppositionally self-dual; in them there is a perfect balance of blue and green).

Moreover, as one can see, one diagonal of the parrots originating the books has become an arrow (contradiction has become subalternation, i.e. implication), so that from strong to book, through weak and parrot, we observe, so to say, a growing degeneration of the chromatic internal structure of the arrow-hexagons (i.e. a loss of balanced symmetries) (Fig. 53).

The criteria for ordering such hexagons are not straightforward – they consist in comparing the diagonals, the simplexes and the vertices – so that in some cases such an ordering can be seen as a matter of choice. We will try to keep a linear order (although in

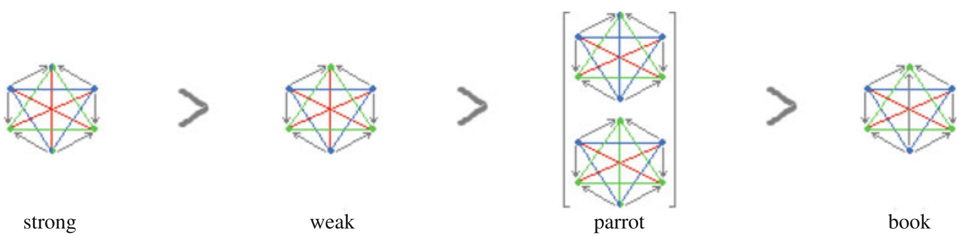
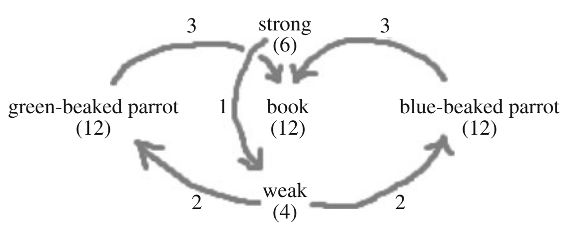


Fig. 53 Arrow-hexagons, so far, can be put in a linear order of growing "oppositional degeneration"

Fig. 54 Morphogenetic chart of "oppositional degeneration": a possible way of modelling "oppositional entropy"



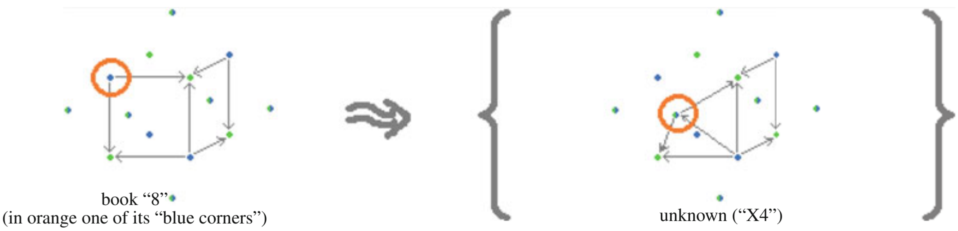


Fig. 55 The only possible 1-move transformation of a book's "blue corner"

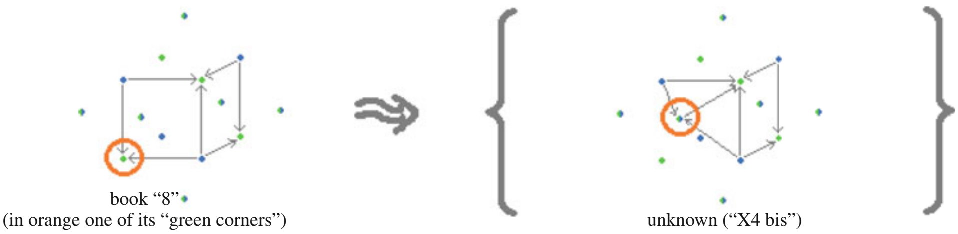


Fig. 56 The only possible 1-move transformation of a book's "green corner"

some cases, as we will see, a fuzzy order relation would be more appropriate). The linear order of the four families of arrow-hexagons considered so far ("strong > weak > parrot > book") gives the following alternative morphogenetic chart (expressing some kind of "degeneration dynamics") (Fig. 54).

This might be read as a "tendency" (or propensity) maybe pertaining to some kind of "oppositional entropy".

What are all the possible 1-move transformations of a book? Do these 1-moves yield new arrow-hexagons? Let us first consider what happens when one of the two "blue corners" (the two being symmetrical and, therefore, leading to symmetrical results) of a book is left free to move (we consider here book "8" of our previous list) (Fig. 55).

The only possible transformation leads to a figure that has not yet been seen, which we will therefore momentarily call "X4" (we will study it later on in this chapter).

Secondly, let us consider what happens if we move one of its two "green corners" (symmetry considerations will spare us the analysis of the other green corner) (Fig. 56).

The only possible transformation leads to a figure that has not yet been seen but is very similar (i.e. oppositionally dual) to the one previously called "X4". Consequently, we will call it momentarily "X4 bis" (and study it later together with "X4").

What are the possible 1-move transformations reached by a book when its "green spine" moves? (Fig. 57).

As we can see, the only 1-move transformation is the one leading to a new instance of parrot (blue-beaked), which is something we know already. Another transformation is possible, but it is a 2-moves transformation (that goes through the previous parrot), the result of which is a weak hexagon, also something that we already know.

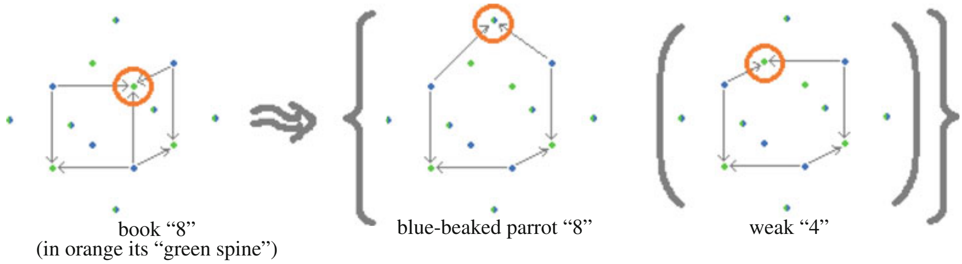


Fig. 57 The possible 1-move (and 2-moves) transformation(s) of a book's "green spine"

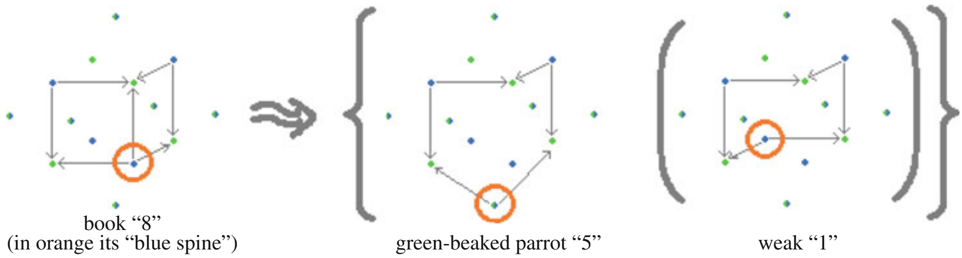


Fig. 58 The possible 1-move (and 2-moves) transformation(s) of a book's "blue spine"

Finally, what are the possible transformations of a book when the vertex left free for all possible 1-moves is the "blue spine" (we can expect this case to be oppositionally dual with respect to the previous one)? (Fig. 58)

The answer is that in this case (as in the previous one) there is only one 1-move transformation; it leads to a (green-beaked) parrot, and this is something that we already know. A 2-moves transformation (going through this parrot) leads to a weak hexagon, also something that we already know.

So, let us now turn to the study of the next new shape of the arrow-hexagon, the one momentarily called by us "X4".

3.5 *The Unknown Case "X4" Consists of "Half-Crab" Hexagons*

As we have just seen in the previous section, another transformation of the books is possible. This gives a new kind of arrow-hexagon, which we called momentarily "X4" (and "X4 bis"). From now on let us call this arrow-hexagon a "half-crab". As it happens, there are neither one (as in the case of the strong, weak and book hexagons), nor two (as in the case of the parrot hexagons), but four kinds of such half-crab hexagons. As we shall see, this is due to the fact that two symmetries are at stake here: an oppositional duality blue/green and a geometrical 3-D symmetry left/right (like for some chemical molecules, which admit a levorotatory and dextrorotatory version). Complexively, the tetrahexahe-

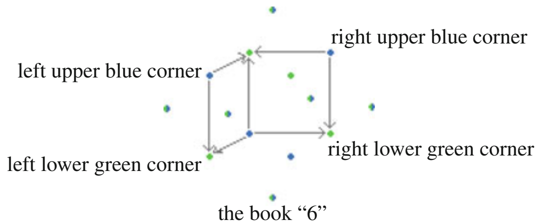


Fig. 59 Arbitrary qualitative terminology useful for handling book hexagons so to reach half-crab hexagons

dron contains 48 half-crabs. Since the only way we know (at least so far) to obtain half-crabs is from books, it is upon books that we must rely for starting our exploration. To make what follows easier, let us give fixed qualitative names to four of the six vertices of the book hexagon (Fig. 59).

The first of the four variants is what we will call the "left blue half-crabs". It is obtained by moving the "right lower green corner" of each book hexagon. As usual, to make things easier we present the 12 of them as consisting of three quartets.

The first quartet results from a 1-move (in the book hexagons 1, 2, 3 and 4, respectively) of the vertices g4, g1, g2 and g3 (Fig. 60).

The second quartet consists of a 1-move (in the book hexagons 5, 6, 7 and 8, respectively) of the vertices g3, g1, g4 and g2 (Fig. 61).

The third quartet consists of a 1-move (in the book hexagons 9, 10, 11 and 12, respectively) of the vertices g2, g4, g3 and g1 (Fig. 62).

The chromatic signature of the left blue half-crabs is the following (here we consider element "11" in our list) (Fig. 63).

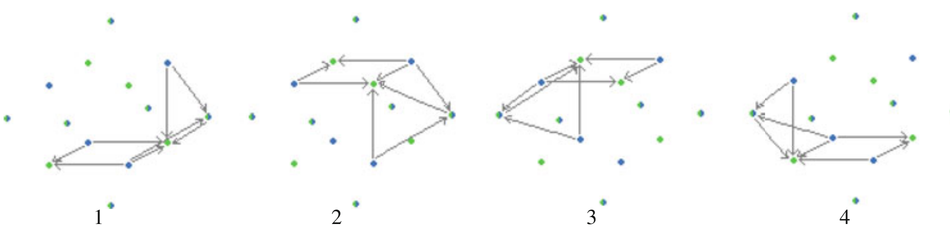


Fig. 60 The first quartet of "left blue half-crabs" (1–4)

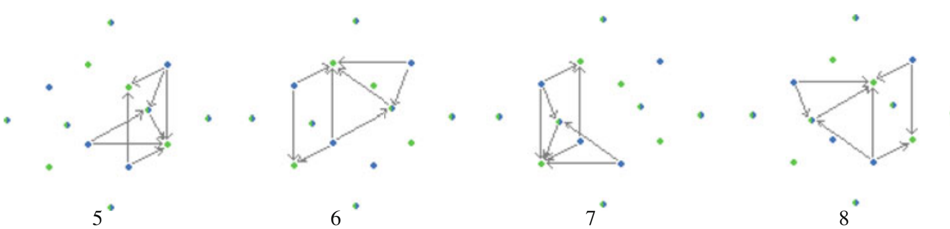


Fig. 61 The second quartet of "left blue half-crabs" (5–8)

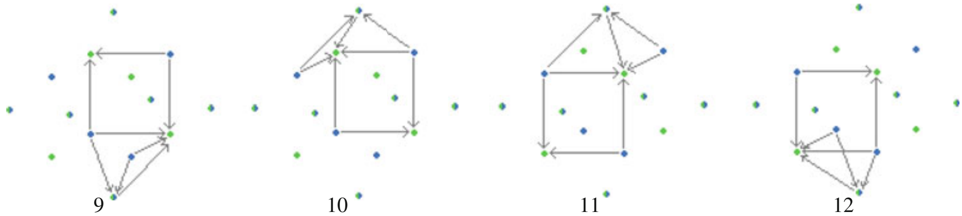


Fig. 62 The third quartet of the "left blue half-crabs" (9–12)

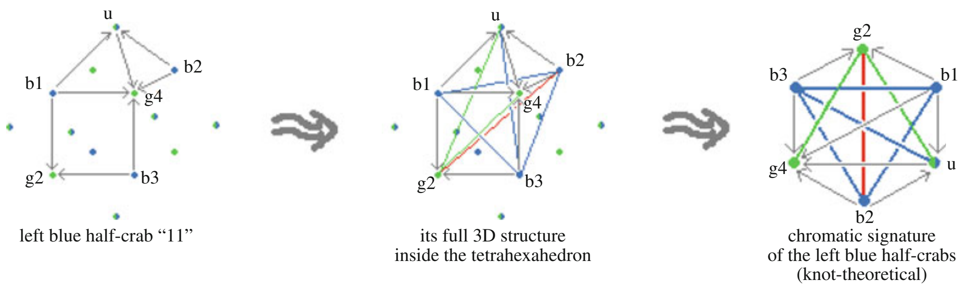


Fig. 63 The chromatic signature of the "left blue half-crab" arrow-hexagons

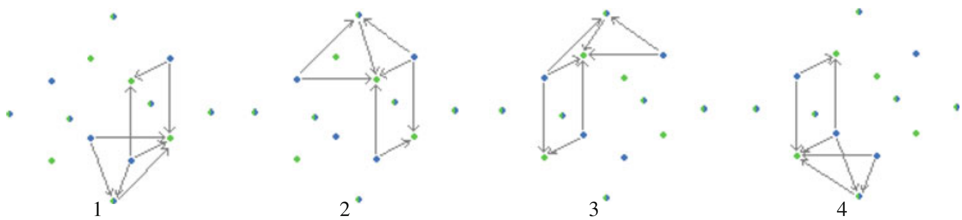


Fig. 64 The first quartet of the "right blue half-crabs" (1-4)

The second of the four kinds of half-crabs is what we will call the "right blue half-crabs". It is obtained by moving the "left lower green corner" of each book hexagon. Again, it consists of 12 elements, which we here present in three quartets.

The first quartet consists of a 1-move (in the book hexagons 1, 2, 3 and 4, respectively) of the vertices g2, g3, g4 and g1 (Fig. 64).

The second quartet consists of a 1-move (in the book hexagons 5, 6, 7 and 8, respectively) of the vertices g4, g2, g3 and g1 (Fig. 65).

The third quartet consists of a 1-move (in the book hexagons 9, 10, 11 and 12, respectively) of the vertices g3, g1, g2 and g4 (Fig. 66).

The chromatic signature of the right blue half-crabs is the following (we consider here element "2" in our list) (Fig. 67).

As we can remark, these two kinds of half-crabs (out of the four) differ with respect to their 3-D geometry (they are like mirror images, levorotatory and dextrorotatory), and this

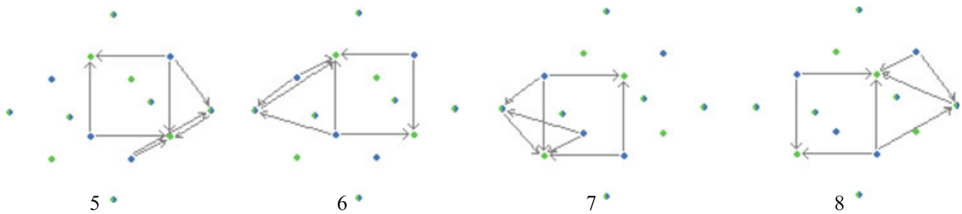


Fig. 65 The second quartet of the "right blue half-crabs" (5–8)

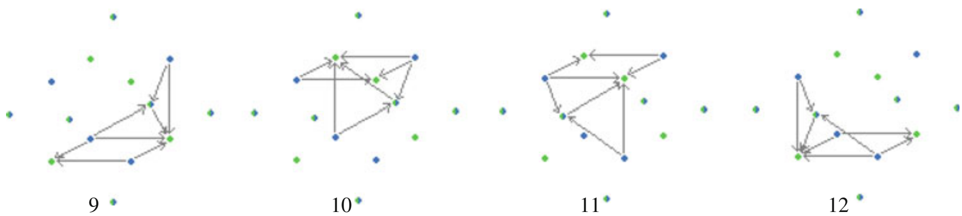


Fig. 66 The third quartet of the "right blue half-crabs" (9–12)

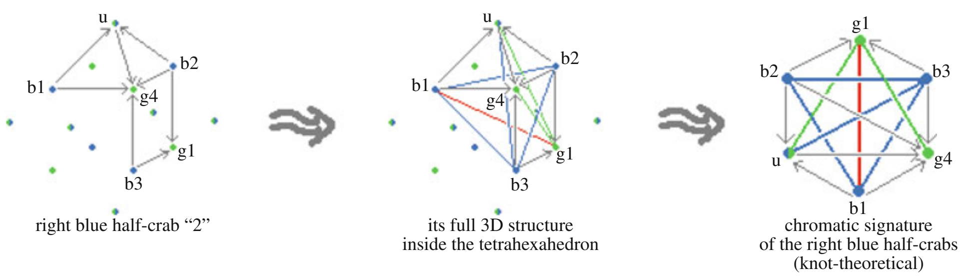


Fig. 67 The chromatic signature of the "right blue half-crab" arrow-hexagons

Fig. 68 Comparison of the (knot-theoretical) chromatic signatures of the left and right "blue half-crabs"; they are 3-D mirror images

can also be expressed with respect to their "chromatic signature", i.e. this left–right 3-D symmetry can be expressed even in 2-D space, provided that one distinguishes between what is crossing and what is crossed in an intersection (i.e. provided one resorts to knot theory instead of graph theory for building the chromatic signature) (Fig. 68).

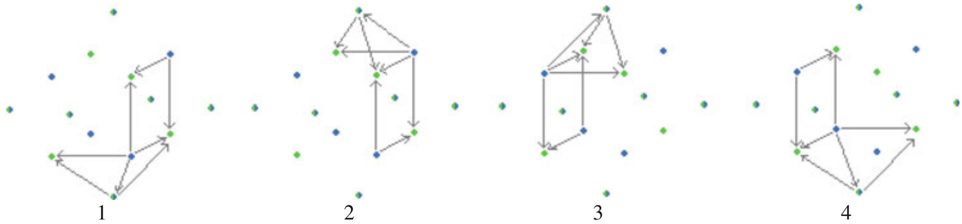


Fig. 69 The first quartet of the "left green half-crabs" (1–4)

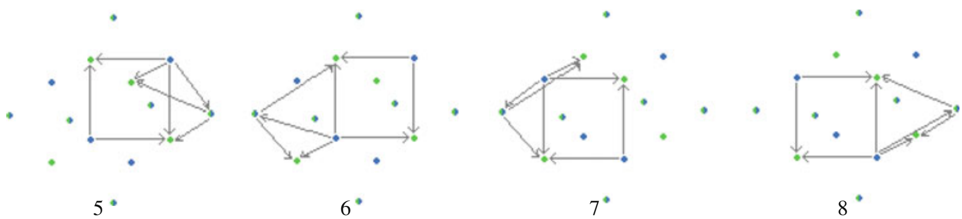


Fig. 70 The second quartet of the "left green half-crabs" (5–8)

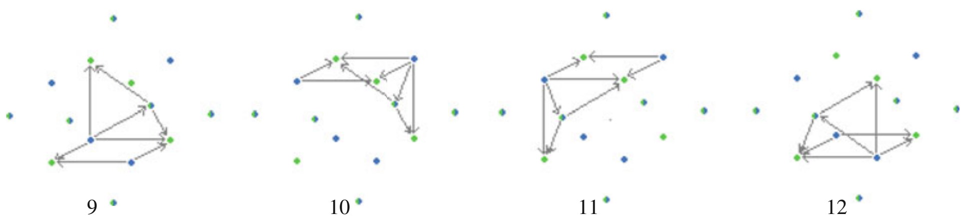


Fig. 71 The third quartet of the "left green half-crabs" (9–12)

The third of the four kinds of half-crabs is what we propose calling the "left green half-crabs". It is obtained by moving the "left upper blue corner" of each book hexagon. There are 12 instances of it, which we present here in three quartets.

The first quartet consists of a 1-move (in the book hexagons 1, 2, 3 and 4, respectively) of the vertices b4, b1, b2 and b3 (Fig. 69).

The second quartet consists of a 1-move (in the book hexagons 5, 6, 7 and 8, respectively) of the vertices b3, b1, b4 and b2 (Fig. 70).

The third quartet consists of a 1-move (in the book hexagons 9, 10, 11 and 12, respectively) of the vertices b2, b4, b3 and b1 (Fig. 71).

The chromatic signature of this first group of green half-crabs is the following (here we consider the left green half-crab "6" of our list) (Fig. 72).

The fourth and last kind of half-crab is composed of what we propose calling "right green half-crabs". They are obtained by moving the "right upper blue corner" of each book hexagon. As previously, there are 12 instances of this kind, which for graphical ease we present in three quartets.

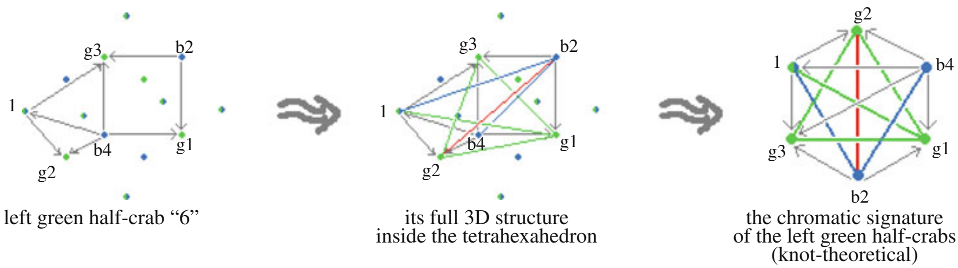


Fig. 72 The chromatic signature of the "left green half-crab" arrow-hexagons

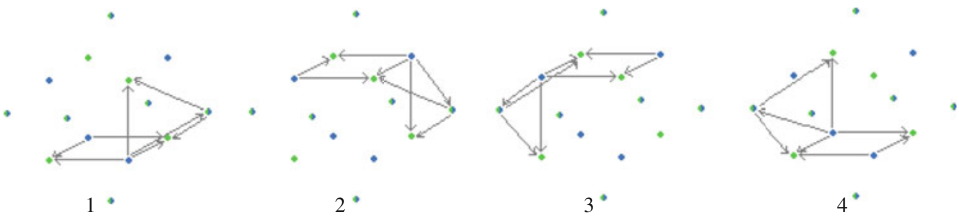


Fig. 73 The first quartet of the "right green half-crabs" (1–4)

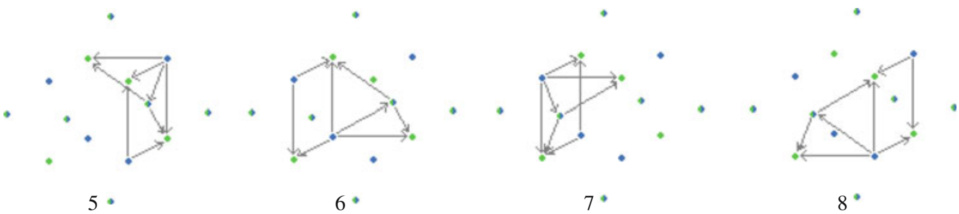


Fig. 74 The second quartet of the "right green half-crabs" (5–8)

The first quartet consists of a 1-move (in the book hexagons 1, 2, 3 and 4, respectively) of the vertices b2, b3, b4 and b1 (Fig. 73).

The second quartet consists of a 1-move (in the book hexagons 5, 6, 7 and 8, respectively) of the vertices b4, b2, b3 and b1 (Fig. 74).

The third quartet consists of a 1-move (in the book hexagons 9, 10, 11 and 12, respectively) of the vertices b3, b1, b2 and b4 (Fig. 75).

The chromatic signature of this second group of green half-crabs is the following (here we consider element "9" in our list) (Fig. 76).

As previously, for the first and the second kind of half-crabs (the left and right blue), the third and the fourth kind here (the left and right green) differ not only in their 3-D shape (they are 3-D mirror images, levorotatory and dextrorotatory), but also in their chromatic signature, provided that the expression of the latter graphically distinguishes the segments crossing from the segments being crossed (again, that is: resorting to knot theory instead of to graph theory) (Fig. 77).

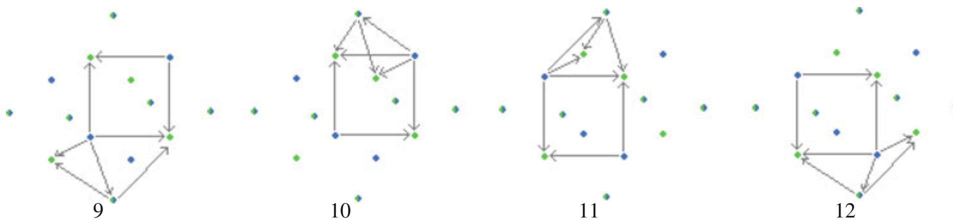


Fig. 75 The third quartet of the "right green half-crabs" (9–12)

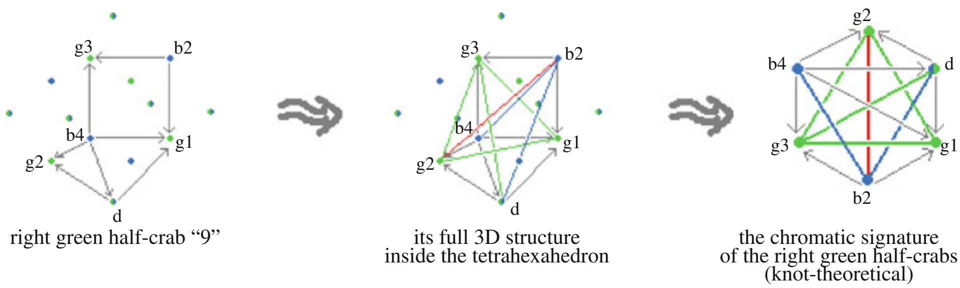


Fig. 76 The chromatic signature of the "right green half-crab" arrow-hexagons

Fig. 77 Comparison of the (knot-theoretical) chromatic signatures of the left and right "green half-crabs"; they are 3-D mirror images

With the further 3D rule (for 3D oppositional duality) of "exchanging the crossing order", the four chromatic signatures of the half-crabs are such that each one admits, among the other three, one which is oppositionally dual to it, in fact the left blue with the right green and the right blue with the left green (Fig. 78).

The half-crab can pave the surface of the oppositional tetrahexahedron, provided that one remarks that each half-crab leaves one vertex untouched in its middle (let us call this "the half-crab's prisoner"). An example is a paving of the tetrahexahedron by two blue half-crabs (one left and one right) and two green half-crabs (one left and one right). We depict these, respectively, in light blue, grey, violet and orange (Fig. 79).

At this level the chart of the possible transformations between arrow-hexagons of the $\beta3$-structure is, therefore, the following (Fig. 80).

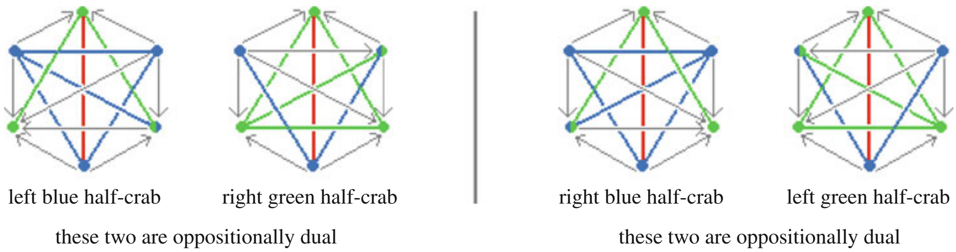


Fig. 78 The chromatic signatures of the blue and of the green half-crabs (left with right, right with left) are oppositionally dual

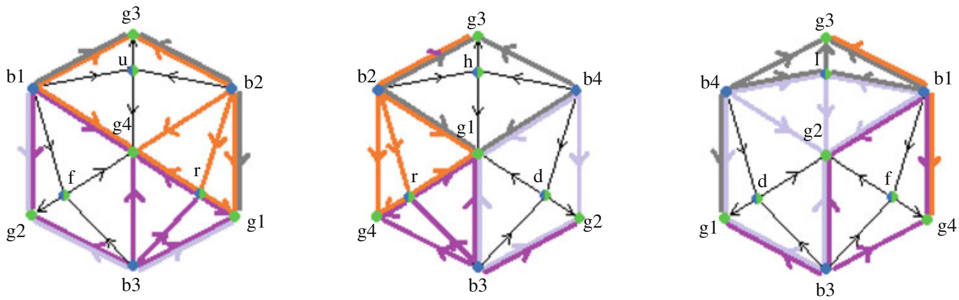
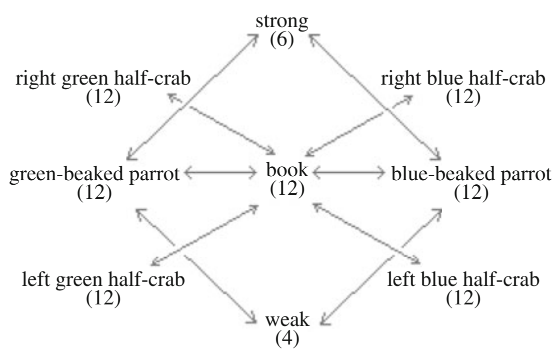


Fig. 79 The half-crab hexagons can pave (but leaving holes) the surface of the oppositional tetrahexahedron

Fig. 80 The morphogenetic chart of the arrow-hexagons of the oppositional tetrahexahedron know so far



The left–right symmetry, which is a blue–green duality, is kept. However, here a further symmetry appears, that between "right" and "left" half-crabs, which can be expressed by turning this morphogenetic chart from 2-D to 3-D. The right "isomers" (of the half-crabs), i.e. the levorotatory instances, belong to the "rear", while the left "isomers" (of the half-crabs), i.e. the dextrorotatory instances, belong to the front of the chart.

One should remark here that the chromatic signature of the half-crabs becomes even more different – with respect to the one of the strong and weak hexagons – than that of the parrots or of the books: one diagonal has become an arrow, another one has become

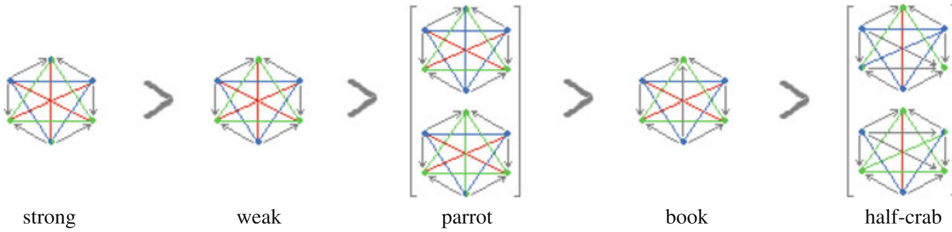
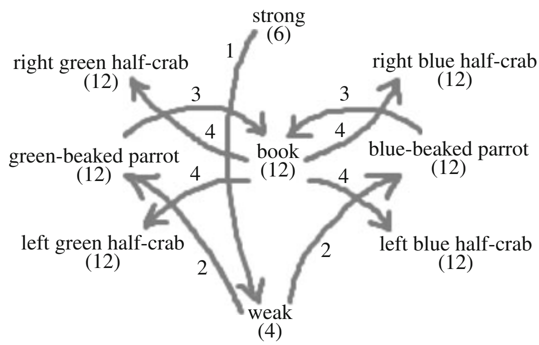


Fig. 81 Arrow-hexagons, so far, can be put in linear order of growing "oppositional degeneration"

Fig. 82 The "morphogenetic degeneration chart": a possible way of modelling "oppositional entropy"



green (or blue) and one side of the green (or blue) triangle has become an arrow. If one tries to catch this change with respect to the "degree of chromatic degeneration", this gives a linear ordering (strong > weak > parrot > book > half-crab) of the arrow-hexagons (Fig. 81).

This linear ordering of the five families of arrow-hexagons considered so far gives, in turn, the following updated version of the "morphogenetic degeneration chart" (Fig. 82).

Are there any more possible transformations? In a half-crab all the vertices are qualitatively different (for there are no symmetries inside the half-crab itself). So, all vertices, and their movements, must be studied singularly.

Let us first consider the 1-move transformations of a half-crab made possible when its "pincer" is left free to move. Here we consider the left blue half-crab number "8" in our previous list (Fig. 83).

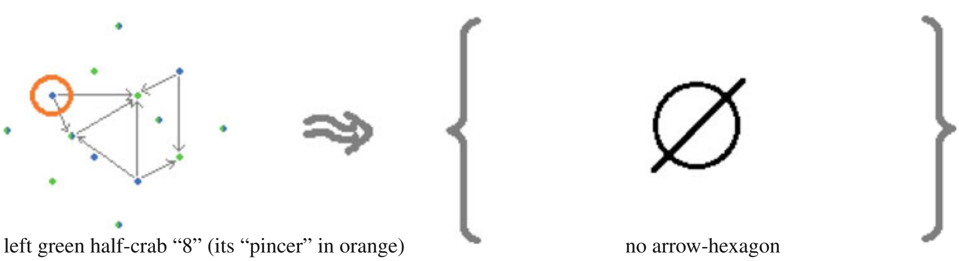
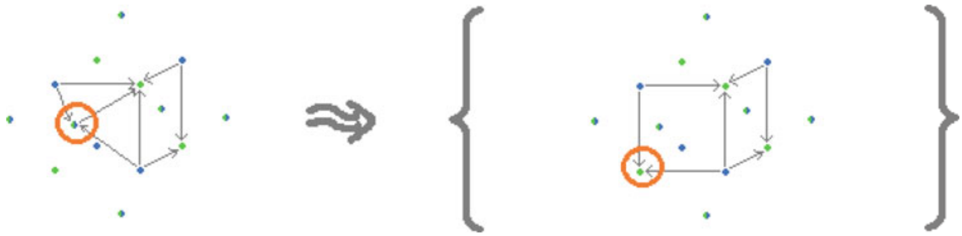
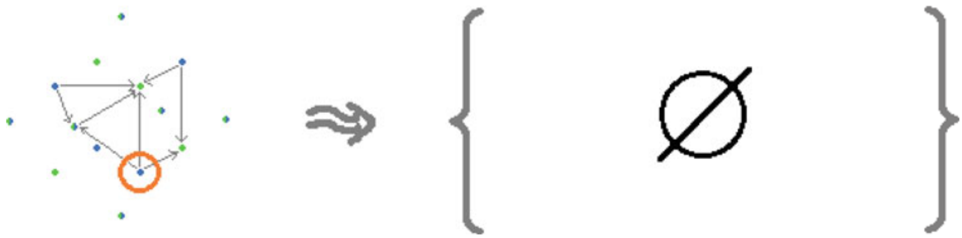


Fig. 83 A half-crab's "pincer" admits neither a 1-move nor a 2-moves transformation



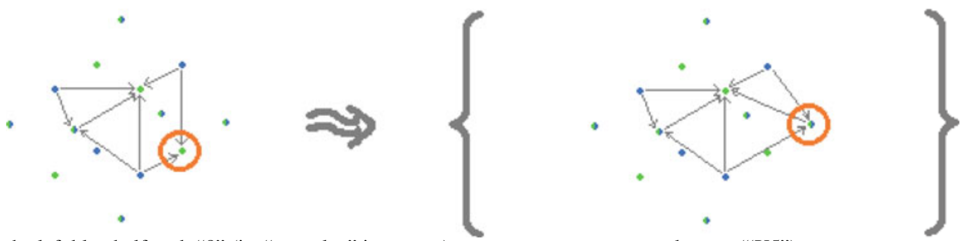
the left blue half-crab “8” (its “elbow” in orange) book “8”

Fig. 84 The only possible 1-move transformation of a half-crab through its “elbow” leads to a book



the left blue half-crab “8” (its “inner leg” in orange) no arrow-hexagon

Fig. 85 Through its “inner leg” a half-crab admits neither a 1-move nor a 2-moves transformation



the left blue half-crab “8” (its “outer leg” in orange) unknown (“X5”)

Fig. 86 The only possible 1-move transformation of a half-crab through its “outer leg” leads to a new kind of arrow-hexagon

As it happens, no transformation is possible in this case.

Secondly, let us consider what happens when it is a half-crab’s “elbow” that is left free to move (Fig. 84).

As it turns out, only one 1-move transformation is possible in this case, namely one which gives a book (here: book “8”).

Thirdly, let us consider what transformations of a half-crab are possible when what moves is its “inner leg” (i.e. the “leg” closest to the half-crab’s pincer) (Fig. 85).

As it turns out, no transformation is possible in this case.

Fourthly, let us consider the 1-move transformations that are possible when what is left free to move is the half-crab’s “outer leg” (i.e. the “leg” most distant from the pincer) (Fig. 86).

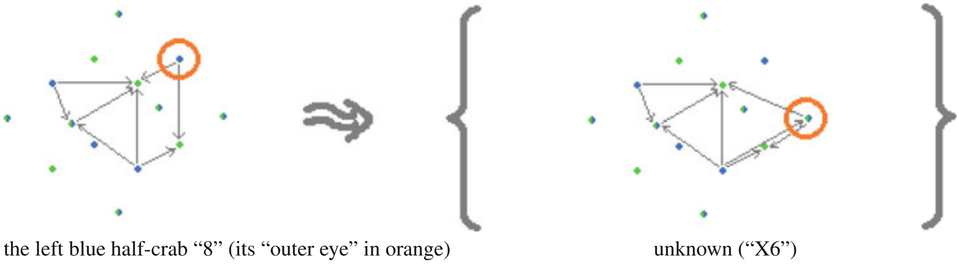


Fig. 87 The only possible 1-move transformation of a half-crab through its "outer eye" leads to a new kind of arrow hexagon

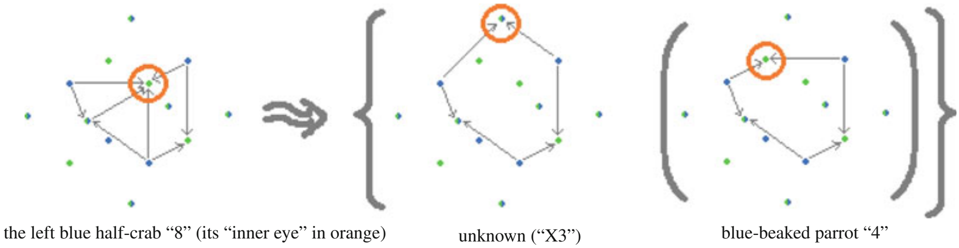


Fig. 88 The possible 1-move (and 2-moves) transformation(s) of a half-crab's "inner eye" leads to a new kind of arrow-hexagon

As it happens, the only possible outcome is a new kind of arrow-hexagon, still unknown, which we, therefore, will call momentarily "X5".

Fifthly, let us consider the 1-move transformations that are possible when the free vertex is the half-crab's "outer eye" (i.e. the "eye" most distant with respect to the pincer) (Fig. 87).

The only possible result in that case is an arrow-hexagon of a new kind, never seen before, which we will therefore call momentarily "X6".

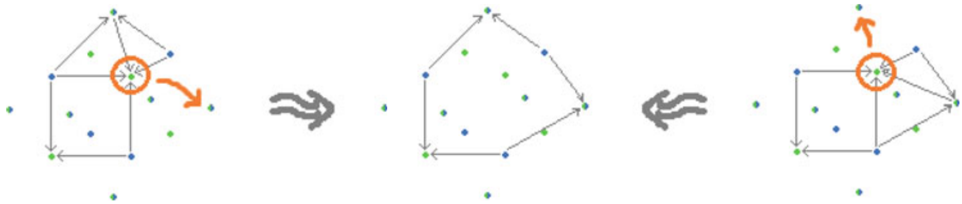
Finally, let us consider the 1-move transformations possible when the free vertex of the half-crab is its "inner eye" (i.e. the "eye" closest to the pincer) (Fig. 88).

As one can see, the only possible viable 1-move generates a transformation of the half-crab into a new kind of arrow-hexagon, which we already saw was reachable from the parrots, and we call it "X3". Another transformation is possible, but it is a 2-moves one, leading to a (here blue-beaked) parrot.

So, being finished with half-crabs, let us now study the arrow-hexagon "X3".

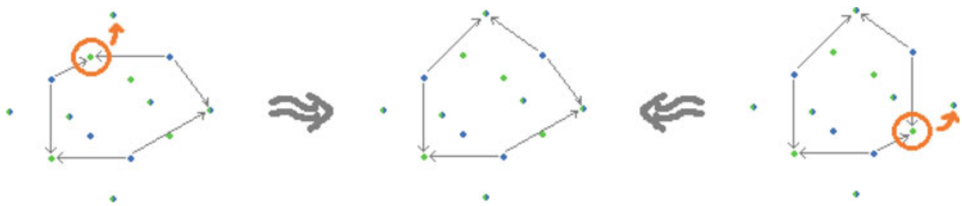
3.6 *The Unknown Case "X3" Consists of "Tortoise" Hexagons*

As we have seen, a new kind of arrow-hexagon, "X3", can be reached by a simple transformation in two ways. The first way is from half-crabs: a given blue (or green) "tortoise"



left blue half-crab "11" blue tortoise (= "X3") right blue half-crab "8"

Fig. 89 "Tortoise" arrow-hexagons can be obtained by a 1-move transformation of half-crab arrow-hexagons



blue-beaked parrot "12" blue tortoise (= "X3") blue-beaked parrot "8"

Fig. 90 "Tortoise" arrow-hexagons can be obtained by a 1-move transformation of parrot arrow-hexagons

– from now on this is what we will call the "X3" – is reachable either from a left or from a right blue (or green) half-crab (Fig. 89).

The second way is from parrots: any given blue (or green) tortoise is reachable from a pair of blue (or green) parrots (Fig. 90).

As for the several other kinds of arrow-hexagons that we have already seen, tortoises can be blue or green. Each of these two variants of tortoises has 12 instances, which we present in three quartets for graphical ease.

The first quartet of blue tortoises (labelled 1–4) can first be obtained from the blue half-crabs; either from the left blue half-crabs, namely 9, 11, 10, 12, by a 1-move of, respectively, their vertices g1, g4, g3, g2; or from the right blue half-crabs, namely 5, 8, 6, 7, by a 1-move of, respectively, their vertices g1, g4, g3, g2. Alternatively, this first quartet can be obtained from the blue-beaked parrots: either from 11, 12, 9, 10, by a 1-move of, respectively, their vertices g2, g3, g4, g1; or from 7, 8, 5, 6, by a 1-move of, respectively, their vertices g4, g1, g2, g3 (Fig. 91).

By a similar reasoning, the second quartet of blue tortoises (5–8) can be obtained first from the half-crabs; either from the left blue half-crabs, namely 1, 3, 4, 2, by a 1-move of, respectively, their vertices g1, g3, g2, g4; or from the right blue half-crabs, namely 9, 10, 12, 11, by a 1-move of, respectively, their vertices g1, g3, g2, g4. Alternatively, this second quartet can be obtained from the blue-beaked parrots; either from 3, 1, 2, 4, by a 1-move of, respectively, their vertices g4, g2, g3, g1; or from 11, 9, 10, 12, by a 1-move of, respectively, their vertices g3, g1, g4, g2 (Fig. 92).

Finally, the third and last quartet of blue tortoises (9–12) can be obtained either from the left blue half-crabs, namely 5, 6, 8, 7, by a 1-move of, respectively, their vertices g1,

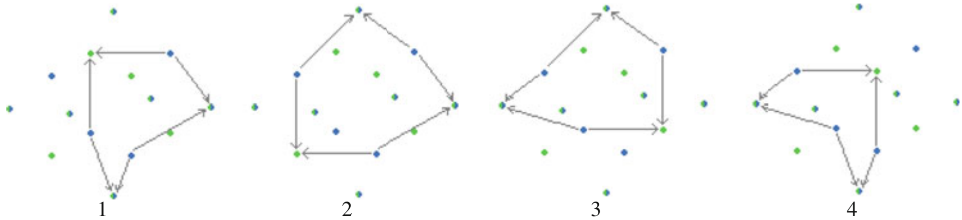


Fig. 91 The first quartet of "blue tortoises" (1–4)

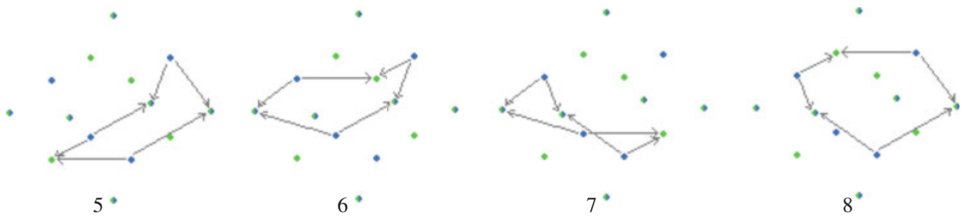


Fig. 92 The second quartet of "blue tortoises" (5–8)

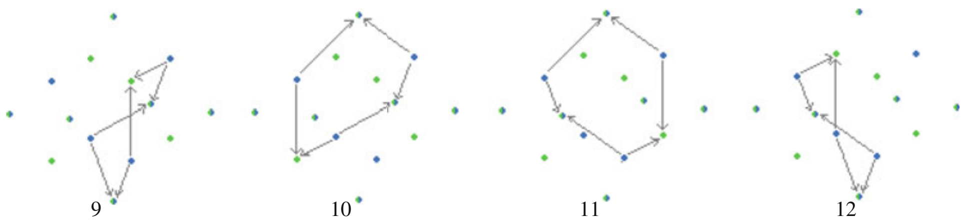


Fig. 93 The third quartet of "blue tortoises" (9–12)

g3, g4, g2; or from the right blue half-crabs, namely 1, 3, 2, 4, by a 1-move of, respectively, their vertices g1, g3, g4, g2. Alternatively, this third quartet can be obtained from the blue-beaked parrots: either from 3, 1, 4, 2, by a 1-move of, respectively, their vertices g2, g4, g3, g1; or from 7, 5, 8, 6, by a 1-move of, respectively, their vertices g3, g1, g2, g4 (Fig. 93).

Consideration of the signature of the blue tortoises obliges us to introduce an interesting new feature: the oppositional relation of "independence" (represented in orange). The independence relation, which had already been discussed *per se* by other authors (e.g. A. Sion), was discovered *geometrically* by H. Smessaert in about 2010. In the $\beta 3$-structure it holds precisely between each pair of "spikes" (with the exception of those pairs that are centrally symmetric, and therefore mutually contradictory). Considered as a whole, the 12 orange segments of independence of an oppositional tetrahexahedron form an "octahedron of independence" (Fig. 94).

If we come back now to the blue tortoises by considering their chromatic signature, we can see that in each of them a relation of independence appears and it concerns one of the three sides of what, in a standard oppositional hexagon (i.e. a strong or a weak one),

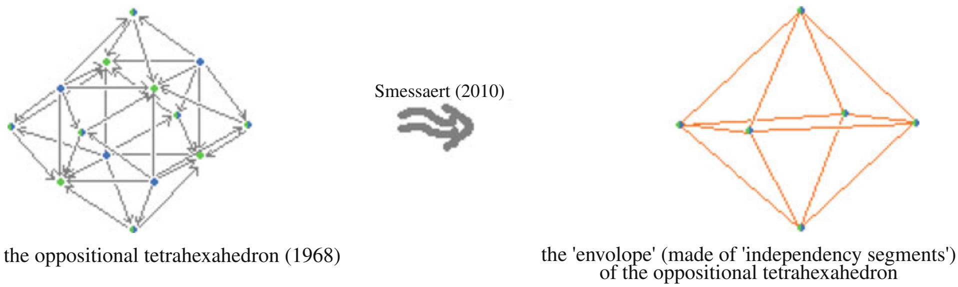


Fig. 94 Each oppositional tetrahexahedron has an octahedral "envelope" made of "independence relations" (after Smessaert)

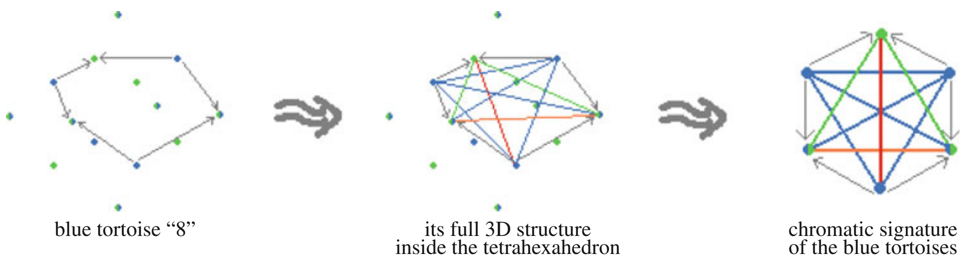


Fig. 95 The chromatic signature of the "blue tortoise" arrow-hexagons

would have been its green hexagon of subcontrariety: a side that has now become orange (here we consider blue tortoise "8" in our list) (Fig. 95).

When the aforementioned transformation of parrots into tortoises concerns green-beaked parrots or green half-crabs (left or right), the result is a "green tortoise". Here as well, there are 12 of them, which for graphical ease we present in three quartets.

The first quartet of green tortoises (1–4) can be obtained either from the left green half-crabs, namely 8, 5, 7, 6, by a 1-move of, respectively, their vertices b3, b2, b1, b4; or from the right green half-crabs, namely 12, 10, 11, 9, by a 1-move of, respectively, their vertices b3, b2, b1, b4. Alternatively, this first quartet can be obtained from the green-beaked parrots: either from 9, 10, 11, 12, by a 1-move of, respectively, their vertices b4, b1, b2, b3; or from 5, 6, 7, 8, by a 1-move of, respectively, their vertices b2, b3, b4, b1 (Fig. 96).

The second quartet of green tortoises (5-8) can be obtained either from the left green half-crabs, namely 10, 9, 11, 12, by a 1-move of, respectively, their vertices b2, b4, b1, b3; or from the right green half-crabs, namely 2, 4, 3, 1, by a 1-move of, respectively, their vertices b2, b4, b1, b3. Alternatively, this second quartet can be obtained from the green-beaked parrots: either from 2, 4, 3, 1, by a 1-move of, respectively, their vertices b3, b1, b4, b2; or from 10, 12, 11, 9, by a 1-move of, respectively, their vertices b4, b2, b3, b1 (Fig. 97).

The third and last quartet of green tortoises (9–12) can be obtained either from the left green half-crabs, namely 4, 2, 3, 1, by a 1-move of, respectively, their vertices b4, b2, b1,

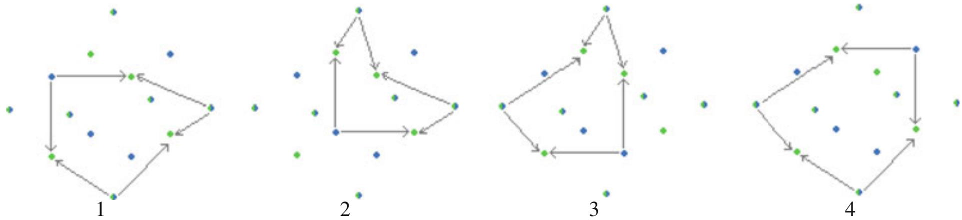


Fig. 96 The first quartet of the "green tortoises" (1–4)

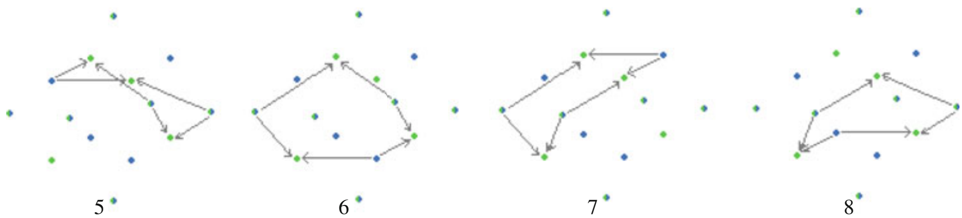


Fig. 97 The second quartet of the "green tortoise" (5–8)

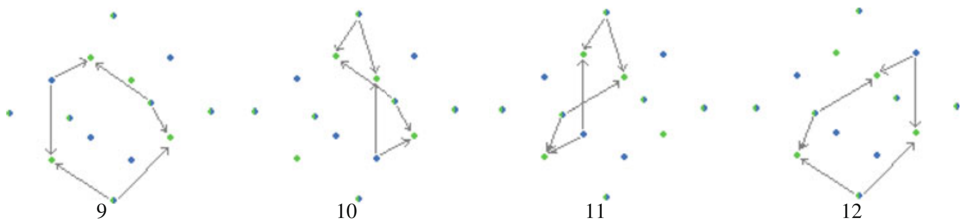


Fig. 98 The third quartet of "green tortoises" (9–12)

b3; or from the right green half-crabs, namely 6, 5, 7, 8, by a 1-move of, respectively, their vertices b4, b2, b1, b3. Alternatively, this third quartet can be obtained from the green-beaked parrots: either 4, 2, 3, 1, by a 1-move of, respectively, their vertices b3, b1, b2, b4; or from 8, 6, 7, 5, by a 1-move of, respectively, their vertices b2, b4, b3, b1 (Fig. 98).

Green tortoises have, of course, a chromatic signature of their own (here we consider green tortoise number "6" in our list) (Fig. 99).

As one can see, the chromatic signature of the green tortoises is oppositionally dual to that of the blue tortoises (Fig. 100).

As it happens, this kind of arrow-hexagon is the one we found while studying the oppositional geometry of partial order as distinct from the simpler oppositional geometry of total order – this is Sesmat–Blanché's hexagon of orders (cf. Sect. 2). So it can be suggested (at least conjecturally) that blue tortoises are "attractors" with respect to what we propose calling an "oppositional expansion"; each time a $\beta 2$-structure expands into a $\beta 3$-structure (because some new oppositional position is added, like "||" in the case of order), the six entities which initially decorated the $\beta 2$-structure (i.e. the six vertices of

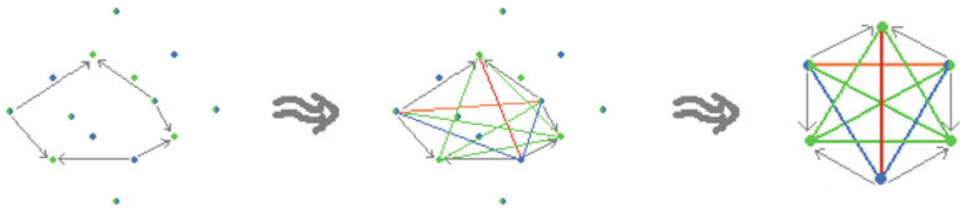


Fig. 99 The chromatic signature of the "green tortoise" arrow-hexagons

Fig. 100 From the viewpoint of their chromatic signature the blue and the green tortoises are oppositionally dual

a strong hexagon) now decorate a *blue* tortoise hexagon. Let us call this blue tortoise the "oppositional shadow" in the $\beta 3$-structure of the previous $\beta 2$-structure. A natural question is that of knowing whether the green tortoises are, as the blue ones, attractors of something. If this is the case, it could be said that tortoises have an oppositional role.

As for paving (by comparison with what we saw in Sects. 3.4 and 3.5 above on the paving power of books and half-crabs), one can remark that tortoises (as well as half-crabs) cannot pave the tetrahexahedron *stricto sensu*: they "cut" its volume (as do the strong hexagons) instead of uniformly covering regions of its surface (by "possessing" all the vertices of the covered region). Inside the perimeter of a tortoise hexagon two vertices of the tetrahexahedron remain, so to say, "imprisoned" (as we have seen, each half-crab has one "prisoner"). Nevertheless, in a broader sense a tortoise can be seen as paving the tetrahexahedron's surface; for a tortoise can be seen as a "tile" that covers 8 out of the 24 triangular faces of the tetrahexahedron. In this respect, a complete pavement of the tetrahexahedron needs three tortoises (here we depict the pavement by three blue tortoises, respectively depicted in orange, violet and grey). Conversely, given that there are 24 tortoises, there should be at least eight different ways of paving a tetrahexahedron with tortoises (Fig. 101).

We are now in the position of drawing a further, more complex version of the chart representing the hexagonal morphodynamics of the $\beta 3$-structure with respect to the 11 patterns (and six families) of arrow-hexagons seen so far (Fig. 102).

It seems easy to compare, as for the "degree of oppositional degeneration", the half-crabs and the tortoises. One can argue that half-crabs are more degenerate (because inside a single half-crab there is no left/right symmetry). If one says that, as we will, then the "degeneration ordering" of the arrow-hexagons known so far can be seen as remaining linear (Fig. 103).

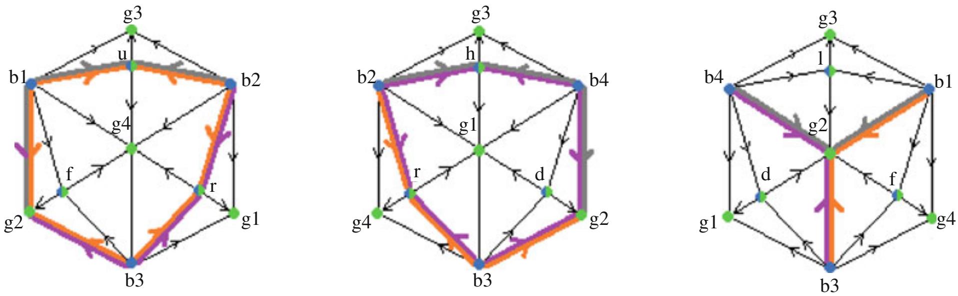


Fig. 101 Tortoise hexagons can pave (but with two holes each) the surface of the oppositional tetrahexahedron

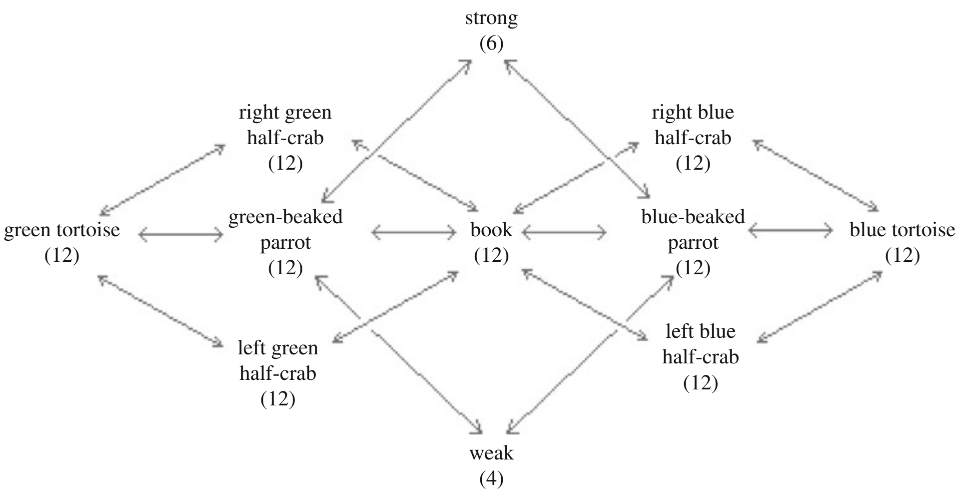


Fig. 102 The morphogenetic chart of the arrow-hexagons of the oppositional tetrahexahedron known so far

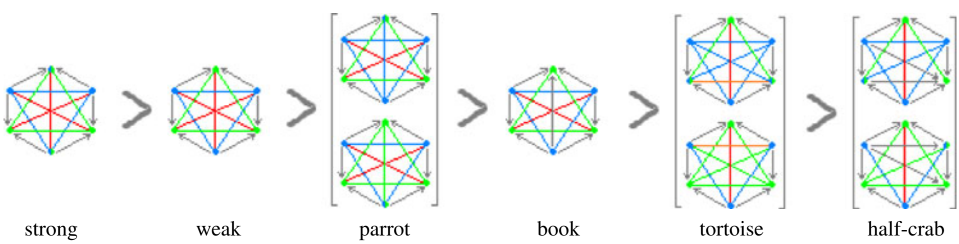


Fig. 103 Arrow-hexagons, so far, can be put in a linear order of growing "oppositional degeneration"

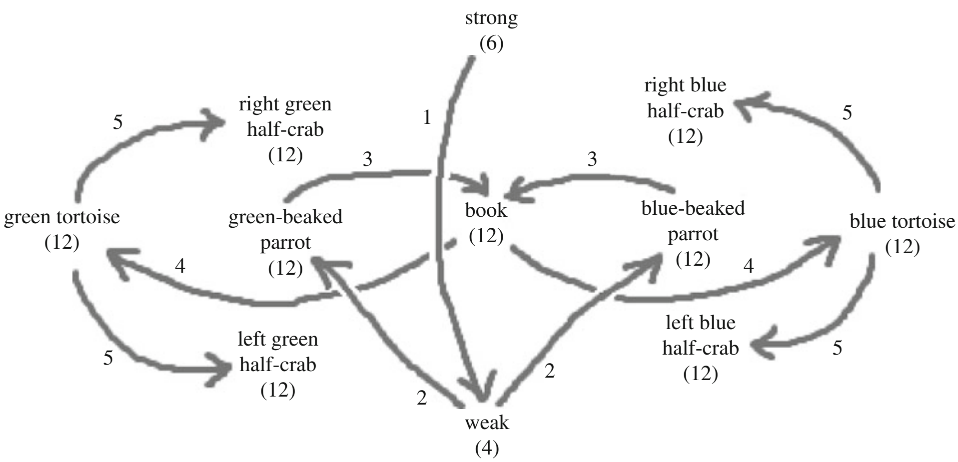


Fig. 104 The "degeneration morphogenetic chart" of the arrow-hexagons of the oppositional tetrahexahedron known so far

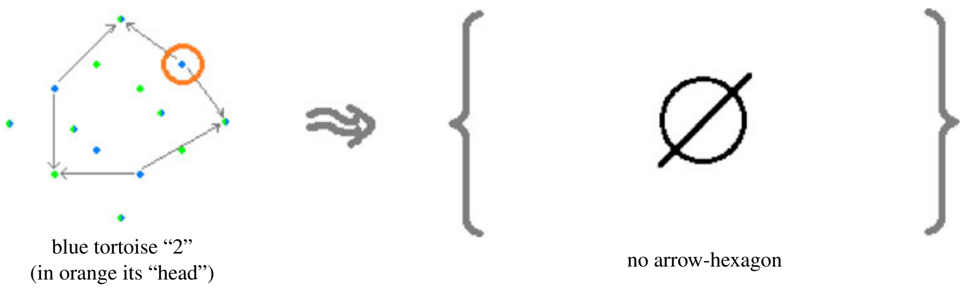


Fig. 105 No tortoise leads to a new arrow-hexagon through a 1-move transformation of its "head"

This yields the following "degeneration morphogenetic chart" of the arrow-hexagons reached so far (Fig. 104).

Again, this might turn out to be meaningful in future studies of oppositional complexity with respect to the idea of something like an "oppositional entropy".

Lastly, what are the possible 1-move transformations of a tortoise (whatsoever)? Let us consider first the 1-move transformations of a tortoise when its moving vertex is its "head" (in the rest of this section we consider the blue tortoise number "2") (Fig. 105).

As it happens, no other arrow-hexagon can be reached this way.

Secondly, let us consider the 1-move transformations of the same (blue) tortoise when its free vertex is its "left arm" (for symmetry reasons, the case of the "right arm" is comparable) (Fig. 106).

As it happens, two arrow-hexagons can be reached this way: one is a (blue-beaked) parrot (here: "12") and the other is a (right blue) half-crab (here: "8").

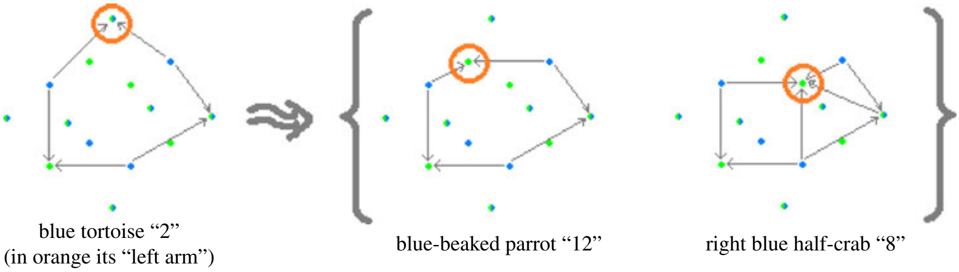


Fig. 106 The possible 1-move transformations of a tortoise through its “left arm”

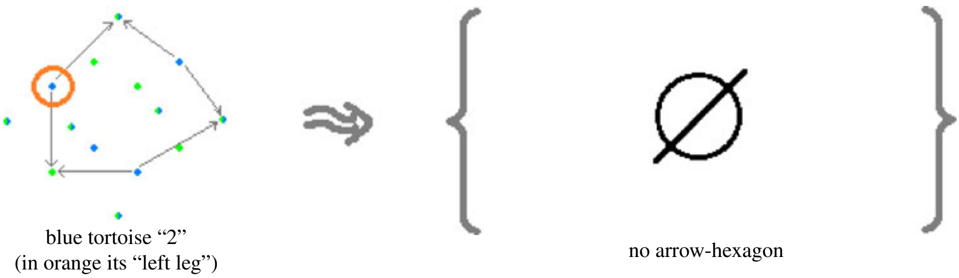


Fig. 107 No tortoise leads to another arrow-hexagon through a 1-move (or a 2-moves) transformation of one of its “legs”

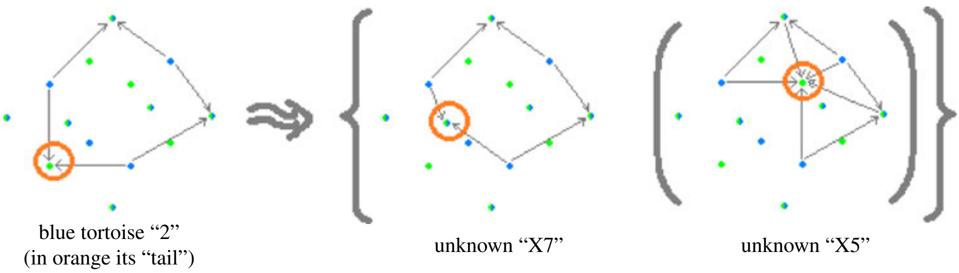


Fig. 108 The possible 1-move (and 2-moves) transformation(s) of a tortoise through its “tail”

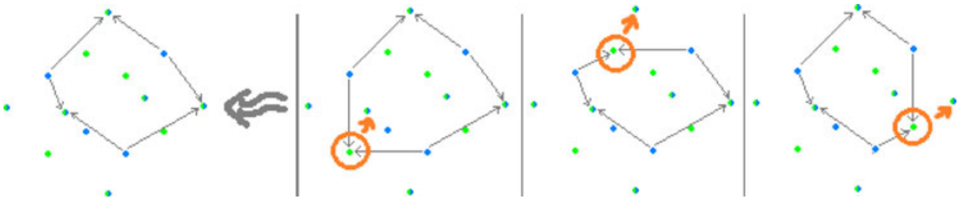
Thirdly, let us consider the outcomes of the transformation resulting from a 1-move of its “left leg” (again, because of symmetry the case of the “right leg” will be comparable) (Fig. 107).

As it turns out, no arrow-hexagon can be reached this way.

Finally, let us test the transformations resulting from the 1-move transformations of the tortoise’s “tail” (Fig. 108).

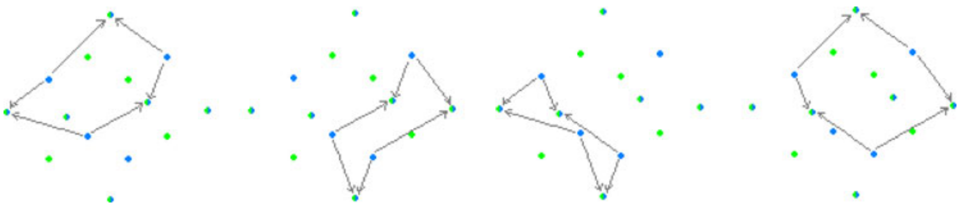
In this case, the only possible useful 1-move transformation generates a new kind of arrow-hexagon, the “X7”. Another transformation is possible by moving the tortoise’s tail, but it is a 2-moves one, which also leads to a new kind of arrow-hexagon, “X5”.

Having finished the study of the tortoises let us now turn to one of the still unknown new families of arrow-hexagons, “X7”.



blue flying squirrel (= “X7”) blue tortoise “2” blue tortoise “8” blue tortoise “11”

Fig. 109 Each “flying squirrel” can be obtained by three different tortoises of the same colour (each time by a 1-move transformation)



1 2 3 4

Fig. 110 The “blue flying squirrels” (1–4)

3.7 The Unknown Case “X7” Consists of “Flying-Squirrel” Hexagons

As we have seen, tortoise hexagons can be transformed into a new kind of arrow-hexagon, called momentarily “X7” and which we will from now on call the “flying squirrel” (Fig. 109).

The correlation is such that blue tortoises give rise to “blue flying squirrels”, whereas green tortoises give rise to green flying squirrels. There are only four blue flying squirrels (we label them 1, 2, 3 and 4). The first blue flying squirrel can be obtained from any one of the three blue tortoises 3, 6, 10, by a 1-move of, respectively, their vertices g1, g4, g2; the second blue flying squirrel is obtained from any of the three blue tortoises 1, 5, 9, by a 1-move of, respectively, their vertices g3, g2, g4; the third of the blue flying squirrels is obtained from any of the three blue tortoises 4, 7, 12, by a 1-move of, respectively, the vertices g4, g1, g3; and the fourth blue flying squirrel is obtained by any of the three blue tortoises 2, 8, 11, by a 1-move transformation of, respectively, the vertices g2, g3, g1 (Fig. 110).

The chromatic signature of the blue flying squirrels is the following (Fig. 111).

As we shall recall later, this chromatic signature seems to be rather important (for it appears in quite interesting contexts, cf. Sect. 2 above).

When instead of from blue we start with green tortoises, we obtain “green flying squirrels”. There are four of them (which we label 1, 2, 3 and 4). The first green flying squirrel is obtained from any of the three green tortoises 1, 8, 12, by a 1-move of, respectively, their vertices b1, b4, b2; the second green flying squirrel is obtained from any of the three green tortoises 3, 7, 11, by a 1-move of, respectively, their vertices b3, b2, b4; the third

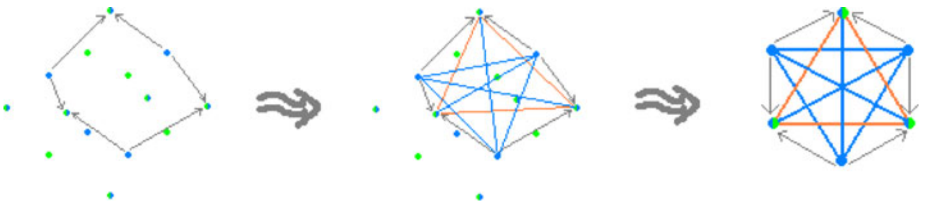


Fig. 111 The chromatic signature of the “blue flying squirrel” arrow-hexagons

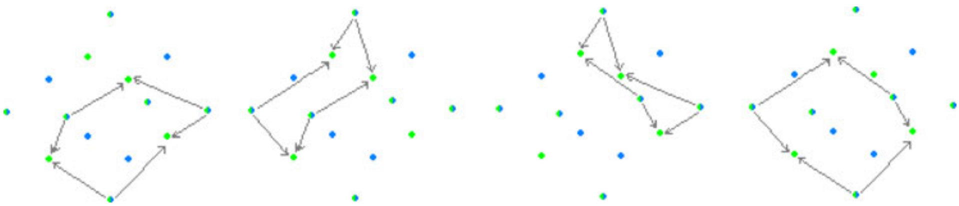


Fig. 112 The “green flying squirrels” (1–4)

green flying squirrel is obtained from any of the three green tortoises 2, 5, 10, by a 1-move of, respectively, their vertices b4, b1, b3; and the fourth green flying squirrel is obtained from any of the three green tortoises 4, 6, 9, by a 1-move of, respectively, their vertices b2, b3, b1 (Fig. 112).

The chromatic signature of the green flying squirrels is the following (Fig. 113).

Again, as we suggested when speaking about “oppositional distribution” (cf. Sect. 2), this chromatic signature seems to be quite meaningful.

As usual, the two signatures, the one of the blue and the one of the green flying squirrels, are oppositionally dual (Fig. 114).

It must be remarked that these two kinds of arrow-hexagons are attractors with respect to oppositional ∧-distribution and ∨-distribution, respectively (cf. Sect. 2 above). In other words, flying squirrels have an “oppositional role”.

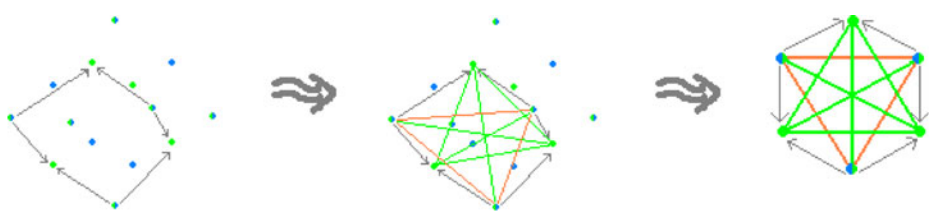


Fig. 113 The chromatic signature of the “green flying squirrel” arrow-hexagons

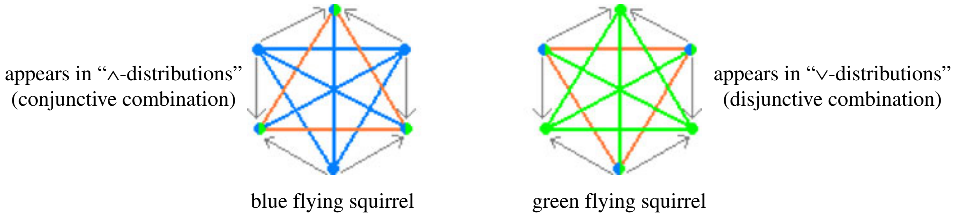


Fig. 114 From the viewpoint of their chromatic signatures the blue and the green flying squirrels are oppositionally dual

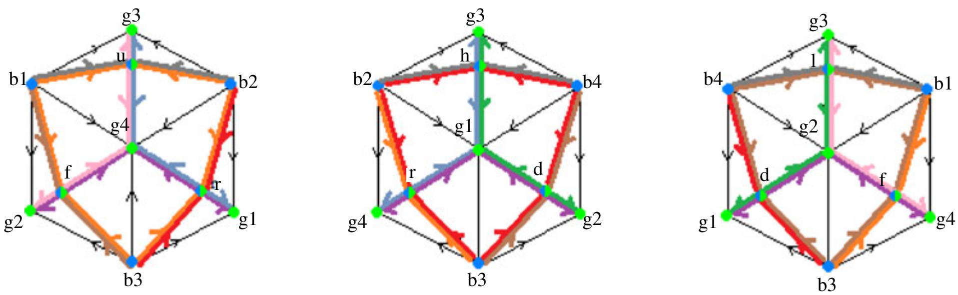


Fig. 115 The four blue flying squirrels and the four green flying squirrels can pave (but each with a hole) the tetrahexahedron

Can the squirrels "pave" the tetrahexahedron? In a sense, they can, assumed, however, that they keep one vertex inside their perimeter untouched; this is the flying squirrel's "prisoner". Since squirrels cover six faces and since the tetrahexahedron has 24 faces, a tetrahexahedron can be paved by four squirrels at a time. One pavement is given by the four blue flying squirrels (here we depict them in orange, grey, red and brown), the other is given by the four green flying squirrels (here we depict them in blue, violet, pink and green) (Fig. 115).

We can now reshape our provisory morphogenetic chart taking into account the flying squirrels (Fig. 116).

If we try to add the case of the flying squirrels to our study of the degeneration ordering of the arrow-hexagons things seem to be less simple than previously: (a) flying squirrels are indeed very symmetric, (b) but they lose much "colour" with respect to the standard of the strong hexagons. These two features make it difficult to compare the flying squirrels to the half-crabs. Nevertheless, we will suppose that it makes sense to see the flying squirrels as more degenerated than the half-crabs, thus arriving at the following linear ordering of the arrow-hexagons known so far (Fig. 117).

If we accept this ordering (but some alternatives, including partial order, could be possible without harm), this gives the following "oppositional morphogenetic degeneration chart" (Fig. 118).

Are there more arrow-hexagons reachable from the flying squirrels? In order to check this, let us study the possible changes of a flying squirrel. Flying squirrels are very sym-

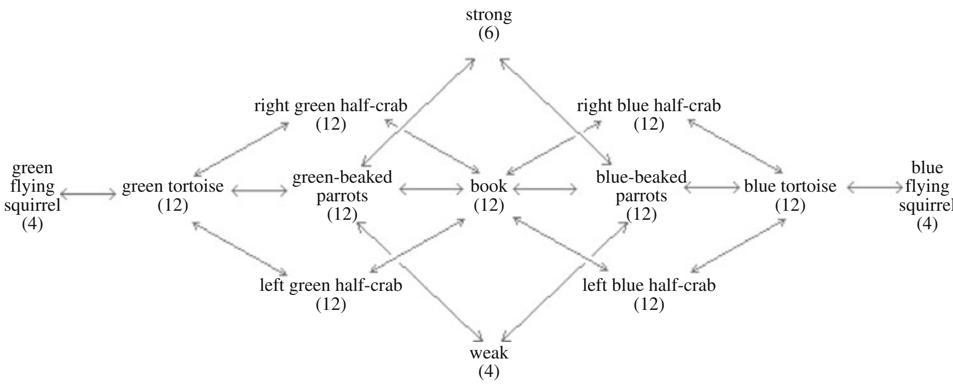


Fig. 116 The morphogenetic chart of the arrow-hexagons of the oppositional tetrahexahedron known so far

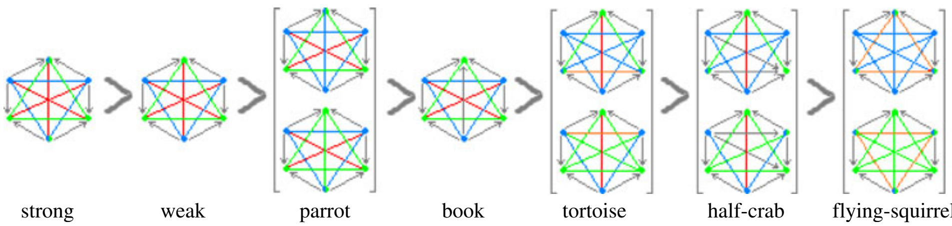


Fig. 117 Arrow-hexagons, so far, can be put in a linear ordering of growing "oppositional degeneration"

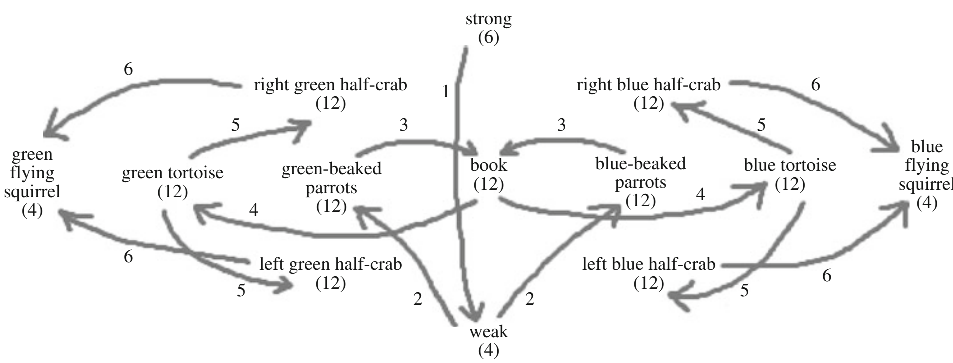


Fig. 118 The degeneration chart of the arrow-hexagons of the oppositional tetrahexahedron known so far

metrical, they have a perfect ternary symmetry, and qualitatively speaking they have only two different kinds of vertices: either blue (or green) or blue–green.

Let us consider first the transformations of a flying squirrel through a 1-move of its blue (or green) "head" (here we consider blue flying squirrel number "4") (Fig. 119).

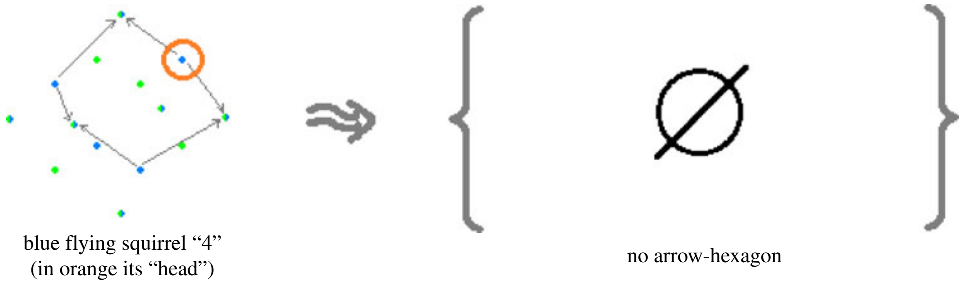


Fig. 119 No flying squirrel leads to a new arrow-hexagon through a 1-move of its "head"

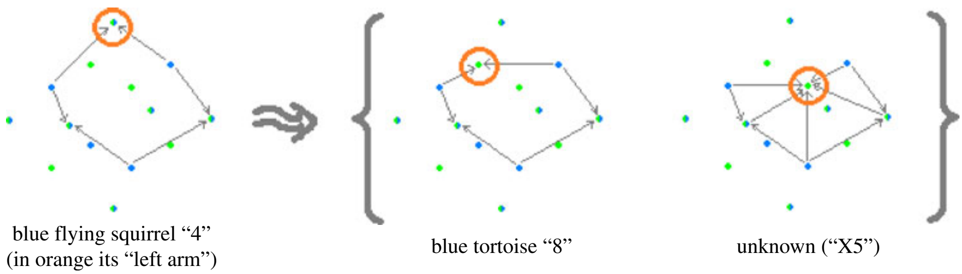


Fig. 120 The possible 1-move transformations of a flying squirrel through its "left arm"

As it happens, this leads to no other arrow-hexagon. Given the perfect ternary symmetry of the flying squirrels this negative result also concerns the other two blue (or green) vertices of the blue (or green) flying squirrel.

Secondly, let us consider the transformations occasioned by a 1-move of the flying squirrel's green-blue "left arm" (Fig. 120).

In this case, two arrow-hexagons can be obtained: either a (blue) tortoise (here: number "8"), or an unknown arrow-hexagon momentarily called "X5" (given its perfect ternary symmetry, the same result concerns the other two green–blue vertices of the flying squirrel).

Being done with the flying squirrels let us turn to one of the still unknown new families of arrow-hexagons, "X5".

3.8 *The Unknown Arrow-Hexagon "X5" Is a "Shell"*

Let us call "shells" the new kind of arrow-hexagon of the oppositional tetrahexahedron that we momentarily labelled "X5". A shell can be reached in two independent ways. Firstly, it can be obtained from the half-crabs (left or right) (Fig. 121).

Secondly, the shell can be obtained from the flying squirrels (Fig. 122).

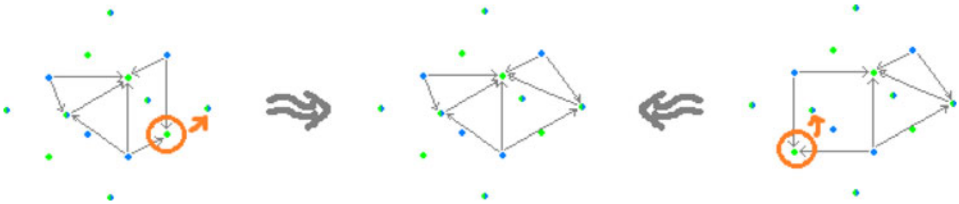


Fig. 121 "Shell" arrow-hexagons can be obtained by a 1-move transformation of half-crab arrow-hexagons

Fig. 122 "Shell" arrow-hexagons can be obtained by a 1-move transformation of flying squirrel arrow-hexagons

As it happens, in the tetrahexahedron there are 24 shells, 12 of which are "blue shells", which we present, for graphical ease, in three quartets.

The first quartet of blue shells can be obtained in three different ways. Either from the left blue half-crabs 1, 2, 3, 4, by a 1-move, respectively, of their vertices g2, g3, g4, g1; or from the right blue half-crabs 1, 2, 3, 4, by a 1-move, respectively, of their vertices g4, g1, g2, g3; or, finally, from the blue flying squirrels 2, 4, 1, 3, by a 1-move, respectively, of their vertices h, f, h, f (Fig. 123).

The second quartet of blue shells can be obtained either from the left blue half-crabs 5, 6, 7, 8, by a 1-move, respectively, of their vertices g4, g2, g3, g1; or from the right blue half-crabs 5, 6, 7, 8, by a 1-move, respectively, of their vertices g3, g1, g4, g2; or, finally,

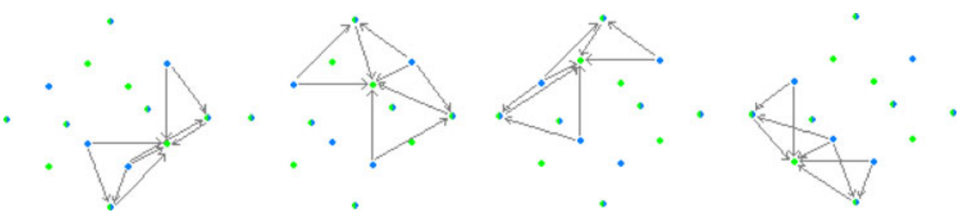


Fig. 123 The first quartet of the "blue shells" (1–4)

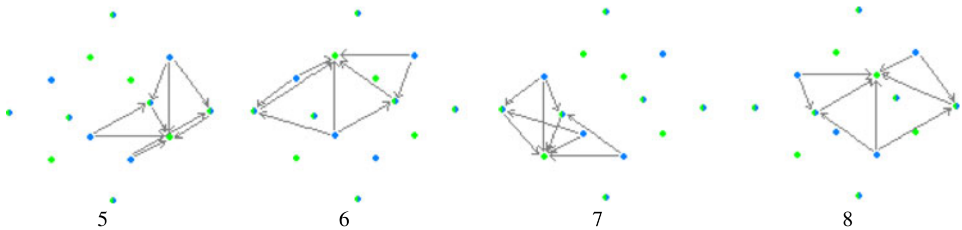


Fig. 124 The second quartet of the "blue shells" (5–8)

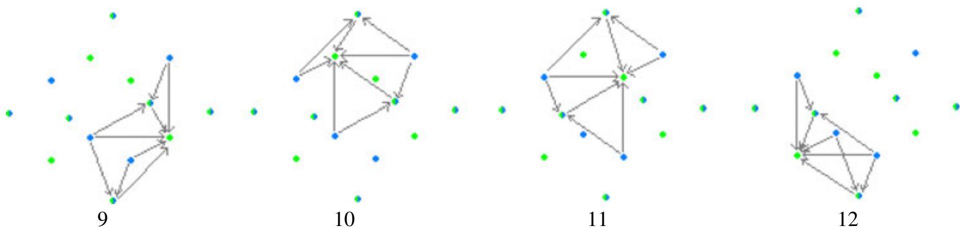


Fig. 125 The third quartet of the "blue shells" (9-12)

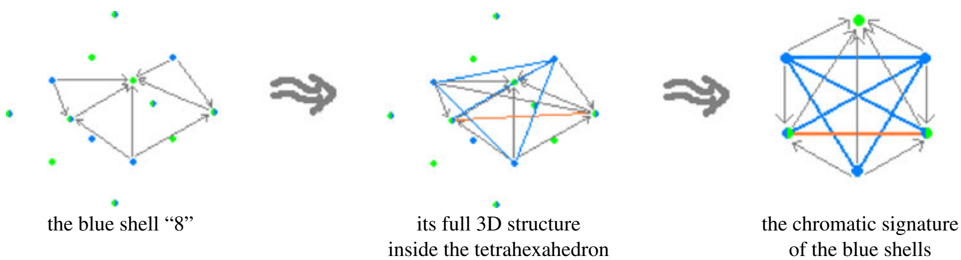


Fig. 126 The chromatic signature of the "blue shell" arrow-hexagons

from the blue flying squirrels 2, 1, 3, 4, by a 1-move, respectively, of their vertices d, u, d, u (Fig. 124).

The third quartet of blue shells can be obtained either from the left blue half-crabs 9, 10, 11, 12, by a 1-move, respectively, of their vertices g3, g1, g2, g4; or from the right blue half-crabs 9, 10, 11, 12, by a 1-move, respectively, of their vertices g2, g4, g3, g1; or, finally, from the blue flying squirrels 2, 1, 4, 3, by a 1-move, respectively, of their vertices r, l, r, l (Fig. 125).

As one can see, the blue shells have the following chromatic signature (here we consider blue shell number "8" in our list) (Fig. 126).

This signature is, so to say, weirder than any previous one. The only thing of the interior (of the hexagon) that is kept with respect to the standard oppositional hexagon (strong or weak) is the blue triangle of contrariety.

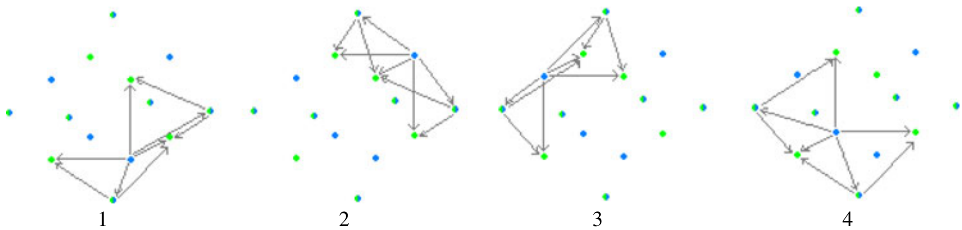


Fig. 127 The first quartet of the "green shells" (1–4)

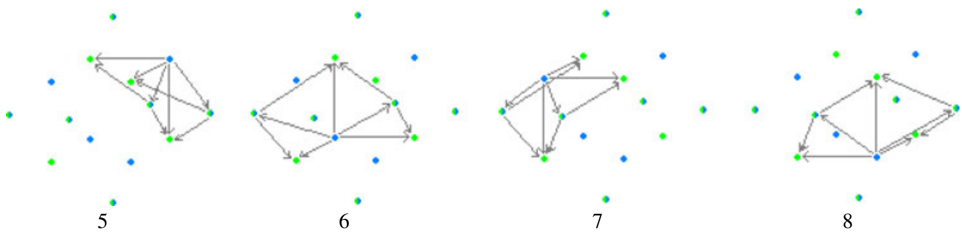


Fig. 128 The second quartet of "green shells" (5–8)

Similarly, there are 12 green shells, which we present, for graphical ease, in three quartets.

The first quartet of green shells can be obtained in three different ways. Either from the left green half-crabs 1, 2, 3, 4, by a 1-move, respectively, of their vertices b2, b3, b4, b1; or from the right green half-crabs 1, 2, 3, 4, by a 1-move, respectively, of their vertices b4, b1, b2, b3; or, finally, from the green flying squirrels 1, 2, 3, 4, by a 1-move, respectively, of their vertices f, h, f, h (Fig. 127).

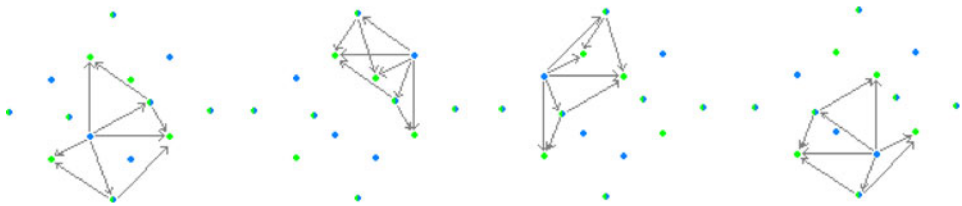
The second quartet of green shells can be obtained either from the left green half-crabs 5, 6, 7, 8, by a 1-move, respectively, of their vertices b4, b2, b3, b1; or from the right green half-crabs 5, 6, 7, 8, by a 1-move, respectively, of their vertices b3, b1, b4, b2; or, finally, from the green flying squirrels 3, 4, 2, 1, by a 1-move, respectively, of their vertices u, d, u, d (Fig. 128).

The third quartet of green shells can be obtained either from the left green half-crabs 9, 10, 11, 12, by a 1-move, respectively, of their vertices b3, b1, b2, b4; or from the right green half-crabs 9, 10, 11, 12, by a 1-move, respectively, of their vertices b2, b4, b3, b1; or, finally, from the green flying squirrels 4, 3, 2, 1 by a 1-move, respectively, of their vertices l, r, l, r (Fig. 129).

Green shells have the following chromatic signature (here we consider green shell number "12" in our list) (Fig. 130).

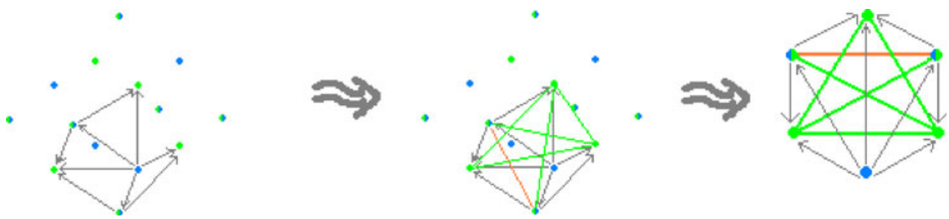
As usual, the chromatic signatures of the blue and green shells are mutually oppositionally dual (Fig. 131).

It is worth mentioning that in some sense, at least visually, this kind of hybrid hexagon appears in the aforementioned study of 1968 by Sauriol (cf. Sect. 2 above). Notice also that one can see that shells can "pave" the tetrahexahedron's surface, and differently from



9 10 11 12

Fig. 129 The third quartet of the "green shells" (9–12)



the green shell "12" — its full 3D structure inside the tetrahexahedron — the chromatic signature of the green shells

Fig. 130 The chromatic signature of the "green shell" arrow-hexagons

Fig. 131 From the viewpoint of their chromatic signatures the blue and the green shells are oppositionally dual

chromatic signature of the blue shells — chromatic signature of the green shells

the half-crabs, the tortoises and the flying squirrels, the shells do pave *perfectly* (i.e. without holes inside their own hexagonal perimeter). Since a shell covers four out of the 24 triangular faces of the tetrahexahedron, we have that one can pave the latter with six shells (Fig. 132).

From this viewpoint the shells (with four faces and six instances) and the flying squirrels (with six faces and four instances) could even be "dual" in some new geometric sense.

With this eighth new kind of oppositional arrow-hexagon we can update our provisory morphogenetic chart accordingly (Fig. 133).

If we now address the question of the "degeneration ordering" of the eight families of arrow-hexagons, the answer is not immediately straightforward. The chromatic signature of the shells has a strong left/right symmetry, but at the same time it has lost many standard colours, so it is difficult to compare it with some of the most degenerate arrow-hexagons. Let us suppose, nevertheless, that it makes sense to put the shells below the flying squirrels in the degeneration order (Fig. 134).

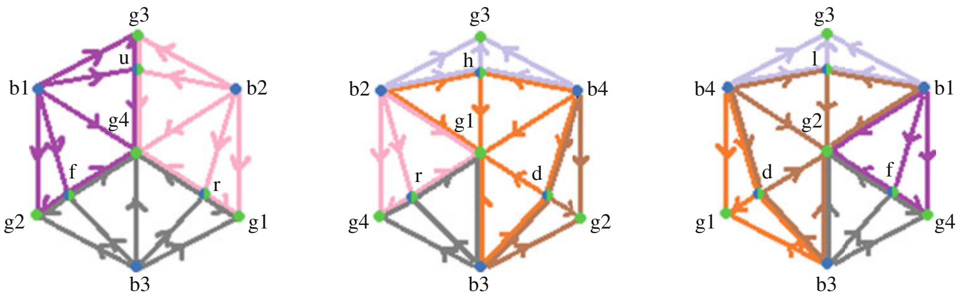


Fig. 132 One of the possible ways of paving the oppositional tetrahexahedron with six shell hexagons (three blue and three green)

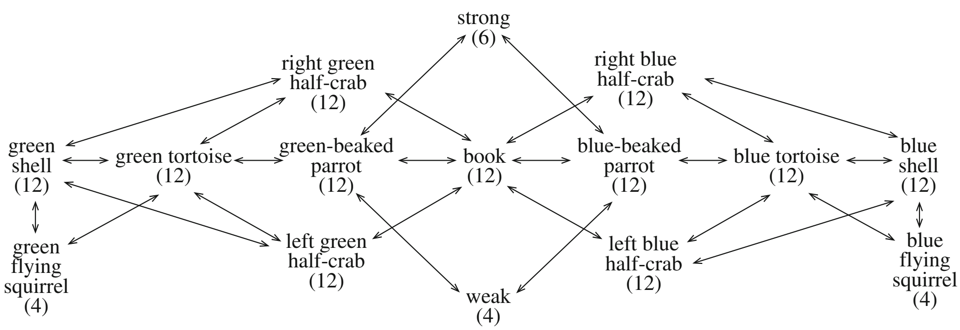


Fig. 133 The morphogenetic chart of the arrow-hexagons of the oppositional tetrahexahedron known so far

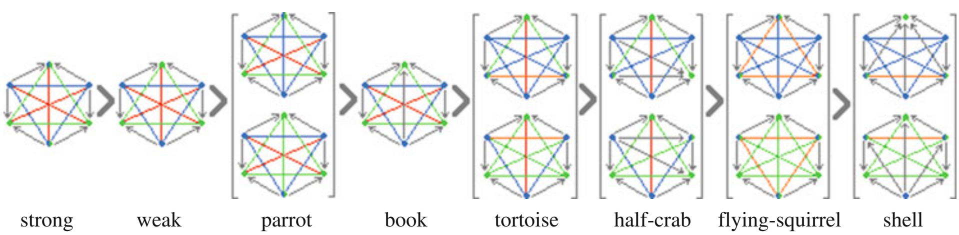


Fig. 134 Arrow-hexagons, so far, can be put in a linear order of growing "oppositional degeneration"

If we adopt this hypothesis (the linear character of the ordering of the degeneration of the signatures of the eight arrow-hexagons known so far) we can express it through the following refined version of the "morphogenetic degeneration chart" (Fig. 135).

Are there any more arrow-hexagons reachable from here? In order to answer this question we now have to examine the possible transformations of a shell.

Let us first consider the 1-move transformations of a shell made possible when its "head" is left free to move (here we consider blue shell number "8") (Fig. 136).

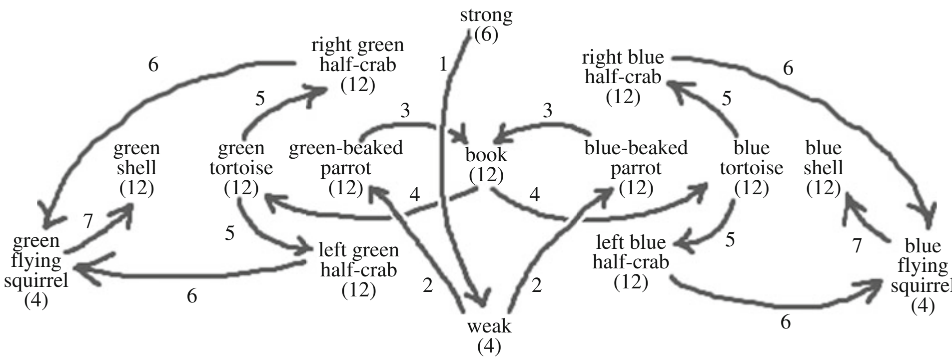


Fig. 135 The morphogenetic "degeneration chart" of the arrow hexagons of the oppositional tetrahexahedron known so far

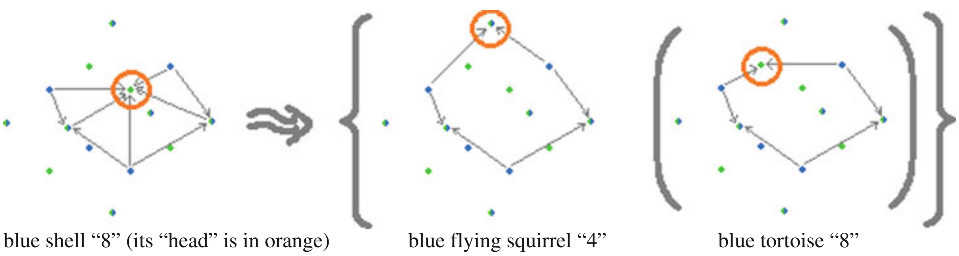


Fig. 136 The possible 1-move (and 2-moves) transformation(s) of a shell through its "head"

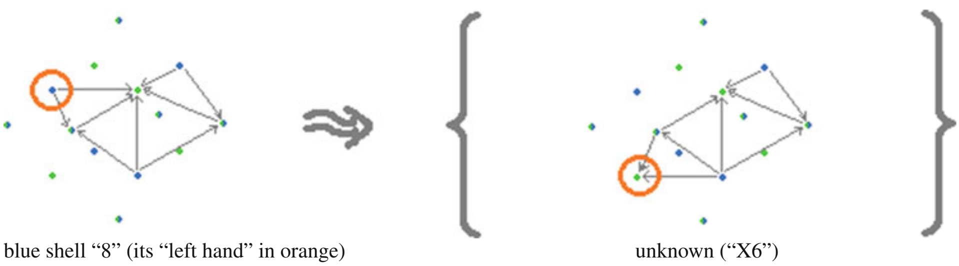


Fig. 137 The only possible 1-move transformation of a shell through one of its "hands" (left or right) leads to a new arrow-hexagon

As one can see, the only possible 1-move metamorphosis transforms the (blue) shell into a (blue) flying squirrel; another possible transformation leads to a (blue) tortoise, but this is reached through a 2-moves transformation, so that it does not count here.

Secondly, let us consider the possible transformations of a shell when the 1-move transformation that it undergoes concerns its "left hand" (the right hand will give a symmetric case) (Fig. 137).

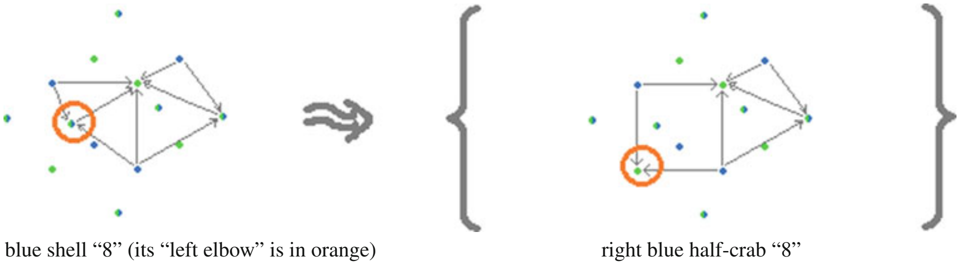


Fig. 138 The only possible 1-move transformation of a shell through its "left elbow"

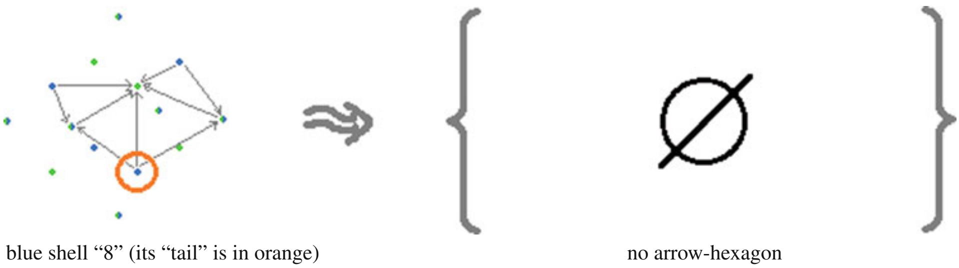


Fig. 139 No shell leads to another arrow-hexagon through a 1-move transformation of its "tail"

As we can see, the only possible viable outcome of this is a new kind of arrow-hexagon, namely one already mentioned by us but not yet treated (we will do this in the next section), "X6".

Thirdly, let us consider the 1-move transformations of a shell's "elbow" (Fig. 138).

As one can see, there is only one possible viable outcome here, which is a right (blue) half-crab.

Finally, let us consider the 1-move transformations of a shell's "tail" (Fig. 139).

As one can see, this does not lead to a viable solution (i.e. no arrow-hexagon can be obtained this way).

Having examined the shells, let us now turn to the still unknown family of arrow-hexagons, "X6".

3.9 The Unknown Arrow-Hexagons "X6" Are "Shrimps"

As we have seen, one more family of arrow-hexagons of the oppositional tetrahexahedron, "X6" – from now on "shrimps" – can be reached in two ways. Firstly, the shrimps can be reached from the shells (Fig. 140).

Secondly, the shrimps can be reached from the half-crabs (Fig. 141).

There are 24 shrimps, 12 left shrimps and 12 right shrimps.

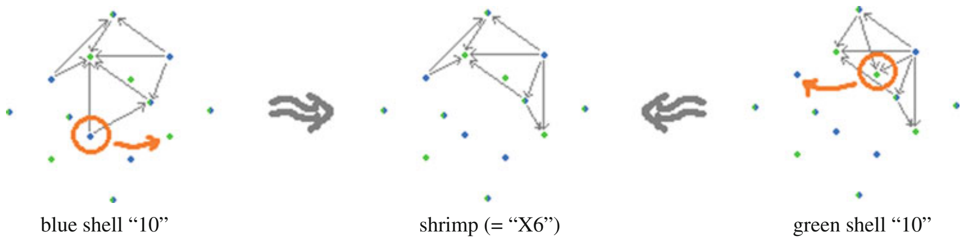


Fig. 140 "Shrimp" arrow-hexagons can be obtained by a 1-move transformation of shell arrow-hexagons

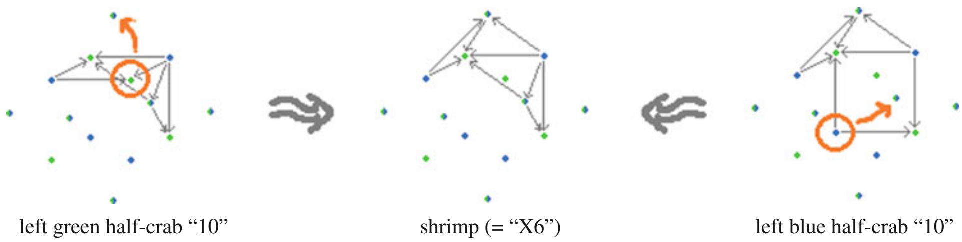


Fig. 141 "Shrimp" arrow-hexagons can be obtained by a 1-move transformation of half-crab arrow-hexagons

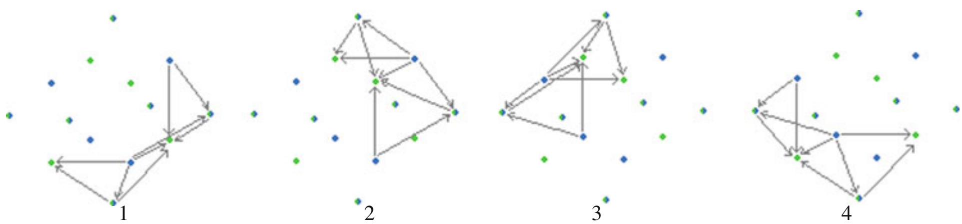


Fig. 142 The first quartet of "left shrimps" (1–4)

The first quartet of left shrimps can be obtained either from the blue shells 1, 2, 3, 4, by a 1-move of, respectively, their vertices b4, b1, b2, b3; or from the green shells 1, 2, 3, 4, by a 1-move of, respectively, their vertices g4, g1, g2, g3. Alternatively, these left shrimps 1-4 can be derived either from the left green half-crabs 1, 2, 3, 4, by a 1-move of, respectively, their vertices g4, g1, g2, g3; or from the left blue half-crabs 1, 2, 3, 4, by a 1-move of, respectively, their vertices b4, b1, b2, b3 (Fig. 142).

Similarly, the second quartet of left shrimps can be obtained either from the blue shells 5, 6, 7, 8, by a 1-move of, respectively, their vertices b3, b1, b4, b2; or from the green shells 5, 6, 7, 8, by a 1-move of, respectively, their vertices g3, g1, g4, g2. Alternatively, these left shrimps 5–8 can be derived either from the left green half-crabs 5, 6, 7, 8, by a 1-move of, respectively, their vertices g3, g1, g4, g2; or from the left blue half-crabs 5, 6, 7, 8, by a 1-move of, respectively, their vertices b3, b1, b4, b2 (Fig. 143).

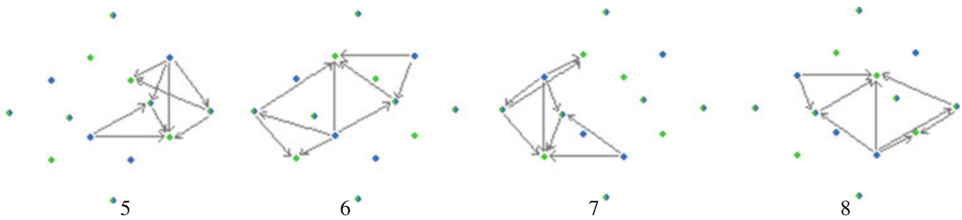


Fig. 143 The second quartet of “left shrimps” (5–8)

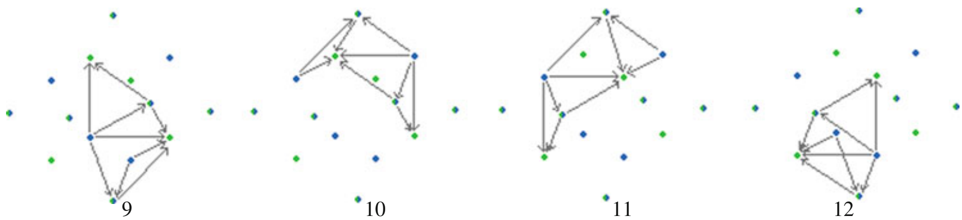


Fig. 144 The third quartet of “left shrimps” (9–12)

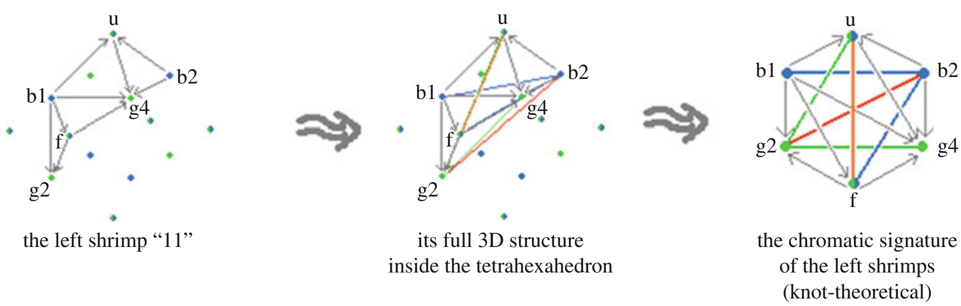


Fig. 145 The chromatic signature of the “left shrimp” arrow-hexagons

In a similar way, the third quartet of left shrimps can be obtained either from the blue shells 9, 10, 11, 12, by a 1-move of, respectively, the vertices b2, b4, b3, b1; or from the green shells 9, 10, 11, 12, by a 1-move of, respectively, their vertices g2, g4, g3, g1. Alternatively, these left shrimps 9–12 can be derived either from the left green half-crabs 9, 10, 11, 12, by a 1-move of, respectively, their vertices g2, g4, g3, g1; or from the left blue half-crabs 9, 10, 11, 12, by a 1-move of, respectively, their vertices b2, b4, b3, b1 (Fig. 144).

The chromatic signature of the left shrimps is the following (here we consider the left shrimp number “11”) (Fig. 145).

The right shrimps behave in a similar way.

The first quartet of right shrimps can be obtained either from the blue shells 1, 2, 3, 4, by a 1-move of, respectively, their vertices b2, b3, b4, b1; or from the green shells 1,

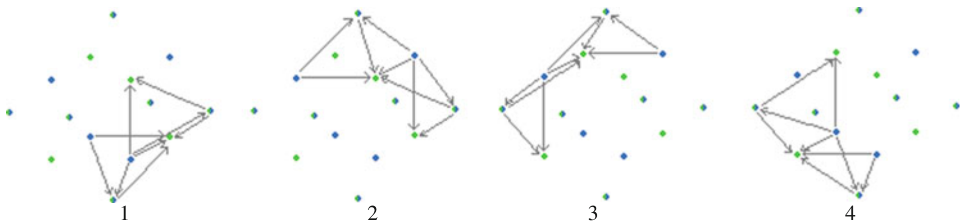


Fig. 146 The first quartet of "right shrimps" (1–4)

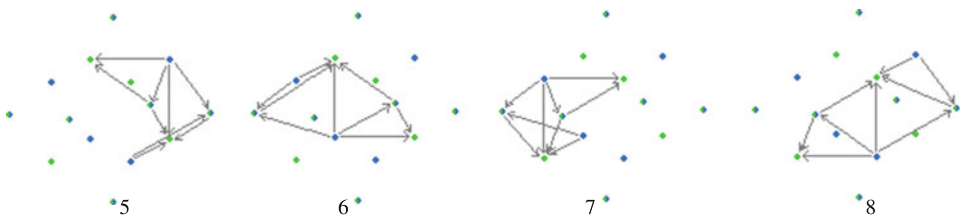


Fig. 147 The second quartet of "right shrimps" (5–8)

2, 3, 4, by a 1-move of, respectively, their vertices g2, g3, g4, g1. Alternatively, the right shrimps 1-4 can be derived either from the right green half-crabs 1, 2, 3, 4, by a 1-move of, respectively, their vertices g2, g3, g4, g1; or from the right blue half-crabs 1, 2, 3, 4, by a 1-move of, respectively, their vertices b2, b3, b4, b1 (Fig. 146).

Analogously, the second quartet of right shrimps can be obtained either from the blue shells 5, 6, 7, 8, by a 1-move of, respectively, their vertices b4, b2, b3, b1; or from the green shells 5, 6, 7, 8, by a 1-move of, respectively, their vertices g4, g2, g3, g1. Alternatively, these right shrimps 5–8 can be derived either from the right green half-crabs 5, 6, 7, 8, by a 1-move of, respectively, their vertices g4, g2, g3, g1; or from the right blue half-crabs 5, 6, 7, 8, by a 1-move of, respectively, their vertices b4, b2, b3, b1 (Fig. 147).

The third and last quartet of right shrimps can be obtained either from the blue shells 9, 10, 11, 12, by a 1-move of, respectively, their vertices b3, b1, b2, b4; or from the green shells 9, 10, 11, 12, by a 1-move of, respectively, their vertices g3, g1, g2, g4. Alternatively, these right shrimps 9–12 can be derived either from the right green half-crabs 9, 10, 11, 12, by a 1-move of, respectively, their vertices g3, g1, g2, g4; or from the right blue half-crabs 9, 10, 11, 12, by a 1-move of, respectively, their vertices b3, b1, b2, b4 (Fig. 148).

The chromatic signature of the right shrimps is the following one (here we consider the right shrimp number "12") (Fig. 149).

As for the chromatic signatures of the left and the right shrimps, similarly to what we saw with half-crabs (left blue and right blue, or left green and right green), they are the same if expressed through graph-theoretical tools (in which case each of them is oppositionally self dual), whereas they differ if they are expressed (as by us here) through knot-theoretical tools (in which case they are 3D mirror images and oppositionally dual) (Fig. 150).

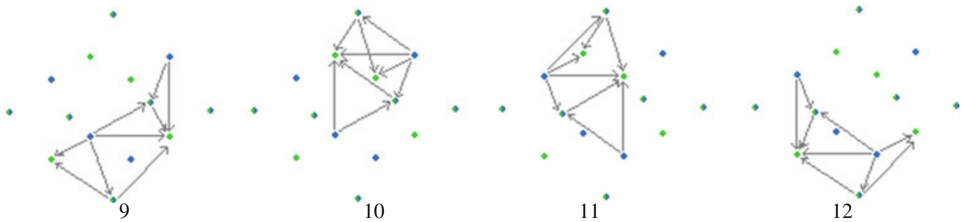
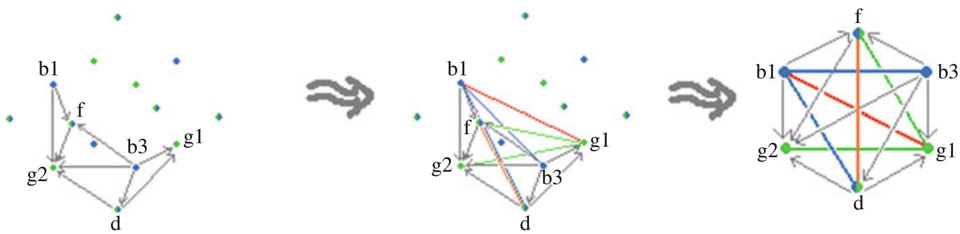


Fig. 148 The third quartet of “right shrimps” (9–12)



the right shrimp “12”

its full 3D structure
inside the tetrahexahedron

the chromatic signature
of the right shrimps
(knot-theoretical)

Fig. 149 The chromatic signature of the “right shrimp” arrow-hexagons

the chromatic signature of the left shrimps
(knot-theoretical)

the chromatic signature of the right shrimps
(knot-theoretical)

Fig. 150 3-D mirror images from the viewpoint of the chromatic signatures of the left shrimps and the right shrimps (moreover they are oppositionally dual)

This reflects the fact that in the 3-D space of the tetrahexahedron left and right shrimps are the levorotatory and dextrorotatory versions of the same structure (to speak as Kant, they are “incongruent counterparts”).

As in the case of the half-crabs, the tortoises, the flying squirrels and the shells (cf. Sect. 3.5, 3.6, 3.7 and 3.8), the shrimps can also pave the surface of the oppositional tetrahexahedron. However, shrimps can pave in a perfect way, without holes (shrimps do not have prisoners). As it happens, the shrimps can pave the tetrahexahedron in several different ways. Given that two shrimps, if joined suitably, can together cover the same surface as a tortoise (but with no hole), shrimps can also pave the tetrahexahedron as the

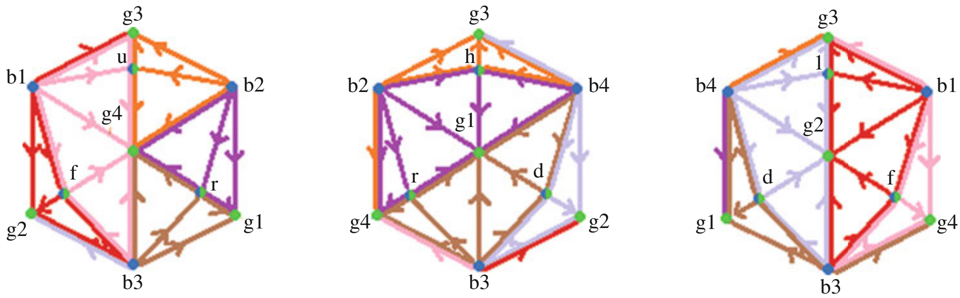


Fig. 151 The shrimp hexagons can pave the surface of the oppositional tetrahexahedron in several ways

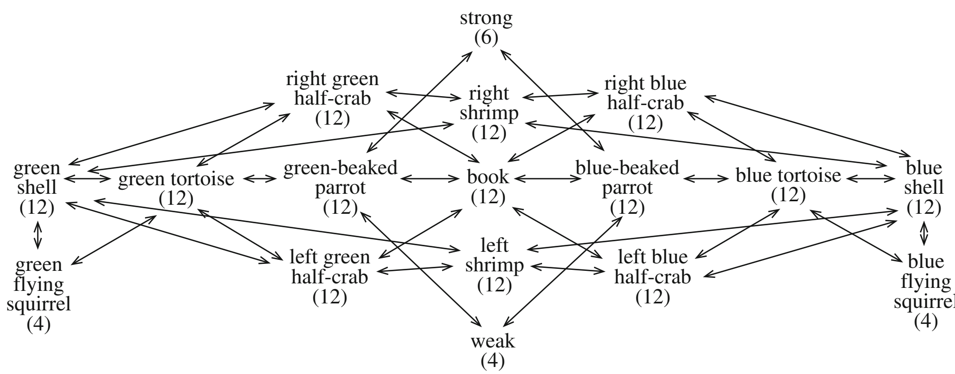


Fig. 152 The morphogenetic chart of the arrow-hexagons of the oppositional tetrahexahedron so far known

tortoises do (cf. Sect. 3.6). Here we give a more original pavement, made of six shrimps (depicted in pink, orange, red, violet, brown and light blue) (Fig. 151).

We can refine our previous oppositional morphogenetic chart of arrow-hexagons by adding left and right shrimps to it (Fig. 152).

If we try to locate the shrimps' chromatic signature with respect to the linear order of oppositional degeneration, the most appropriate place to put them, although this remains debatable (a fuzzy order relation would probably be more appropriate), seems to be in between the half-crabs and the flying squirrels (Fig. 153).

If we accept this linear ordering, then we can draw the following oppositional "degeneration chart" for the nine families of arrow-hexagons known so far (Fig. 154).

Are more arrow-hexagons reachable from the shrimps?

Let us first examine the possible 1-moves of a shrimp's "blue head" (the "green tail" will give symmetric, i.e. dual, results). Here we consider the right shrimp number "12" (Fig. 155).

As it turns out, only one 1-move transformation is possible in this case, namely one which gives a (here: green) shell.

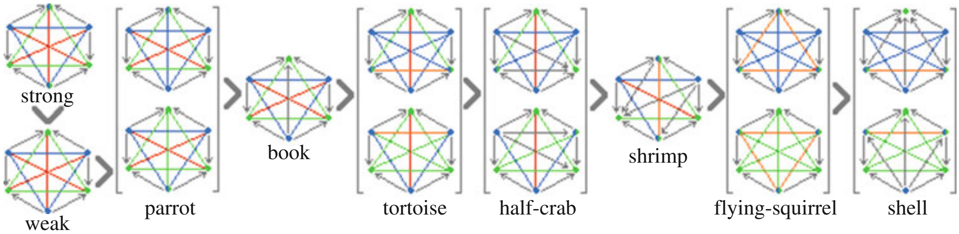


Fig. 153 Arrow-hexagons, so far, can be put in a linear order of growing "oppositional degeneration"

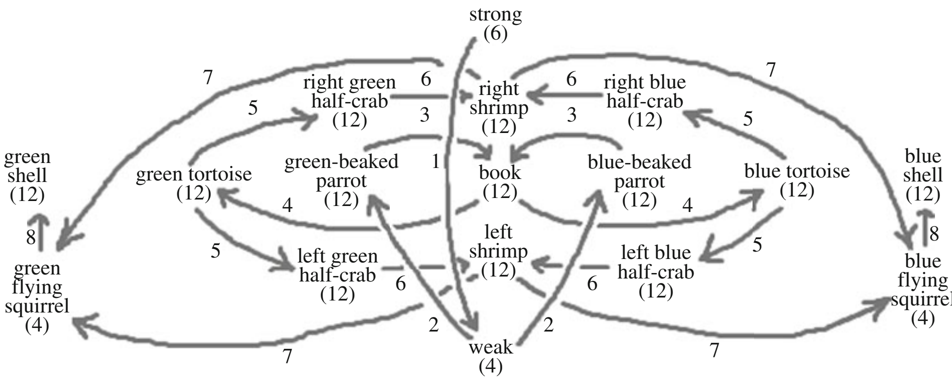


Fig. 154 The morphogenetic "degeneration chart" of the arrow-hexagons of the oppositional tetrahexahedron known so far

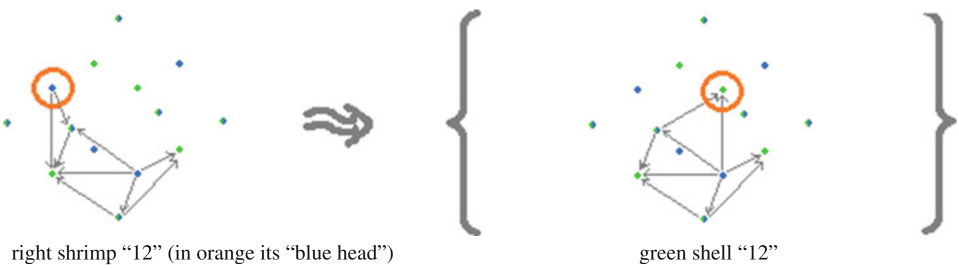


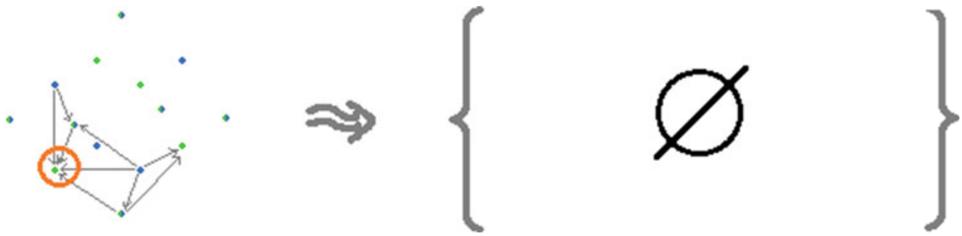
Fig. 155 The only possible transformation of a shrimp through a 1-move of its "blue head" leads to a shell

Secondly, let us examine the transformations obtained from the possible 1-moves of a shrimp's "left green eye" (Fig. 156).

As it happens, no viable transformation (into an arrow-hexagon) is available here.

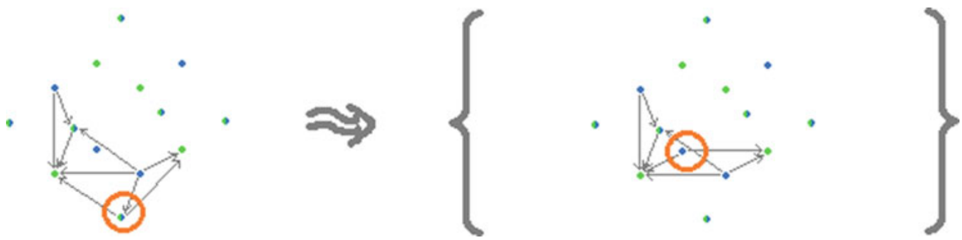
Thirdly, let us consider the possible 1-move transformations of a shrimp's "left blue–green leg" (Fig. 157).

As it turns out, this gives a right blue half-crab.



right shrimp “12” (in orange its “green eye”) no arrow-hexagon

Fig. 156 A shrimp admits neither a 1-move nor a 2-moves transformation into another arrow-hexagon through its “green eye”



right shrimp “12” (in orange its “blue-green eye”) right blue half-crab “12”

Fig. 157 The only possible 1-move transformation of a shrimp through its “blue–green eye”

So, seemingly no new arrow-hexagons are reachable from the shrimps. For the first time in our transformation game we have no more unknown arrow-hexagons to explore. Are we finished? What could this mean?

4 Some Consequences

Let us now try to draw some consequences from our study. We started by asking whether arrow-hexagons, in their then mysterious variety, have interesting invariant properties. To start with, is their variety less mysterious by now? What is this variety?

4.1 Have We Reached a Closure of the Arrow-Hexagons with These 9 Families?

It would seem that the nine families of arrow-hexagons presented here constitute a closed structure with respect to our transformation game. Any of the nine can be reached, in a sufficient finite number of steps, from any other one, and no other arrow-hexagon can be obtained, step by step, from any of the nine families of arrow-hexagons presented here.

However, can we be sure that no other arrow-hexagons exist (inside the β3-structure)? For it could be the case that not all the arrow-hexagons of the β3-structure are morphogenetically *connected* and, therefore, it could be the case that some further arrow-hexagon exists, even though our game will never be able to reach it by connexity. As it happens, one can prove that the nine arrow-hexagons we obtained are the only existing ones. We will not give the proof here, which simply consists in considering, by graphical construction of the β3-structure, either one of the 12 arrows of the cube or one of the 24 arrows of the spike, these being the only two possible kinds of arrows of a tetrahexahedron, and in showing that whatever "hexagon" of alternated arrows (in its perimeter) one constructs inside the β3-structure upon this starting arrow the result is one of the nine kinds of arrow-hexagons that we presented here.

4.2 In the Tetrahexahedron There Are Also "Folded Arrow-Hexagons"

As it happens, more kinds of arrow-hexagons can indeed be found inside the oppositional tetrahexahedron, provided, however, that one allows the appearance (in the "inside" of the arrow-hexagon, i.e. inside its chromatic signature) of "bi-directional arrows" (i.e. the logical biconditional connective). This means considering arrow-hexagons such that they have strictly less than six vertices (i.e. such that some of their vertices play more than one time, and therefore are joined by a double arrow, being equivalent). It must be remarked that in some sense the study of such structures would commit us to adopting at least *de jure* some of the tools of Smessaert and Demey's "logical geometry" (cf. Sect. 2 above, Fig. 8). Their "implication geometry" deals, among other things, with double-sided arrows of logical equivalence (unless what is needed here is rather "equality" than logical equivalence). In any case, the idea of "counting a vertex two times" would contradict the so far standard methodology of oppositional geometry: which therefore should/could be enriched duly. Just to give an intuition, here is a small visual list of such possible "folded arrow-hexagons" (butterflies, tadpoles, snakes, bacteria. . .) (Fig. 158).

Their chromatic signature is fully comparable to those we have seen in our study, with the additional, crucial feature, however, that now such a chromatic signature also contains double-sided arrows (of logical equivalence) (Fig. 159).

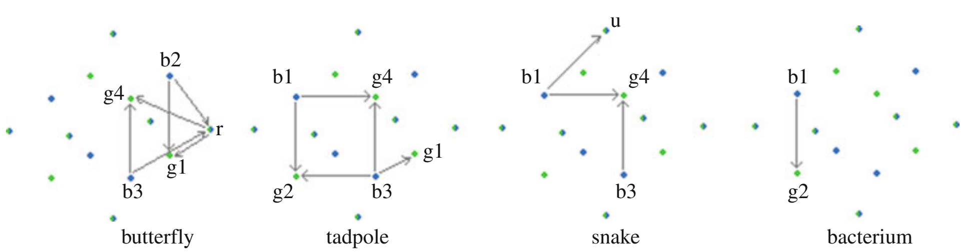


Fig. 158 Some "folded arrow-hexagons", i.e. arrow-hexagons with strictly less than six distinct vertices

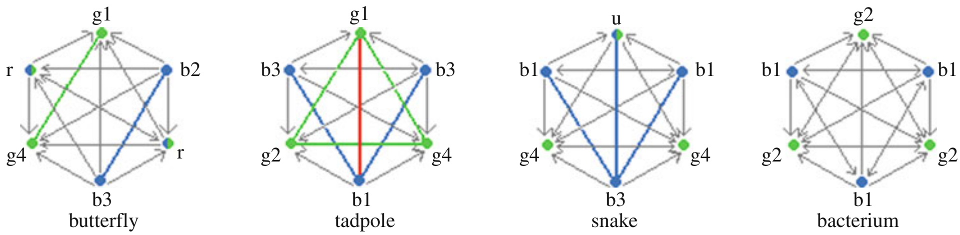


Fig. 159 The chromatic signature of some "folded arrow-hexagons", i.e. arrow-hexagons with double-sided arrows

Note also that these "folded arrow-hexagons", which contain at least one double-sided arrow, are not counted among the 30 030 "hexagons" given by the formula C_{14}^{6}. So, including them in our catalogue gives an even larger number of arrow-hexagons (all contained in the tetrahexahedron)!

Note also that there could be other interesting ways of conceiving morphogenetic degeneration charts, for instance by using the dimension "height" of a 3-D representation for measuring degeneration (thus using "pits" and the like). It also seems likely that our morphogenetic charts are not yet satisfactory either. For instance, they do not yet express well the "morphogenetic distance" between two kinds of arrow-hexagons, for example, the distance – of four moves – between the strong and the shrimps seems to be expressed suboptimally (in our chart of Fig. 152 the two seem to be close). It might be the case that the relevant ordering of the nine families of arrow-hexagons must be a four-dimensional space.

4.3 The Place of the Arrow-Hexagons Among All Hybrid Hexagons

As it happens, the construction which gave, by "unfolding" an arrow of subalternation, an arrow-hexagon "bacterion" (cf. Figs. 158 and 159), can give, when applied to the other four possible kinds of segments of the tetrahexahedron (i.e. the red one of contradiction, the blue one of contrariety, the green one of subcontrariety, the orange one of independence), comparable strange "folded hybrid hexagons". These keep the same "bacterial structure" of two entangled triangles (each one made of a closed chain of three double arrows) which characterised the bacterion but have a different perimeter and different diagonals according to the oppositional quality of the starting segment. Let us call these folded hybrid hexagons "segment-hexagons" (or "bacterial hexagons") (Fig. 160).

This leads us to an important general point. Outside the strong constraint of keeping the perimeter made of alternated arrows intact (suggested *in primis* by the logical hexagon and its intriguing mathematical power) there is, truly speaking, a huge quantity of other possible oppositional "hexagons", regular or not. Most of them are highly irregular, but even in that case they exist by pairs of "oppositionally dual" hexagons, and as any other hybrid hexagon, these can be represented by their chromatic signature (Fig. 161).

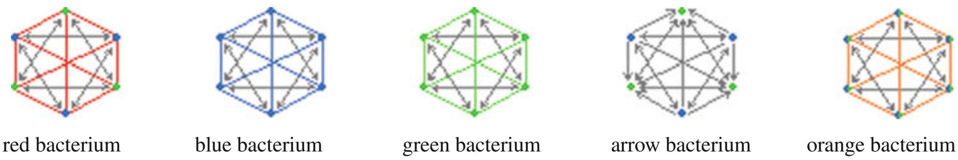
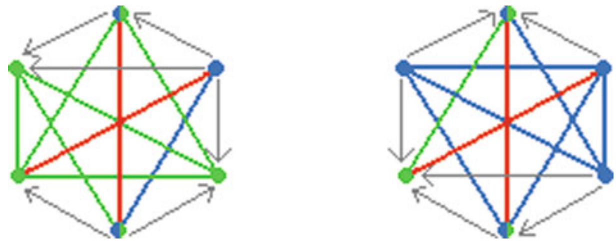


Fig. 160 Another interesting family of regular folded hybrid hexagons: the "bacterial hexagons"

Fig. 161 Some hybrid hexagons, regular or not, contained in the $\beta3$-structure that are not arrow-hexagons



There seems to be reason to believe that several of these formal entities possess still unknown properties that are quite interesting and might be very useful somewhere in the future.

In particular (cf. [24]), some hybrid hexagons, containing orange segments (of independence, cf. Sect. 3.6, Fig. 94), are already known to be quite useful. They are the attractors (i.e. the expected outcome) of the "partial superposition" of oppositional segments" (Fig. 162).

Such superpositions admit some families of invariant solutions, among them the following three orange hybrid hexagons (we give their chromatic structure in Fig. 163)

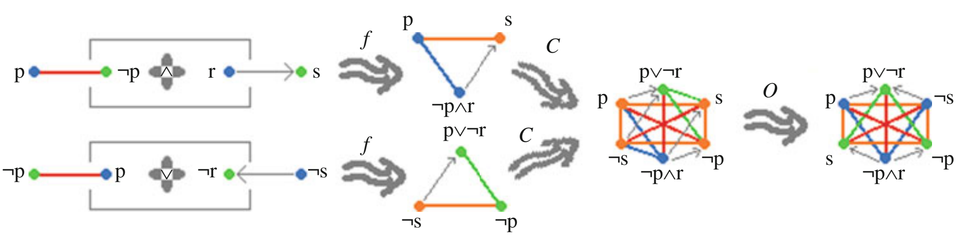


Fig. 162 The partial ∧-superposition of red and from-arrow and the ∨-superposition of red and to-arrow are oppositionally dual

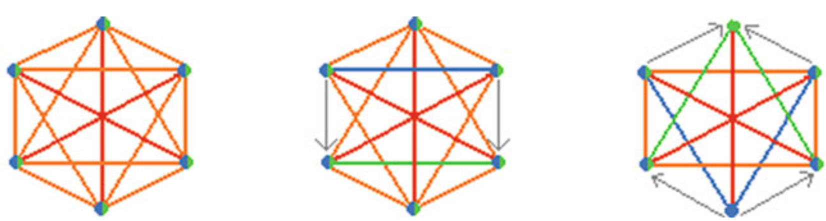


Fig. 163 Some hybrid hexagons, regular or not, contained in the $\beta3$-structure that are not arrow-hexagons

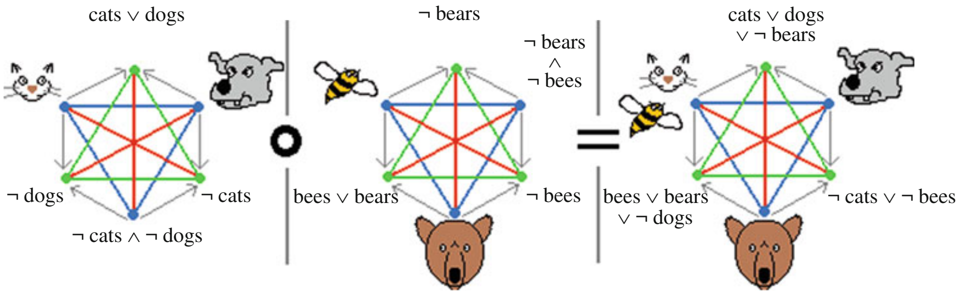


Fig. 164 The (non-uniform) superposition of two logical hexagons can lead to a suboptimal result (a strong or weak hexagon)

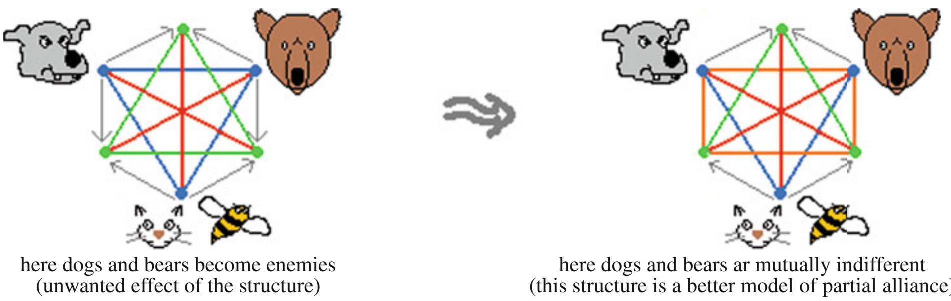


Fig. 165 Resorting to suitable hybrid hexagons can lead to better formal tools for expressing "oppositional dynamics"

Moreover, when two logical hexagons are combined by "oppositional superposition" (which, in this case, is a technique for combining three-oppositional structures to create models of enmities and alliances) one of these three orange hybrid hexagons expresses the kind of result that we would like to obtain as an outcome of our model. That is, we do not want, because cats and bees have made an alliance while their respective enemies (dogs and bears) do not have, that dogs and bears nevertheless become *ipso facto* mutual enemies: which would be yet the case if the hexagon obtained by superposition were a strong or even a weak one (in that case dogs and bears would be joined, inside such an hexagon, by a blue segment of contrariety) (Fig. 164).

Resorting to one of the orange hybrid hexagons of Figs. 162 and 163 offers the means for a much better solution to the problem of expressing oppositional partial alliance (cf. Fig. 165).

So, provided that we think that there can be quite interesting properties behind different kinds of hybrid oppositional structures, it seems natural to think that any tool allowing "navigation" inside the tetrahexahedron would be very welcome. This chapter tried to produce one. To give an idea of possible future formal combinatorial games here is a simple, very small attempted typology (Fig. 166).

Some research is being done in this direction. Smessaert and Demey, for instance, are investigating, beyond hexagons (but inside the $\beta 3$-structure), the formal properties of op-

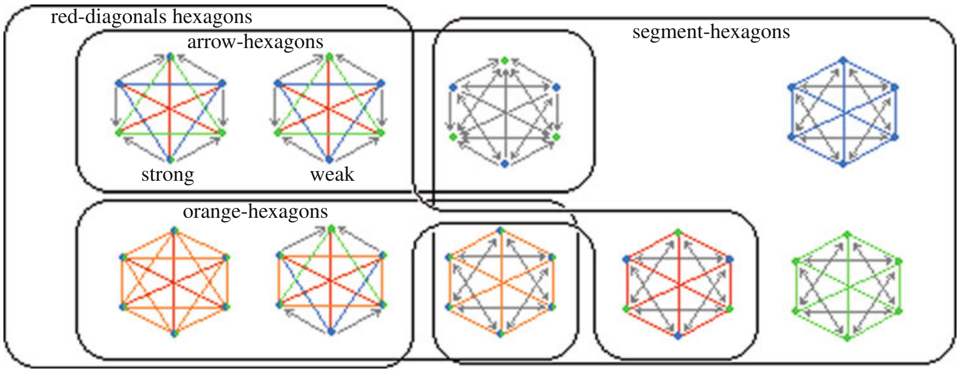


Fig. 166 Some possible, very general families of hybrid hexagons (arrow-hexagons are a very particular case)

positional structures that keep red diagonals (cf. [33]). However, a general "cartography" of the hybrid hexagons of the $\beta 3$-structure seems to be something that is still missing, and the more general task of studying hybrid oppositional structures beyond hexagons is something even further away.

5 Béziau's "Red Hexagon" Does Not Belong to the Tetrahexahedron!

Yet what about Béziau's provocative and stimulating "red arrow-hexagon" (Fig. 9) that we mentioned at the beginning of this study? Doesn't it contrast, by its simple existence, our claim according to which there are only the nine families of arrow-hexagons listed in Sects. 3.1–3.9 above? For it does not belong to any of them. As it happens, the "red hexagon" proposed by Béziau cannot be exactly as it was first proposed. It is not possible to have a red "triangle of contradiction" (if p is contradictory to q and q is contradictory to r, r cannot be contradictory to p; rather, r is then equivalent to p). So, instead, Béziau's very interesting red hexagon, which is indeed some kind of very singular oppositional hexagon, turns out to be equivalent to the following *non-arrow-hexagon*, which does not have an intact arrow-perimeter and is not uniquely constituted (in its inside) of red segments of contradiction (Fig. 167).

This hexagon, which is totally new in the field, is clearly not an arrow-hexagon. But even after correction, Béziau's new hexagon remains weird, for in some sense it does not belong to the $\beta 3$-structure, which is very strange. Béziau, who contributes so importantly in various ways to the field, so to say happens now to be a "hacker" of oppositional geometry. He finds the way of constructing something that breaks the framework! The explanation for this puzzle is the following: "p or not p" (used by Béziau's red hexagon) is equivalent to "$\top$", whereas "p and not p" (also used by Béziau for building his red hexagon) is equivalent to "$\bot$". However, oppositional geometry has a very special treat-

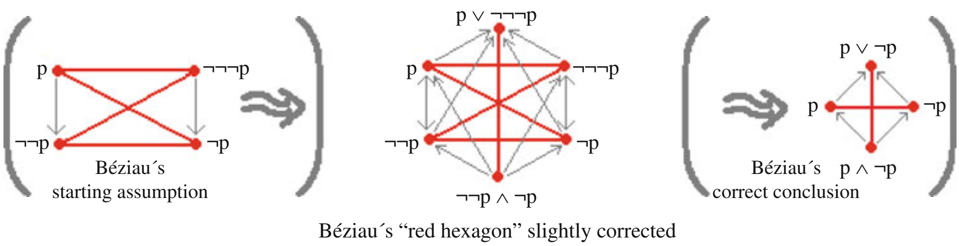


Fig. 167 Béziau's "red arrow-hexagon" is in fact a hybrid hexagon, but not an arrow-hexagon; its perimeter is broken

ment for ⊤ and ⊥: on the one hand, they cannot be denoted explicitly by the vertices (this is a technical point and a very important "trademark" of the theory: explicit (i.e. simplicial) contrariety cannot be trivial). On the other hand, the two trivial logical constants are nevertheless "expressed geometrically" by the "symmetry centre" of *any* oppositional figure, this visible centre (which, again, is not a vertex!) embodying at the same time "⊤" and "⊥" (this fact has been established independently by Angot-Pellissier and Smessaert). Therefore, if his strange red hexagon, which exists and does not exist, turned out to be a solid paraconsistent singularity of oppositional geometry, Béziau might be opening a new chapter of the theory . . .

6 From the Arrow-Hexagons to the "Arrow-βn-Structures"

A final natural question is that of knowing whether the structures highlighted in this study (i.e. the arrow hexagons and, more generally, the hybrid hexagons) admit an equivalent at higher levels of oppositional geometry. The present study has, for instance, focussed on the "avatars" (the "oppositional shadows") of the $\beta 2$-structure inside the $\beta 3$-structure that replaces it when the oppositional space is "inflated" (as we have seen, this is the role played by the blue tortoises, as we recalled with the example of the hexagon of total order

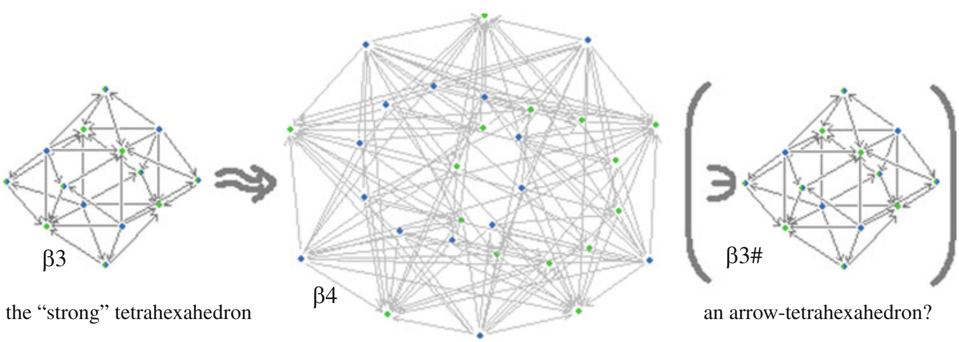


Fig. 168 Are there hybrid tetrahexahedra and, in particular, arrow-tetrahexahedra inside the $\beta 4$-structure?

that becomes a tetrahexahedron of partial order, Fig. 17). The question then is: is something comparable to this to be expected when a $\beta 3$-structure becomes a $\beta 4$-structure? This will be the possible subject of a future study: studying hybrid instances of the oppositional tetrahexahedron inside the $\beta 4$-structure, such that the tetrahexahedron's 3-D network of 36 alternated arrows remains constant while its inner relations (of contradiction, contrariety, subcontrariety and independence) change (Fig. 168).

7 Conclusion

In this paper we proposed giving a unified treatment to a mysterious phenomenon encountered independently by different scholars, including Béziau (cf. Sect. 2). As a methodology for this we proposed focussing on the notion of the "arrow-hexagon", which is a generalisation of the notion of the logical hexagon. In order to develop this, we proposed some preliminary notions and strategies through a geometric game, such as "hybrid hexagons", "chromatic signature", "oppositional duality", "oppositional morphogenetics", "oppositional shadow", "oppositional degeneration" and "oppositional role". We were able to determine that there are nine families of arrow-hexagons, which are mutually related and bear some roles, and we started to describe them. Among them, four already have an identified important role, but five remain poorly understood so far. If several oppositional roles seem to be missing so far for the arrow-hexagons, some important oppositional roles are already available for hybrid hexagons that are not arrow-hexagons. This seems promising. Oppositional roles do exist and deserve to be studied (and used in models). Moreover, this suggests that this kind of enquiry should be generalized, by studying new families of oppositional fragments of the tetrahexahedron and, more generally, the fragments of any βn-structure. Our aim in the long run consists in establishing formal tools for modelling "*dynamic* oppositional phenomena". This aim is followed both by studying the formal properties of the hybrid oppositional structures (and beyond) and by trying to give to them, whenever possible, an intuitive pragmatic interpretation.

References

1. Angot-Pellissier, R.: 2-opposition and the topological hexagon. In: Béziau, J.-Y., Payette, G. (eds.) The Square of Opposition – a General Framework for Cognition. Peter Lang, Bern (2012)
2. Angot-Pellissier, R.: Many-valued logical hexagons in a 3-oppositional trisimplex (draft)
3. Angot-Pellissier, R., Moretti, A.: Many-valuedness, paracompleteness and paraconsistency in a 3-oppositional quadri-simplex of sheaves. Talk at the 5th World Congress on Paraconsistency, Kolkata, India, 13–17 February 2014
4. Ben Aziza, H. (eds.): Le carré et ses extensions: aspects théoriques, pratiques et historiques. Publications de la faculté des sciences humaines et sociales de Tunis, Tunis (2015)
5. Béziau, J.-Y.: New light on the square of oppositions and its nameless corner. Log. Investigat. **10**, 218–233 (2003)
6. Béziau, J.-Y.: The power of the Hexagon. Log. Univers. **6**(1–2), 1–43 (2012)

7. Béziau, J.-Y., Costa-Leite, A., Facchini, A. (eds.): Aspects of Universal Logic. Travaux de logique, vol. 17. University of Neuchâtel, Neuchâtel (2004)
8. Béziau, J.-Y., Jacquette, D. (eds.): Around and Beyond the Square of Opposition. Birkhäuser, Basel (2012)
9. Béziau, J.-Y., Payette, G. (eds.): The Square of Opposition – A General Framework for Cognition. Peter Lang, Bern (2012)
10. Blanché, R.: Sur l'opposition des concepts. Theoria **19**, 89–130 (1953)
11. Blanché, R.: Structures intellectuelles. Essai sur l'organisation systématique des concepts. Vrin, Paris (1966)
12. Blanché, R.: Raison et discours. Défense de la logique discursive. Vrin, Paris (1967)
13. Chatti, S., Schang, F.: The Cube, the Square and the Problem of Existential Import. Hist. Philos. Log. **34**(2), 101–132 (2013)
14. Dubois, D., Prade, H.: From Blanché's hexagonal organization of concepts to formal concept analysis and possibility theory. Log. Univers. **6**(1-2), 149–169 (2012)
15. Jacoby, P.: A triangle of opposites for types of propositions in Aristotelian logic. The New Scholasticism **24**, 32–56 (1950)
16. Moretti, A.: Politics from the viewpoint of oppositional geometry (draft)
17. Moretti, A.: Geometry for modalities? Yes: through n-opposition theory. In: Béziau, J.-Y., Costa-Leite, A., Facchini, A. (eds.) Aspects of Universal Logic. Travaux de logique, vol. 17. University of Neuchâtel, Neuchâtel (2004)
18. Moretti, A.: The geometry of logical opposition. PhD Thesis, University of Neuchâtel, Neuchâtel, Switzerland (2009)
19. Moretti, A.: The geometry of standard deontic logic. Log. Univers. **3**(1), 19–57 (2009)
20. Moretti, A.: The critics of paraconsistency and of many-valuedness and the geometry of oppositions. Log. Log. Philos. **19**, 63–94 (2010)
21. Moretti, A.: Why the logical hexagon? Log. Univers. **6**(1–2), 69–107 (2012)
22. Moretti, A.: Was Lewis Carroll an amazing oppositional geometer? Hist. Philos. Log. **35**, 4 (2014)
23. Moretti, A.: Le retour du refoulé: l'hexagone logique qui est derrière le carré sémiotique. In: Ben Aziza, H. (eds.) Le carré et ses extensions: aspects théoriques, pratiques et historiques. Publications de la faculté des sciences humaines et sociales de Tunis, Tunis (2015)
24. Moretti, A.: The partial superposition of two 2-oppositional segments (draft)
25. Pellissier, R.: "Setting" n-opposition. Log. Univers. **2**(2) 235–263 (2008)
26. Sauriol, P.: Remarques sur la théorie de l'hexagone logique de Blanché. Dialogue **7**, 374–390 (1968)
27. Sauriol, P.: La structure tétrahexaédrique du système complet des propositions catégoriques. Dialogue **15**, 479–501 (1976)
28. Sesmat, A.: Logique – II. Les raisonnements, la logistique. Hermann, Paris (1951)
29. Smessaert, H.: On the 3D visualisation of logical relations. Log. Univers. **3**(2), 303–332 (2009)
30. Smessaert H. The classical Aristotelian hexagon versus the modern duality hexagon. Log. Univers. **6**(1–2), 171–199 (2012)
31. 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 Artificial Intelligence, vol. 73 520. Springer, Heidelberg (2012)
32. Smessaert, H., Demey, L.: Logical geometries and information in the square of oppositions. J. Log. Lang. Inf. **23**(4), 527–565 (2014)
33. Smessaert, H., Demey, L.: La logique géométrique du dodécaèdre rhombique des oppositions. In: Ben Aziza, H. (eds.) Le carré et ses extensions: aspects théoriques, pratiques et historiques. Publications de la faculté des sciences humaines et sociales de Tunis, Tunis (2015)

A. Moretti (✉)
Nice, France
e-mail: alessio.moretti@unine.ch

The Distributed Ontology, Modeling, and Specification Language – DOL

Till Mossakowski, Mihai Codescu, Fabian Neuhaus, and Oliver Kutz

Abstract There is a diversity of ontology languages in use, among them OWL, RDF, OBO, Common Logic, and F-logic. Related languages such as UML class diagrams, entity-relationship diagrams and object role modeling provide bridges from ontology modeling to applications, e.g., in software engineering and databases. Also in model-driven engineering, there is a diversity of diagrams: UML consists of 15 different diagram types, and SysML provides further types. Finally, in software and hardware specification, a variety of formalisms are in use, like Z, VDM, first-order logic, temporal logic etc.

Another diversity appears at the level of ontology, model and specification modularity and relations among ontologies, specifications, and models. There is ontology matching and alignment, module extraction, interpolation, ontologies linked by bridges, interpretation and refinement, and combination of ontologies, models and specifications.

The *d*istributed *o*ntology, modeling and specification *l*anguage (DOL) aims at providing a unified metalanguage for handling this diversity. In particular, DOL provides constructs for (1) "as-is" use of ontologies, models, and specifications (OMS) formulated in a specific ontology, modeling or specification language, (2) OMS formalized in heterogeneous logics, (3) modular OMS, (4) mappings between OMS, and (5) networks of OMS. This chapter sketches the design of the DOL language. DOL has been submitted as a proposal within the OntoIOp (ontology, model, specification integration and interoperability) standardisation activity of the object management Group (OMG).

Keywords Heterogeneous ontologies · Modularity · Interoperability · Institutions

Mathematics Subject Classification (2000) Primary 68T30 · Secondary 03C95

1 Introduction

Logical languages are used in several fields of computing for the development of formal, machine-processable texts that carry a formal semantics. Among those fields are (1) **O**ntologies formalizing domain knowledge, (2) (formal) **M**odels of systems, and (3) the formal **S**pecification of systems. Ontologies, models and specifications will (for the purpose of this paper) henceforth be abbreviated as **OMS**.

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_21

An OMS provides formal descriptions which range in scope from domain knowledge and activities (ontologies, models) to properties and behaviours of hardware and software systems (models, specifications). While the use of OMS varies considerably, there are two recurring challenges: *reusability* and *interoperability*.

Reusability is an issue because the development of OMS is typically done manually by experts and, thus, an expensive process. Hence, it is desirable to be able to reuse existing OMS during the development of new OMS. This presupposes a framework that allows to build *structured OMS* by identifying modules and their relationships to each other. For example, it requires the ability to combine two existing OMS in a way that handles the namespaces of the OMS in an appropriate way. Further, the reuse of an existing OMS often requires that the OMS is *adapted* for its new purpose. For example, the adaption may require the extension of the OMS by new axioms, or the extraction of a subset of the OMS, or the change of its semantics from open world to closed world.

The interoperability challenge is closely related to the reusability challenge. Since the development of OMS is not an exact science and is usually driven by project specific requirements, two OMS that have been developed independently will represent the same domain in different and, often, conflicting ways. They may differ, for example, with respect to the terminology, or with respect to the definitions of the underlying concepts, or with respect to the perspective from which they represent their domain. Thus, in a situation where two independently developed OMS are supposed to be reused as modules of a larger OMS, the differences between these OMS will typically prevent them from working together properly. Overcoming this lack of interoperability may require an alignment or even an integration of these OMS. This typically involves the identification of synonyms, homonyms, and the development of bridge axioms, which connect the two OMS appropriately.

Both the reusability and the interoperability challenges are amplified by the diversity of OMS languages that are in use. For ontologies these include OWL, RDF, OBO, Common Logic, and F-logic. Related languages such as UML class diagrams, entity-relationship diagrams and object role modeling provide bridges from ontology modeling to applications, e.g., in software engineering and databases. Also in model-driven engineering, there is a diversity of diagrams: UML consists of 15 different diagram types, and SysML provides further types. Finally, in software and hardware specification, a variety of formalisms are in use, like Z, VDM, first-order logic, temporal logic etc. These languages do not just differ with respect to their syntax, but with respect to their semantics and to their levels of expressiveness.

To address both challenges we propose the *d*istributed *o*ntology, modeling and specification *L*anguage (DOL). DOL is a metalanguage that enables the reuse, integration, and alignment of existing OMS – even if they are written in different formalisms. The underlying methodological stance is that it would be futile to attempt to develop yet another OMS language that would subsume all the others; instead we have to accept the diversity of OMS languages and the diversity of perspectives that are represented by different OMS. DOL provides a sound and formal semantic basis for specifying structured OMS, which may reuse as modules several existing OMS (possibly written in different languages) without requiring any changes to these modules. Further, DOL allows us to specify mappings between different OMS (e.g., alignments and logical entailments).

In particular, DOL enjoys the following distinctive features:

- modular OMS and OMS networks are specially supported,
- OMS can not only be aligned (as in BioPortal [40] and NeON [17]), but also combined along alignments,
- mappings between OMS (interpretation of theories, conservative extensions etc.) are supported;
- it supports a variety of OMS languages (OWL, RDF, Common Logic, first-order logic, CASL; planned: UML, relational database schema, F-logic, distributed description logics, and more);
- OMS can be translated to other OMS languages, and compared with OMS in other languages;
- heterogeneous OMS (i.e., structured OMS with modules written in different languages) can be built;
- OMS languages and OMS language translations are first-class citizens and are available on the Web as linked data.

The paper is organized as follows: we first discuss the theoretical foundations of DOL in Sect. 2, followed by a sketch of the DOL language itself in Sect. 3. Section 4 briefly discusses the DOL-enabled, web-based OMS repository engine Ontohub, and Sect. 5 concludes.

2 Foundations of the Distributed Ontology, Modeling and Specification Language (DOL)

The DOL[1] aims at providing a unified framework for (1) "as-is" use of OMS formulated in a specific OMS language, (2) modular OMS, (3) mappings between OMS, (4) OMS networks, and (5) OMS formalized in heterogeneous logics. Historically, the design of DOL has inherited many ideas and features (1) discussed in the Workshop on Modular Ontologies series [15, 16, 24, 28, 43, 45], (2) from the Alignment API [10], (3) from CLEAR, ASL and specifications in an arbitrary institution [5, 41, 42, 47], and (4) from the CASL (common algebraic specification language) and HetCASL (CASL's heterogeneous extension) languages, standardized in IFIP WG 1.3[2] (Foundations of System Specification) [2, 25, 30, 35].

A library in DOL consists of modules formalized in *basic OMS languages*, such as OWL (based on description logic) or Common Logic (based on first-order logic with some second-order features). These modules are serialized in the existing syntaxes of these languages in order to facilitate reuse of existing OMS. DOL adds a meta-level on top, which

[1] DOL has formerly been standardized within ISO/TC 37/SC 3. The OntoIOp (ontology, modeling and specification integration and interoperability) activity is now being continued at OMG, see the project page at http://ontoiop.org.
[2] See http://ifipwg13.informatik.uni-bremen.de

allows for expressing heterogeneous OMS and mappings between OMS.[3] Such mappings include (heterogeneous) *imports* and *alignments*, *conservative extensions* (important for studying OMS modules), and *theory interpretations* (important for reusing proofs). Thus, DOL gives OMS interoperability a formal grounding and makes heterogeneous OMS and services based on them amenable to automated verification. The basic syntax and semantics of DOL has been introduced in [37, 38], and the general theory of heterogeneous specifications for OMS in [27]. DOL uses internationalized resource identifiers (IRIs, the Unicode-aware superset of URIs) for all entities of OMS libraries to make them referenceable on the Web.

2.1 Foundations

The large variety of logical languages in use can be captured at an abstract level using the concept of *institutions* [12]. This allows us to develop results independently of the particularities of a logical system and to use the notions of institution and logical language interchangeably throughout the rest of the chapter.

The main idea is to collect the non-logical symbols of the language in signatures and to assign to each signature the set of sentences that can be formed with its symbols. For each signature, we provide means for extracting the symbols it consists of, together with their kind. Signature morphisms are mappings between signatures. We do not assume any details except that signature morphisms can be composed and that there are identity morphisms; this amounts to a category of signatures. Readers unfamiliar with category theory may replace this with a partial order (signature morphisms are then just inclusions). See [37] for details of this simplified foundation.

Institutions also provide a model theory, which introduces semantics for the language and gives a satisfaction relation between the models and the sentences of a signature. The main restriction imposed is the satisfaction condition, which captures the idea that truth is invariant under change of notation (and enlargement of context) along signature morphisms. This relies on two further components of institutions: the translation of sentences along signature morphisms, and the reduction of models against signature morphisms (generalizing the notion of model reduct known from logic).

Definition 2.1 An **institution** [12] is a quadruple $I = (\mathbf{Sign}, \mathbf{Sen}, \mathbf{Mod}, \models)$ consisting of the following:

- a category **Sign** of *signatures* and *signature morphisms*,
- a functor $\mathbf{Sen} : \mathbf{Sign} \to \mathbb{S}et$[4] giving, for each signature Σ, the set of *sentences* $\mathbf{Sen}(\Sigma)$, and for each signature morphism $\sigma : \Sigma \to \Sigma'$, the *sentence translation map* $\mathbf{Sen}(\sigma) : \mathbf{Sen}(\Sigma) \to \mathbf{Sen}(\Sigma')$, where often $\mathbf{Sen}(\sigma)(\varphi)$ is written as $\sigma(\varphi)$,

[3] The languages that we call "basic" OMS languages here are usually limited to one logic and do not provide meta-theoretical constructs.
[4] $\mathbb{S}et$ is the category having all sets as objects and functions as arrows.

- a functor $\mathbf{Mod} : \mathbf{Sign}^{op} \to \mathbb{C}at$[5] giving, for each signature Σ, the category of *models* $\mathbf{Mod}(\Sigma)$, and for each signature morphism $\sigma\colon \Sigma \longrightarrow \Sigma'$, the *reduct functor* $\mathbf{Mod}(\sigma) : \mathbf{Mod}(\Sigma') \to \mathbf{Mod}(\Sigma)$, where often $\mathbf{Mod}(\sigma)(M')$ is written as $M'|_\sigma$, and $M'|_\sigma$ is called the σ*-reduct* of M', while M' is called a σ*-expansion* of $M'|_\sigma$,
- a satisfaction relation $|_\Sigma \subseteq |\mathbf{Mod}(\Sigma)| \times \mathbf{Sen}(\Sigma)$ for each $\Sigma \in |\mathbf{Sign}|$,

such that for each $\sigma\colon \Sigma \longrightarrow \Sigma'$ in **Sign** the following **satisfaction condition** holds:

$$(\star) \qquad M' \models_{\Sigma'} \sigma(\varphi) \text{ iff } M'|_\sigma \models_\Sigma \varphi$$

for each $M' \in |\mathbf{Mod}(\Sigma')|$ and $\varphi \in \mathbf{Sen}(\Sigma)$. □

It is also possible to complement an institution with a proof theory, introducing a derivability relation between sentences, formalized as an *entailment system* [33]. In particular, this can be done for all logics that have so far been in use in DOL.

Example 2.2 OWL signatures consist of sets of atomic classes, individuals, object, and data properties. OWL signature morphisms map classes to classes, individuals to individuals, object properties to object properties, and data properties to data properties. For an OWL signature Σ, sentences are subsumption relations between classes or properties, membership assertions of individuals in classes and pairs of individuals in properties, complex role inclusions, and some more. Sentence translation along a signature morphism simply replaces nonlogical symbols with their image along the morphism. The kinds of symbols are class, individual, object property, and data property, respectively, and the set of symbols of a signature is the union of its sets of classes, individuals, and properties. Models are (unsorted) first-order structures that interpret concepts as unary and properties as binary predicates, and individuals as elements of the universe of the structure, and satisfaction is the standard satisfaction of description logics. This gives us an institution for OWL.

Strictly speaking, this institution captures *OWL 2 DL without restrictions* in the sense of [44]. The reason is that in an institution, the sentences can be used for arbitrary formation of theories. This is related to the presence of DOL's union operator on OMS. OWL 2 DL's specific restrictions on theory formation can be modeled *inside* this institution, as a constraint on OMS. This constraint is generally not preserved under unions or extensions. DOL's multi-logic capability allows the clean distinction between ordinary OWL 2 DL and OWL 2 DL without restrictions.

In this framework, a basic OMS O over an institution I is a pair (Σ, E) where Σ is a signature and E is a set of Σ-sentences. Given a basic OMS O, we denote by $\mathsf{Sig}(O)$ the signature of the OMS. An OMS morphism $\sigma : (\Sigma_1, E_1) \to (\Sigma_2, E_2)$ is a signature morphism $\sigma : \Sigma_1 \to \Sigma_2$ such that $\sigma(E_1)$ is a logical consequence of E_2. Here, in an arbitrary institution, notions like logical consequence, satisfiability etc. can be defined in the standard way.

In the following we will need to assume existence of inclusions between signatures and of their unions. These concepts can be captured in a categorical setting using *inclusion*

[5] $\mathbb{C}at$ is the category of categories and functors. Strictly speaking, $\mathbb{C}at$ is not a category but only a so-called quasicategory, which is a category that lives in a higher set-theoretic universe.

systems [11]. However, inclusion systems are too strong for our purposes and therefore we will work under weaker assumptions.

Definition 2.3 An *inclusive category* [14] is a category having a broad subcategory[6] which is a partially ordered class with finite products and coproducts, called intersection (denoted $\cap$) and union (denoted $\cup$) such that for each pair of objects A, B, $A \cup B$ is a pushout of $A \cap B$ in the category.

A category *has pushouts which preserve inclusions* if there exists a pushout

$$\begin{array}{ccc} A & \hookrightarrow & A' \\ \downarrow & & \downarrow \\ B & \hookrightarrow & B' \end{array}$$

for each span where one arrow is an inclusion.

A functor between two inclusive categories is inclusive if it takes inclusions in the source category to inclusions in the target category.

Definition 2.4 An institution is *weakly inclusive* if

- **Sign** is inclusive and has pushouts which preserve inclusions,
- **Sen** is inclusive, and
- each model category have a broad subcategory of inclusions.

Let I be a weakly inclusive institution. We say that I *has differences*, if there is a binary operation $\setminus$ on signatures, such that for each pair of signatures Σ_1, Σ_2, we have

1. $\Sigma_1 \setminus \Sigma_2 \subseteq \Sigma_1$
2. $(\Sigma_1 \setminus \Sigma_2) \cap \Sigma_2 = \emptyset$
3. for any Σ with the properties 1. and 2. above, $\Sigma \subseteq \Sigma_1 \setminus \Sigma_2$.

2.2 Translations Between Institutions

Several notions of *translations* between institutions can be introduced. The most frequently used variant are *institution comorphisms* [13]. A comorphism from institution L_1 to institution L_2 maps L_1-signatures to L_2-signatures along a functor Φ and Σ-sentences in L_1 to $\Phi(\Sigma)$-sentences in L_2, for each L_1-signature Σ, while $\Phi(\Sigma)$-models are mapped to Σ-models. Again, a satisfaction condition has to be fulfilled. For *institution morphisms* [13], the directions of the translation of sentences and models are reversed.

[6] That is, with the same objects as the original category.

Definition 2.5 An **institution comorphism** from an institution $I = (\mathbf{Sign}^I, \mathbf{Mod}^I, \mathbf{Sen}^I, \models^I)$ to an institution $J = (\mathbf{Sign}^J, \mathbf{Mod}^J, \mathbf{Sen}^J, \models^J)$ consists of a functor $\Phi\colon \mathbf{Sign}^I \longrightarrow \mathbf{Sign}^J$, and two natural transformations $\beta\colon \mathbf{Mod}^J \circ \Phi \Longrightarrow \mathbf{Mod}^I$ and $\alpha\colon \mathbf{Sen}^I \Longrightarrow \mathbf{Sen}^J \circ \Phi$, such that

$$M' \models^J_{\Phi(\Sigma)} \alpha_\Sigma(\varphi) \Leftrightarrow \beta_\Sigma(M') \models^I_\Sigma \varphi.$$

holds, called the **satisfaction condition**. □

Here, $\Phi(\Sigma)$ is the translation of the signature Σ from institution I to institution J, $\alpha_\Sigma(\varphi)$ is the translation of the Σ-sentence φ to a $\Phi(\Sigma)$-sentence, and $\beta_\Sigma(M')$ is the translation (or perhaps better: reduction) of the $\Phi(\Sigma)$-model M' to a Σ-model. The naturality of α and β mean that for each signature morphism $\sigma : \Sigma \to \Sigma'$ in I the following squares commute:

$$\begin{array}{ccc} Sen^I(\Sigma) & \xrightarrow{\alpha_\Sigma} & Sen^J(\Phi(\Sigma)) \\ \downarrow{\scriptstyle Sen^I(\sigma)} & & \downarrow{\scriptstyle Sen^J(\Phi(\sigma))} \\ Sen^I(\Sigma') & \xrightarrow[\alpha_{\Sigma'}]{} & Sen^J(\Phi(\Sigma')) \end{array} \qquad \begin{array}{ccc} Mod^J(\Phi(\Sigma')) & \xrightarrow{\beta_{\Sigma'}} & Mod^I(\Sigma') \\ \downarrow{\scriptstyle Mod^J(\Phi(\sigma))} & & \downarrow{\scriptstyle Mod^I(\sigma)} \\ Mod^J(\Phi(\Sigma)) & \xrightarrow[\beta_\Sigma]{} & Mod^I(\Sigma) \end{array}$$

Definition 2.6 An **institution morphism** from an institution $I = (\mathbf{Sign}^I, \mathbf{Mod}^I, \mathbf{Sen}^I, \models^I)$ to an institution $J = (\mathbf{Sign}^J, \mathbf{Mod}^J, \mathbf{Sen}^J, \models^J)$ consists of a functor $\Phi\colon \mathbf{Sign}^I \longrightarrow \mathbf{Sign}^J$, and two natural transformations $\beta\colon \mathbf{Mod}^I \Longrightarrow \mathbf{Mod}^J \circ \Phi$ and $\alpha\colon \mathbf{Sen}^J \circ \Phi \Longrightarrow \mathbf{Sen}^I$, such that

$$M \models^I_\Sigma \alpha_\Sigma(\varphi) \Leftrightarrow \beta_{\Phi(\Sigma)}(M) \models^J_{\Phi(\Sigma)} \varphi.$$

holds, called the **satisfaction condition**.

Mappings of institutions are split along the following dichotomies:

- *Translation* versus *projection*: a translation embeds or encodes a logic into another one, while a projection is a forgetful operation (e.g., the projection from first-order logic to propositional logic forgets predicates with arity greater than zero). It is an interesting informal observation that translations can be formalized as institution comorphisms, and projections as institution morphisms.
- *Plain mapping* versus *simple theoroidal mapping* [13]: while a plain mapping needs to map signatures to signatures, a (simple) theoroidal mapping maps signatures to theories. The latter therefore allows for using "infrastructure axioms": e.g., when mapping OWL to Common Logic, it is convenient to rely on a first-order axiomatization of a transitivity predicate for properties.

Mappings can also be classified according to their accuracy; see [36] for details. *Sublogics* are the most accurate mappings: they are syntactic subsets. *Embeddings* come close to sublogics, like injective functions come close to subsets. A mapping can be *faithful* in the sense that logical consequence (or logical deduction) is preserved and reflected, that is, inference systems and reasoning engines for the target logic can be reused for the source logic (along the mapping). *(Weak) exactness* is a technical property that guarantees this faithfulness even in the presences of OMS structuring operations [4].

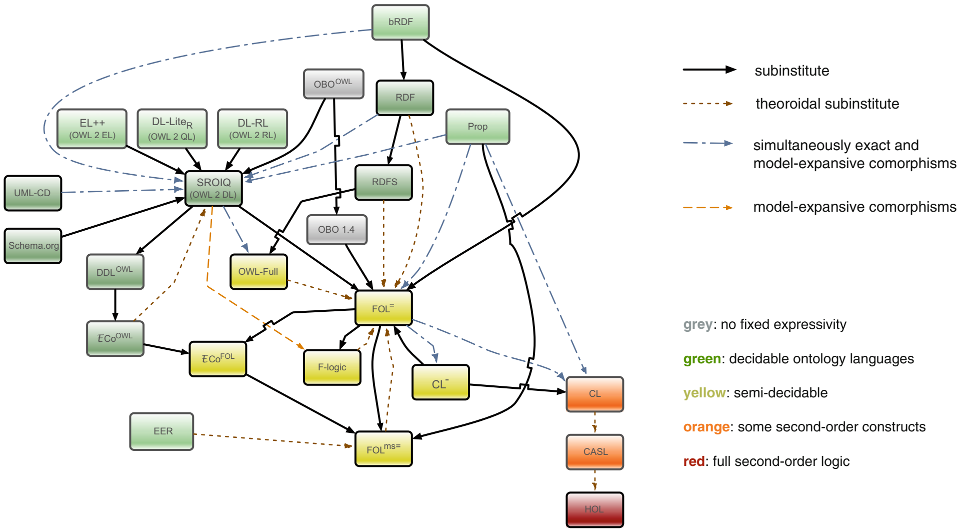


Fig. 1 The current logic translation graph for DOL-conforming languages

2.3 A Graph of Logic Translations

Figure 1 is a revised and extended version of the graph of logics and translations introduced in [36]. New nodes include UML class diagrams, OWL-Full (i.e. OWL with an RDF semantics instead of description logic semantics), and Common Logic without second-order features (CL$^-$). We have defined the translations between most of these logics in earlier publications [36, 38]. The definitions of the DOL conformance of some central standard OMS languages and translations among them will be given as annexes to the standard and published in an open registry, which is also the place where the remaining definitions will be maintained.

3 The Language DOL

3.1 DOL Syntax and Semantics

The DOL language is not "yet another OMS language," but a *metalanguage* for expressing relations between OMS. Therefore, any OMS written in any conforming OMS language also is a DOL OMS. Therefore, when working with DOL users can reuse OMS as they are, no changes are required.

DOL provides abstract syntax categories for:

1. OMS (ontologies, models and specifications). *Basic OMS* are OMS that are written in some OMS language (e.g., OWL or CASL). A *modular* or *structured* OMS is written in

a modular way, with the help of DOL structuring operations. A heterogeneous OMS is a modular OMS that involves modules, which are written in different OMS languages. The semantics of OMS is given by a signature and a class of models. In some cases, we can additionally provide a theory-level semantics of OMS, as a signature and a class of sentences that, if it exists, agrees with the model-level semantics (that is, the model class is equal to the class of models satisfying the theory). We call an OMS *flattenable* if it has a theory-level semantics and *elusive* if it only admits a model-level semantics. Whether an OMS is flattenable can be decided based on the structuring operations on OMS, as follows:

Flattenable OMS: basic OMS are flattenable; if all their components OMS are flattenable, then the following operations on OMS yield flattenable OMS: extension, union, translation, interpolate/forget, extract, reference, qualification, combination.

Elusive OMS: the reduction, minimization, or maximization of an OMS is elusive; further, any OMS containing an elusive OMS is elusive.[7]

For detailed definitions of these types of OMS, see Sect. 3.2.

2. OMS *mappings*. They denote relations between two OMS or OMS networks, typically along a signature morphism. Some mappings may also involve other OMS or other signatures. Examples of OMS mappings are interpretations (specifying a logical consequence relationship between OMS), equivalences of OMS (specifying that their model classes are in bijective correspondence), conservative extensions (between OMS and their modules), OMS alignment. They are presented in Sect. 3.3.
3. *Networks* of OMS. Networks are graphs with nodes labeled with OMS and edges labeled with OMS mappings. The edges show how two OMS are interlinked. The rationale behind networks is that they provide a way to specify or model complex distributed systems (or domains), where a single OMS would become too complex (this especially can be the case if the OMS are formulated in different OMS languages). Instead, the different OMS of the network provide different viewpoints on the system, while their compatibility is ensured via mappings. Networks are discussed in Sect. 3.4.
4. *Libraries* of OMS. OMS, mappings, and networks are organized in libraries. A library consists of a list of declarations involving (possibly modular and/or heterogeneous) OMS. These declarations can be definitions (assigning a name to an OMS, OMS mapping, or network of OMS) and qualifications of the current language, logic and/or serialization. This is detailed in Sect. 3.5.

The semantics of DOL is based on a fixed (but in principle arbitrary) logic graph. A logic graph is given by a collection of institutions, institution morphisms, and institution comorphisms (serving as logics, logic reductions and logic translations). Moreover, some of the institution comorphisms are marked as default translations and some of the institution morphisms are marked as default projection (but only at most one between a given source and target institution).

---

[7] Note that extension, union, translation, reference, qualification and combination are defined for flattenable and elusive OMS, while interpolate/forget and extract are only defined for flattenable OMS.

We assume that for each institution in the logic graph there is a trivial signature $\emptyset$ with model class $\mathcal{M}_\emptyset$ and such that there exists a unique signature morphism from $\emptyset$ to any signature of the institution. Moreover, we assume the existence of a designated error logic in the graph, and a partial union operation on logics, denoted by $\bigcup$: $L_1 \bigcup L_2 = (L, \rho_1 : L_1 \to L, \rho_2 : L_2 \to L)$, when defined.

3.2 Modular and Heterogeneous OMS

Modular and heterogeneous OMS are generated by the following grammar, where Σ is a signature, Δ is a set of sentences over Σ, σ a signature morphism, I an institution, ρ an institution comorphism and μ an institution morphism[8]:

```
OMS ::= ⟨I, Σ, Δ⟩
    | IRI
    | OMS and OMS | OMS then OMS
    | OMS with σ | OMS with translation ρ
    | OMS reveal Σ | OMS hide Σ | OMS hide along μ
    | OMS keep Σ [with I] | OMS keep I | OMS forget Σ [with I] | OMS
       forget I
    | OMS extract Σ | OMS remove Σ
    | OMS select ⟨Σ, Δ⟩ | OMS reject ⟨Σ, Δ⟩
    | minimize OMS | maximize OMS
    | combine Network
```

The semantics of an OMS O has four components:

- the institution of O, denoted $\mathbf{Inst}(O)$,
- the signature of O, denoted $\mathbf{Sign}(O)$ (which is a signature in $\mathbf{Inst}(O)$),
- the models of O, denoted $\mathbf{Mod}(O)$ (which is a class of models over $\mathbf{Sign}(O)$),
- the axioms of O, denoted $\mathbf{Ax}(O)$ (which is a set of sentences over $\mathbf{Sign}(O)$).[9]

For elusive OMS, $\mathbf{Ax}(O)$ is undefined. For flattenable OMS, $\mathbf{Mod}(O)$ can be obtained as $\mathbf{Mod}(O) = \{M \in \mathbf{Mod}(\mathbf{Sign}(O)) \mid M \models \mathbf{Ax}(O)\}$.

In the following, we discuss the different kinds of (possibly modular and/or heterogeneous) OMS.

[8] This is a mathematically abstracted version of DOL. In reality, signatures are represented by symbol sets, and signature morphisms by symbol maps. The details of passing from symbol sets (resp. maps) to signatures (resp. signature morphisms) are left out here. Also, we have left out OMS bridges, since their design is still being discussed.

[9] The *theory* of O, written, $\mathbf{Th}(O)$, is the closure of $\mathbf{Ax}(O)$ under logical entailment. Note, however, that throughout the text we use “theory” also more informally as denoting some set of axioms in a particular signature and logic.

3.2.1 Basic OMS

A *basic OMS O* written inline, in a conforming OMS language and serialization. The semantics is inherited from the OMS language I[10] and results in a theory $\langle \Sigma, \Delta \rangle$ (therefore, for simplicity, in the syntax above, we have identified the basic OMS with $\langle I, \Sigma, \Delta \rangle$). O can also be an OMS fragment, which means that some of the symbols or axioms may refer to symbols declared outside O (i.e., in an imported OMS). This is mainly used for extensions and equivalences. Here are two sample ontologies in OWL (using Manchester syntax) and Common Logic (using CLIF):

```
Class: Woman EquivalentTo: Person and Female
ObjectProperty: hasParent

(cl-module PreOrder
  (forall (x) (le x x))
  (forall (x y z) (if (and (le x y) (le y z)) (le x z))))
```

Formally,

- $\mathbf{Inst}(I, \Sigma, \Delta) = I$
- $\mathbf{Sign}(I, \Sigma, \Delta) = \Sigma$
- $\mathbf{Mod}(I, \Sigma, \Delta) = \{M \in \mathbf{Mod}(\Sigma) \mid M \models \Delta\}$
- $\mathbf{Ax}(I, \Sigma, \Delta) = \Delta$.

3.2.2 IRI Reference

An IRI reference to an OMS existing on the Web,[11] possibly abbreviated using prefixes.[12] For example:

```
  <http://owl.cs.manchester.ac.uk/co-ode-files/ontologies/pizza.owl>
```

or alternatively

```
%prefix(
  co-ode: <http://owl.cs.manchester.ac.uk/co-ode-files/ontologies/> )%
co-ode:pizza.owl
```

The semantics of such an IRI reference would require a *global environment* mapping IRIs to (semantics of) OMS. However, for simplicity, we omit the global environment (and therefore also the semantics of IRI references) here.

[10] I is normally determined by the context of the enclosing library and passed around as an additional parameter of the semantics. For simplicity, here we let I become part of the basic OMS.
[11] Note that not all OMS can be downloaded by dereferencing their IRIs. Implementing a catalogue mechanism in DOL-aware applications might remedy this problem.
[12] Some of the following listings abbreviate IRIs using prefixes but omit the prefix bindings for readability.

3.2.3 Extension

An *extension* of an OMS by new symbols and axioms, written O_1 **then** O_2, where O_2 is an OMS (fragment) in a conforming OMS language. The resulting signature is that of O_1, augmented with the symbols in O_2. A model of an extension OMS is a model of this signature, that satisfies the axioms on O_2 and is (when appropriately reduced) a model of O_1. An extension can optionally be marked as conservative (%mcons or %ccons after the "**then**"). The semantics is that each O_1-model must have at least one expansion to the whole extension O_1 **then** O_2 (for %mcons) resp. that each logical consequence of O_1 **then** O_2 is already one of O_1 if it is over the signature of O_1 (for %ccons). In case that O_2 does not introduce any new symbols, the keyword %implied can be used instead of %ccons or %mcons; the extension then merely states intended logical consequences. The keyword %def stands for definitional extensions. This is similar to %mcons, but the model expansion must always exist uniquely. The following OWL ontology is an example for the latter:

```
  Class Person
  Class Female
then %def
  Class: Woman EquivalentTo: Person and Female
```

The semantics of $O = O_1$ **then** O_2 is

- $\mathbf{Inst}(O) = \mathbf{Inst}(O_1) = \mathbf{Inst}_{O_1}(O_2)$
- $\mathbf{Sign}(O) = \mathbf{Sign}(O_1) \cup \mathbf{Sign}_{O_1}(O_2)$
- $\mathbf{Mod}(O) = \{M \in \mathbf{Mod(Sign}(O)) \mid M|_{\mathbf{Sign}(O_i)} \in \mathbf{Mod}(O_i),\ \text{for } i = 1, 2\}$
- $\mathbf{Ax}(O) = \mathbf{Ax}(O_1) \cup \mathbf{Ax}_{O_1}(O_2)$

where O_2 is analyzed in the context of previous declarations in O_1, as indicated by adding an index in its semantics.

3.2.4 Union

A *union* of two self-contained OMS (not fragments), written O_1 **and** O_2. Models of this union are those models that are (perhaps after appropriate reduction) models of both O_1 and O_2. For example, the class of commutative monoids can be expressed as

```
algebra:Monoid and algebra:Commutative
```

Forming a union of OMS is a particularly common operation in the RDF logic, where it is known as merging graphs [18, section 0.3]; however, the RDF language provides no explicit syntax for this operation. When multiple RDF ontologies ("graphs") contain statements about the same symbol ("resource"), i.e., syntactically, triples having the same subject, the effect is that in the merged graph the resource will have all properties that have previously been stated about it separately. Different kinds of properties, e.g., multilingual labels, geodata, or outgoing links to external graphs, are often maintained in different RDF graphs, which are then merged; consider the following excerpt:

```
{ :OVGU rdfs:label "Otto-von-Guericke-Universität Magdeburg"@de . } and
{ :OVGU geo:lat "52.1403"^^xsd:float . } and
{ :OVGU owl:sameAs[13]
    <http://de.dbpedia.org/page/OvGU> . }
```

The semantics of $O = O_1$ **and** O_2 is

- $\mathbf{Inst}(O) = I$ where $\mathbf{Inst}(O_1) \bigcup \mathbf{Inst}(O_2) = (I, (\Phi_1, \alpha_1, \beta_1) : \mathbf{Inst}(O_1) \to I, (\Phi_2, \alpha_2, \beta_2) : \mathbf{Inst}(O_2) \to I)$
- $\mathbf{Sign}(O) = \Phi_1(\mathbf{Sign}(O_1)) \cup \Phi_2(\mathbf{Sign}(O_2))$
- $\mathbf{Mod}(O) = \{M \in \mathbf{Mod}(\mathbf{Sign}(O)) \mid \beta_{\Sigma_i}(M|_{\Phi_i(\mathbf{Sign}(O_i))}) \in \mathbf{Mod}(O_i),\ \text{for } i = 1, 2\}$
- $\mathbf{Ax}(O) = \alpha_1(\mathbf{Ax}(O_1)) \cup \alpha_2(\mathbf{Ax}(O_2))$.

3.2.5 Translation

A *translation* of an OMS to a different signature (written O **with** σ, where σ is a signature morphism) or into some OMS language (written O **with translation** ρ, where ρ is an institution comorphism). For example, we can combine an OWL ontology with a first-order axiom (formulated in Common Logic) as follows:

```
logic OWL : {
  ObjectProperty: isProperPartOf
  Characteristics: Asymmetric
  SubPropertyOf: isPartOf }
with translation OWL22CommonLogic
then
  (if (and (isProperPartOf x y) (isProperPartOf y z)) (isProperPartOf
      x z))
```

Note that OWL can express transitivity, but not together with asymmetry.

The semantics of $O = O'$ **with** σ is

- $\mathbf{Inst}(O) = \mathbf{Inst}(O')$
- $\mathbf{Sign}(O) = \Sigma'$ where $\sigma : \mathbf{Sign}(O') \to \Sigma'$
- $\mathbf{Mod}(O) = \{M \in \mathbf{Mod}(\Sigma') \mid M|_\sigma \in \mathbf{Mod}(O')\}$
- $\mathbf{Ax}(O) = \sigma(\mathbf{Ax}(O'))$.

The semantics of $O = O'$ **with translation** ρ is

- $\mathbf{Inst}(O) = I$, where $\rho = (\Phi, \alpha, \beta) : \mathbf{Inst}(O') \to I$
- $\mathbf{Sign}(O) = \Phi(\mathbf{Sign}(O))$
- $\mathbf{Mod}(O) = \{M \in \mathbf{Mod}(\mathbf{Sign}(O)) \mid \beta_{\mathbf{Sign}(O)}(M) \in \mathbf{Mod}(O')\}$
- $\mathbf{Ax}(O) = \alpha_{\mathbf{Sign}(O)}(\mathbf{Ax}(O'))$.

[13] While *owl:same as* is borrowed from the *vocabulary* of OWL, it is commonly used in the RDF logic to link to resources in external graphs, which should be treated as if their IRI were the same as the subject's IRI.

3.2.6 Reduction

A *reduction* of an OMS to a smaller signature Σ is written O **reveal** Σ. Alternatively, it can be written O **hide** Σ, where Σ is the set of symbols to be hidden (i.e. this is equivalent to O **reveal** $\mathsf{Sig}(O) \setminus \Sigma$). The effect is an existential quantification over all hidden symbols. For example, when specifying a group in sorted first-order logic, using the CASL language,

```
sort Elem
ops 0: Elem; __+__: Elem * Elem -> Elem; inv: Elem -> Elem
forall x, y, z : Elem
       . 0 + x       = x
       . x + (y + z) = (x + y) + z
       . x + inv(x)  = 0
reveal Elem, 0, __+__
```

revealing everything except the inverse operation `inv` results in a specification of the class of all monoids that can be extended with an inverse operation, i.e., the class of all groups with inverse left implicit.

Here is an example of hiding:

```
ontology Pizza = %% a~simplified remake of the Pizza ontology [19]
  Individual: TomatoTopping
  Individual: MozzarellaTopping DifferentFrom: TomatoTopping
  ObjectProperty: hasTopping
  Class: VegetarianTopping
    EquivalentTo: { TomatoTopping, MozzarellaTopping, ... }
  Class: VegetarianPizza SubClassOf: some hasTopping VegetarianTopping
  ...
end

ontology Pizza_hide_VegetarianTopping =
  Pizza hide VegetarianTopping
end
```

A reduction to a less expressive logic is written O **hide along** μ, where μ is an institution morphism. This is a common operation in TBox/ABox settings, where an ontology in an expressive language provides the terminology (TBox) used in assertions (ABox) stated in a logic that is less expressive but scales to larger data sets; OWL DL (whose logic is $\mathcal{SROIQ}$) vs. RDF is a typical language combination:

```
ontology TBoxABox =
  Pizza hide along OWL22RDF
  then logic RDF : {
    :myPizza :hasTopping
      [ a :TomatoTopping ], [ a :MozzarellaTopping ] .
  }
```

The semantics of $O = O'$ **reveal** Σ' is

- $\mathbf{Inst}(O) = \mathbf{Inst}(O')$
- $\mathbf{Sign}(O) = \Sigma'$
- $\mathbf{Mod}(O) = \{M|_\iota \mid M \in \mathbf{Mod}(O')\}$ where $\iota : \Sigma' \to \mathbf{Sign}(O')$ is the inclusion
- $\mathbf{Ax}(O)$ is undefined.

The semantics of $O = O'$ **hide** Σ' is

- $\mathbf{Inst}(O) = \mathbf{Inst}(O')$
- $\mathbf{Sign}(O) = \mathbf{Sign}(O') \setminus \Sigma'$
- $\mathbf{Mod}(O) = \{M|_\iota \mid M \in \mathbf{Mod}(O')\}$ where $\iota : \mathbf{Sign}(O') \setminus \Sigma' \to \mathbf{Sign}(O')$ is the inclusion
- $\mathbf{Ax}(O)$ is undefined.

The semantics of $O = O'$ **hide along** μ is

- $\mathbf{Inst}(O) = I$ where $\mu = (\Phi, \alpha, \beta) : \mathbf{Inst}(O) \to I$
- $\mathbf{Sign}(O) = \Phi(\mathbf{Sign}(O'))$
- $\mathbf{Mod}(O) = \{\beta_{\mathbf{Sign}(O')}(M) \mid M \in \mathbf{Mod}(O')\}$
- $\mathbf{Ax}(O)$ is undefined.

3.2.7 Filtering

A *filtering* O **select** $\langle\Sigma, \Delta\rangle$, which selects those sentences from O that have signature Σ, plus those in Δ (where Δ is a subset $\mathbf{Ax}(O)$). It can also be written O **reject** $\langle\Sigma, \Delta\rangle$, where Σ is the set of symbols and Δ the set of axioms to be hidden. For example, we can select all axioms of Galen[14] involving Drugs, Joints, or Bodyparts by

```
logic OWL
ontology myGalen =
   <http://example.org/GALEN/galen.owl>
 select Drugs, Joints, Bodyparts
end
```

The semantics of $O = O'$ **select** $\langle\Sigma, \Delta\rangle$ is defined only if $\Sigma \subseteq \mathbf{Sign}(O)$ and $\Delta \subseteq \mathbf{Ax}(O)$, and in that case, it is given by

- $\mathbf{Inst}(O) = \mathbf{Inst}(O')$
- $\mathbf{Sign}(O) = \Sigma'$ where Σ' is the smallest signature with $\Sigma \subseteq \Sigma'$ and $\Delta \subseteq \mathbf{Sen}(\Sigma)$[15]
- $\mathbf{Ax}(O) = (\mathbf{Ax}(O') \cap \mathbf{Sen}(\mathbf{Sign}(O))) \cup \Delta$
- $\mathbf{Mod}(O)$ is the class of all $\mathbf{Ax}(O)$-models.

The semantics of $O = O'$ **reject** (Σ, Δ) is

- $\mathbf{Inst}(O) = \mathbf{Inst}(O')$
- $\mathbf{Sign}(O) = \mathbf{Sign}(O') \setminus \Sigma$
- $\mathbf{Ax}(O) = \mathbf{Ax}(O') \cap \mathbf{Sen}(\mathbf{Sign}(O)) \setminus \Delta$
- $\mathbf{Mod}(O)$ is the class of all $\mathbf{Ax}(O)$-models.

3.2.8 Interpolation

An *approximation* [31] (or technically, *uniform interpolation*) of an OMS, either in a sub-signature or a sublogic (written O **keep** Σ, O **keep** Σ **with** L or O **keep** L, where Σ is

[14] We assume that GALEN is available as an OWL ontology.
[15] If this smallest signature does not exist, the semantics is undefined.

a signature and L is a logic). The effect is that sentences not expressible in Σ (resp. L) are weakened or removed, but the resulting theory still has the same consequences, as fas as these are expressible in Σ (and/or L). Technically, this is a uniform interpolant [32, 46]. For example, we can interpolate the first-order DOLCE mereology in OWL:[16]

```
DOLCE_Mereology keep OWL
```

Dually, O **forget** Σ or O **forget** Σ **with** L interpolates O with the signature $\mathsf{Sig}(O) \setminus \Sigma$, i.e. Σ specifies the symbols that need to be left out (and optionally, L specifies a sublogic that needs to be targeted). Cf. the notion of forgetting in [32, 46]. For example,

```
  Pizza forget VegetarianTopping
```

This has both a model-theoretic and a theory-level semantics, i.e., it yields a theory in the reduced signature (without `VegetarianTopping`). In contrast, `Pizza` **`hide`** `VegetarianTopping` has only a model-level semantics (see also the comparison in Sect. 3.2.12).

The semantics of $O = O'$ **keep** Σ **with** I is

- $\mathbf{Inst}(O) = I$ and $(\Phi, \alpha, \beta) : \mathbf{Inst}(O') \to I$ is the default projection (in case I is missing, it is the identity on $\mathbf{Inst}(O')$)
- $\mathbf{Sign}(O) = \Phi(\Sigma)$
- $\mathbf{Ax}(O) = \alpha^{-1}_{\mathbf{Sign}(O')}(\mathbf{Ax}(O')^{\bullet}) \cap \mathbf{Sen}^{I}(\mathbf{Sign}(O))$,[17] i.e. that part of $\mathbf{Ax}(O')$ that can be expressed in the smaller signature and logic
- $\mathbf{Mod}(O)$ is the class of $\mathbf{Ax}(O)$-models

The semantics of O **forget** Σ' **with** I is the same as the semantics of O **keep** $(\mathbf{Sign}(O) \setminus \Sigma')$ **with** I.

3.2.9 Extraction

A module *extracted* from an OMS, written O **extract** Σ, where Σ is a subsignature of $\mathsf{Sig}(O)$. The extracted module is a subOMS of O with signature larger than (or equal to) Σ, such that O is a conservative extension of the extracted module. Intuitively, a module (in the sense of module extraction) is a small sub-OMS that says the same about Σ as the OMS O itself. For example, we can extract from GALEN a module referring to drugs, joints and body parts:

```
logic OWL
ontology myGalen =
   <http://example.org/GALEN/galen.owl>
 extract Drugs, Joints, Bodyparts
end
```

(This example is continued in Sect. 3.3.5).

[16] Interpolants need not always exist, and even if they do, tools might only be able to approximate them.
[17] In practice, one looks for a finite subset that still is logically equivalent to this set. Note that $\Delta^{\bullet}$ is the set of logical consequences of Δ, i.e. $\Delta^{\bullet} = \mathbf{Th}(\Delta)$.

The semantics of $O = O'$ **extract** Σ is

- $\mathbf{Inst}(O) = \mathbf{Inst}(O')$
- $\mathbf{Sign}(O) = \Sigma \cup \Sigma'$
- $\mathbf{Ax}(O) = \Delta'$
- $\mathbf{Mod}(O)$ is the class of $\mathbf{Ax}(O)$-models

where $\langle \Sigma', \Delta' \rangle$ is the smallest depleting Σ-module [22], i.e. the smallest[18] subtheory $\langle \Sigma', \Delta' \rangle$ of $(\mathbf{Sign}(O'), \mathbf{Ax}(O'))$ such that the following model-theoretic inseparability holds

$$\mathbf{Ax}(O') \setminus \Delta' \equiv_{\Sigma' \cup \Sigma} \emptyset.$$

This means intuitively that $\mathbf{Ax}(O') \setminus \Delta'$ cannot be distinguished from $\emptyset$ (what $\Sigma' \cup \Sigma$ concerns) and formally that

$$\begin{array}{ll} & \{M|_{\Sigma' \cup \Sigma} \mid M \in \mathbf{Mod}(\mathbf{Sign}(O')), M \models \mathbf{Ax}(O') \setminus \Delta'\} \\ = & \{M|_{\Sigma' \cup \Sigma} \mid M \in \mathbf{Mod}(\mathbf{Sign}(O'))\}. \end{array}$$

Dually, O **remove** Σ extracts w.r.t. the signature $\mathsf{Sig}(O) \setminus \Sigma$,[19] i.e. the semantics is given by that of O **extract** $\mathsf{Sig}(O) \setminus \Sigma$.

3.2.10 Combination

A *combination* of OMS, written **combine** N, where N is a network. The simplest example of a combination is a disjoint union (we here translate OWL OMS into many-sorted OWL in order to be able to distinguish between different universes of individuals):

```
ontology Publications1 =
  Class: Publication
  Class: Article SubClassOf: Publication
  Class: InBook SubClassOf: Publication
  Class: Thesis  SubClassOf: Publication
  ...

ontology Publications2 =
  Class: Thing
  Class: Article SubClassOf: Thing
  Class: BookArticle SubClassOf: Thing
  Class: Publication SubClassOf: Thing
  Class: Thesis  SubClassOf: Thing
  ...

logic MS-OWL
```

[18] If the smallest such subtheory does not exist, the semantics is undefined. In [22], it is shown that it does exist in usual institutions.

[19] Note that the resulting module can still contain symbols from Σ, because the resulting signature may be enlarged.

```
network Publications_Network =
 1 : Publications1 with translation OWL2MS-OWL,
 2 : Publications2 with translation OWL2MS-OWL
end

ontology Publications_Combined =
combine
  Publications_Network
  %% implicitly: Article ↦ 1:Article ...
  %%             Article ↦ 2:Article ...
end
```

If mappings or alignments are present, the semantics of a combination is a quotient of a disjoint union (symbols related along the edges are identified). Technically, this is a colimit, see [7, 48]. An example for this is given along with the examples for alignments below.

The semantics of $O = \textbf{combine } N$ is

- $\textbf{Inst}(O) = I$
- $\textbf{Sign}(O) = \Sigma$, where $(I, \Sigma, \{\mu_i\}_{i\in|G|})$ is the colimit of the graph G given by the semantics of N
- $\textbf{Ax}(O) = \cup_{i\in|G|}\mu_i(\textbf{Ax}(O_i))$, where O_i is the OMS label of the node i in G
- $\textbf{Mod}(O) = \{M \in \textbf{Mod}(\Sigma) \mid M|_{\mu_i} \in \textbf{Mod}(O_i), i \in |G|\}$, where O_i is the OMS label of the node i in G.

3.2.11 Minimization

A *minimization* of an OMS imposes a closed-world assumption on part of the OMS. It forces the nonlogical symbols declared in O to be interpreted in a minimal way. This is written **minimize** $\{ O \}$. Symbols declared before the minimized part are considered to be fixed for the minimization (that is, we minimize among all models with the same reduct). Symbols declared after the minimization can be varied. This is borrowed from circumscription [3, 29]. Alternatively, the nonlogical symbols to be minimized and to be varied can be explicitly declared: O **minimize** Σ_1 **vars** Σ_2. For example, in the following OWL theory, B2 is a block that is not abnormal, because it is not specified to be abnormal, and hence it is also on the table.

```
  Class: Block
  Individual: B1 Types: Block
  Individual: B2 Types: Block DifferentFrom: B1
then minimize {
        Class: Abnormal
        Individual: B1 Types: Abnormal }
then
  Class: OnTable
  Class: BlockNotAbnormal EquivalentTo:
    Block and not Abnormal SubClassOf: OnTable
then %implied
  Individual: B2 Types: OnTable
```

The semantics of O = **minimize** O' is

- **Inst**(O) = **Inst**(O')
- **Sign**(O) = **Sign**(O')
- **Mod**$(O) = \{M \in$ **Mod**$(O') \mid M$ is minimal in **Mod**$(O')\}$
- **Ax**(O) is undefined.

Note that for minimality we need the inclusions in model categories. Dually to minimizations, there are also maximizations.

3.2.12 Hide vs. Remove vs. Forget vs. Reject

We have four ways of removing the class `VegetarianTopping` from the ontology `Pizza` using the keywords **hide**, **remove**, **forget**, and **reject**, respectively. Table 1 illustrates some of the connections between Sects. 3.2.6 and 3.2.9.

Using hiding, we keep the model class of `Pizza`, but just remove the interpretation of `VegetarianTopping` from each model. Note that the resulting ontology has

```
VegetarianPizza SubClassOf:
  Annotations: dol:iri (*)
  some hasTopping { TomatoTopping, MozzarellaTopping, ... }
```

as a logical consequence. This is also a consequence of the corresponding uniform interpolant

```
  Pizza forget VegetarianTopping
```

which captures the theory of `Pizza` **`hide`** `VegetarianTopping`. Note that there is a subtle difference between (model-theoretic) hiding and (consequence-theoretic) forgetting: a model satisfying the *theory* of O **hide** Σ might itself not be a model of O **hide** Σ. In examples involving "**with** L," the uniform interpolant can be weaker than the hiding, because it is only required to have the same logical consequences in some language L, and a formula like (*) might not be a formula of L. Also, an extracted module does not contain (*), because it only selects a subontology, and `Pizza` does not contain (*). Finally, `Pizza` **`reject`** `VegetarianTopping` simply drops all sentences involving `VegetarianTopping`, and therefore also consequences like (*) are lost.

Note that while forget/keep and hide/reveal both work w.r.t. smaller signatures and sublogics, remove/extract and select/reject do not work for sublogics. This is because remove/extract must always respect the conservative extension property, which may not be possible when projecting to a sublogic. And if conservativity cannot be guaranteed, then forget/keep can be used in any case. In the case of select/reject, it is unclear what selecting of a sublogic should bring other than projecting to the sublogic using **hide along**.

Proposition 3.1 The following relations among the constructs in Table 1 hold:

$$\begin{aligned} & \mathbf{Mod}(O\ \mathbf{hide}\ \Sigma) \\ = \ & \mathbf{Mod}(O\ \mathbf{remove}\ \Sigma)|_{Sig(O)\setminus\Sigma} \\ \subseteq \ & \mathbf{Mod}(O\ \mathbf{forget}\ \Sigma) \\ \subseteq \ & \mathbf{Mod}(O\ \mathbf{reject}\ \Sigma) \end{aligned}$$

Table 1 Hiding – Extraction – Approximation – Filtering

	Hide/reveal	Remove/extract	Forget/keep	Select/reject
Semantic background	Model reduct	Conservative extension	Uniform interpolation	Theory filtering
Relation to original	Interpretable	Subtheory	Interpretable	subtheory
Approach	Model level	Theory level	Theory level	Theory level
Type of OMS	Elusive	Flattenable	Flattenable	Flattenable
Signature of result	$= \Sigma$	$\geq \Sigma$	$= \Sigma$	$\geq \Sigma$
Change of logic	Possible	Not possible	Possible	Not possible
Application	Specification	Ontologies	Ontologies	Blending

3.3 OMS Mappings

OMS mappings are generated by the following grammar:

```
MappingDefn ::= interpretation NAME : OMS to OMS = σ
        | entailment NAME = OMS entails OMS
        | equivalence NAME : OMS <-> OMS  = ⟨Σ, Δ⟩
        | conservative extension NAME = O1 of O2 for Σ
        | alignment NAME CARD1 CARD2 : OMS to OMS = Correspondences
        | refinement NAME : OMS to OMS = σ
        | refinement NAME = NAME then NAME
```

The semantics of an OMS mapping is given as a graph whose nodes N are labeled with

- $Name(N)$, the name of the node
- $\mathbf{Inst}(N)$, the institution of the node
- $\mathbf{Sign}(N)$, the signature of the node
- $\mathbf{Mod}(N)$, the class of $\mathbf{Sign}(N)$-models of the node
- $\mathbf{Ax}(N)$, the set of $\mathbf{Ax}(N)$-sentences of the node

and whose edges are labeled with signature morphisms between the signatures of the source and target nodes. The theory of a node corresponding to an elusive OMS may be undefined. The class of models of a node corresponding to a flattenable OMS is the class of models of $\mathbf{Ax}(N)$. For brevity, we may write the label of a node as a tuple. We make the simplifying assumption that any OMS is assigned a unique name. The theory-level semantics of an OMS is needed for alignments.

In the following we discuss the different types of OMS mappings.

3.3.1 Interpretation

Theory *interpretations*, written **interpretation** $Id : O_1$ **to** $O_2 = \sigma$, expressing that the σ-reduct of each model of O_2 is a model of O_1. Instead of σ, an institution comorphism can be referred to. For example, we can express that the natural numbers are a total order as follows:

```
interpretation i : TotalOrder to Nat = Elem ↦ Nat
```

Here is a more complex example in Common Logic from the COLORE repository [9]:

```
interpretation geometry_of_time %mcons :
 %% Interpretation of linearly ordered time intervals...
  int:owltime_le
 %% ... that begin and end with an instant as lines
 %% that are incident with linearly ...
  to { ord:linear_ordering and bi:complete_graphical
     %% ... ordered points in a~special geometry, ...
       and  int:mappings/owltime_interval_reduction }
  = int:ProperInterval ↦ int:Interval end
```

The semantics of **interpretation** $N : O_1$ **to** $O_2 = \sigma$ is defined iff σ is a signature morphism from $\mathbf{Sign}(O_1)$ to $\mathbf{Sign}(O_2)$ such that for each $M_2 \in \mathbf{Mod}(O_2)$, $M_2|_\sigma \in \mathbf{Mod}(O_1)$. In that case, the graph of N is $(O_1, \mathbf{Inst}(O_1), \mathbf{Sign}(O_1), \mathbf{Mod}(O_1), \mathbf{Ax}(O_1)) \xrightarrow{\sigma} (O_2, \mathbf{Inst}(O_2), \mathbf{Sign}(O_2), \mathbf{Mod}(O_2), \mathbf{Ax}(O_2))$

3.3.2 Refinement

Refinements, written **refinement** $Id : O_1$ **to** $O_2 = \sigma$, expressing that O_2 is an acceptable realization of O_1. Semantically, this is equivalent with a theory interpretation from O_1 to O_2 along σ. Refinements can be combined using the **then** keyword, as in the example below, where the requirement of implementing a monoid is refined to implementing the monoid of natural numbers with addition, using the representation of numbers as lists of binary digits, for efficiency:

```
spec Monoid =
 sort Elem
 ops 0 : Elem;
         __+__ : Elem * Elem -> Elem, assoc, unit 0
end

spec NatWithSuc =
 free type Nat ::= 0 | suc(Nat)
 op __+__ : Nat * Nat -> Nat, unit 0
 forall x , y : Nat . x + suc(y) = suc(x + y)
 op 1:Nat = suc(0)
end

spec Nat =
  NatWithSuc hide suc
end

refinement R1 =
Monoid refined via Elem |-> Nat to Nat
end
```

```
spec NatBin =
generated type Bin ::= 0 | 1 | __0(Bin) | __1(Bin)

ops __+__ , __++__ : Bin * Bin -> Bin
forall x, y : Bin
 . 0 0 = 0  .  0 1 = 1
 . not  (0 = 1)  .  x 0 = y 0 => x = y
 . not  (x 0 = y 1)  .  x 1 = y 1 => x = y
 . 0 + 0 = 0  .  0 ++ 0 = 1
 . x 0 + y 0 = (x + y) 0  .  x 0 ++ y 0 = (x + y) 1
 . x 0 + y 1 = (x + y) 1  .  x 0 ++ y 1 = (x ++ y) 0
 . x 1 + y 0 = (x + y) 1  .  x 1 ++ y 0 = (x ++ y) 0
 . x 1 + y 1 = (x ++ y) 0  .  x 1 ++ y 1 = (x ++ y) 1
end

refinement R2 =
 Nat refined via Nat |-> Bin to NatBin
end

refinement R3 = R1 then R2
```

The semantics of **refinement** $R : O_1$ **to** $O_2 = \sigma$ is defined iff σ is a signature morphism from $\mathbf{Sign}(O_1)$ to $\mathbf{Sign}(O_2)$ such that for each $M_2 \in \mathbf{Mod}(O_2)$, $M_2|_\sigma \in \mathbf{Mod}(O_1)$. In that case, the graph of N is $(O_1, \mathbf{Inst}(O_1), \mathbf{Sign}(O_1), \mathbf{Mod}(O_1), \mathbf{Ax}(O_1)) \stackrel{\sigma}{\rightarrow} (O_2, \mathbf{Inst}(O_2), \mathbf{Sign}(O_2), \mathbf{Mod}(O_2), \mathbf{Ax}(O_2))$

The semantics of R_1 **then** R_2 is defined if and only if the semantics of R_1 is $(N_1, I_1, \Sigma_1, \mathcal{M}_1, \Delta_1) \stackrel{\sigma_1}{\rightarrow} (N_2, I_2, \Sigma_2, \mathcal{M}_2, \Delta_2)$, the semantics of R_2 is $(N_1', I_1', \Sigma_1', \mathcal{M}_1', \Delta_1') \stackrel{\sigma_2}{\rightarrow} (N_2', I_2', \Sigma_2', \mathcal{M}_2', \Delta_2')$, such that $I_2 = I_1'$, $\Sigma_2 = \Sigma_1'$ and $\mathcal{M}_1' \subseteq \mathcal{M}_2$, and then the graph of the composition is $(N_1'', I_1, \Sigma_1, \{M|_{\sigma_1;\sigma_2} \mid M \in \mathcal{M}_2'\}, \bot) \stackrel{\sigma_1;\sigma_2}{\rightarrow} (N_2', I_2', \Sigma_2', \mathcal{M}_2', \Delta_2')$, where N_1'' is a new name.

3.3.3 Entailment

Entailments, written **entailment** $Id = O_1$ **entails** O_2, express that O_2 is logically entailed by O_1. For example, we can express that in a group, the inverse of an element still exists after hiding the explicit inverse operation from the specification as follows:

```
logic CASL

spec InterpolatedGroup =
  sort Elem
  ops 0:Elem; __+__:Elem*Elem->Elem; inv:Elem->Elem
  forall x, y, z: Elem . x+0=x
                       . x+(y+z) = (x+y)+z
                       . x+inv(x) = 0
  forget inv
end

entailment ent = InterpolatedGroup
```

```
  entails { . forall x:Elem . exists y . Elem . x+y=0 }
end
```

The semantics of **entailment** $N = O_1$ **entails** O_2 is defined iff $\mathbf{Sign}(O_1) = \mathbf{Sign}(O_2)$ and $\mathbf{Mod}(O_1) \models \mathbf{Ax}(O_2)$. In that case, the graph of N is $(O_1, \mathbf{Inst}(O_1), \mathbf{Sign}(O_1), \mathbf{Mod}(O_1), \mathbf{Ax}(O_1)) \xrightarrow{id} (O_2, \mathbf{Inst}(O_2), \mathbf{Sign}(O_2), \mathbf{Mod}(O_2), \mathbf{Ax}(O_2))$

3.3.4 OMS Equivalence

OMS *equivalences*, written **equivalence** $Id : O_1 \leftrightarrow O_2 = O_3$, expressing that O_1 and O_2 have model classes that are in bijective correspondence. This is done by providing a (fragment) OMS O_3 such that O_i **then** O_3 is a definitional extension [27]. For example, Boolean algebras are equivalent to Boolean rings:

```
equivalence e : algebra:BooleanAlgebra ↔ algebra:BooleanRing =
forall x, y : Elem
. x ∧ y = x*y
. x ∨ y = x + y + x*y
. ¬x = 1 + x
. x*y = x ∧ y,
. x+y = (x ∨ y) ∧ ¬(x ∧ y).
end
```

The semantics of **equivalence** $N : O_1 \leftrightarrow O_2 = O_3$ is defined iff for each model $M_i \in \mathbf{Mod}(O_i)$ there exists a unique model $M \in \mathbf{Mod}_{(\mathbf{Sign}(O_1)\cup\mathbf{Sign}(O_2),\emptyset)}(O_3)$ such that $M|_{\mathbf{Sign}(O_i)} = M_i$. In that case, the graph of N is $(O_1, I, \mathbf{Sign}(O_1), \mathbf{Mod}(O_1), \mathbf{Ax}(O_1)) \xrightarrow{\iota_1} (O_3, I, \mathbf{Sign}_{(\mathbf{Sign}(O_1)\cup\mathbf{Sign}(O_2),\emptyset)}(O_3), \mathbf{Mod}_{(\mathbf{Sign}(O_1)\cup\mathbf{Sign}(O_2),\emptyset)}(O_3), \mathbf{Ax}_{(\mathbf{Sign}(O_1)\cup\mathbf{Sign}(O_2),\emptyset)}(O_3)) \xleftarrow{\iota_2} (O_2, I, \mathbf{Sign}(O_2), \mathbf{Mod}(O_2), \mathbf{Ax}(O_2))$ where ι_i are inclusions.

3.3.5 Conservative Extension

A *conservative extension* is written as **conservative extension** $Id\ c : O_1$ **of** O_2 **for** Σ. This expresses that O_2 contains all knowledge about the signature Σ from the O_1 or, more precisely, O_1 is a conservative extension of O_2 with restriction signature Σ and conservativity c. If c is %mcons, this means that every Σ-reduct of an O_2-model can be expanded to an O_1-model. If c is %ccons, this means that every Σ-sentence φ following from O_1 already follows from O_2. This relation shall hold for any module O_2 extracted from O_1 using the **extract** construct. For example, we can specify that we obtained a module of GALEN by extracting the parts corresponding to drugs, joints, and body parts as follows:

```
module myGalenIsAModule : myGalen of
  <http://example.org/GALEN/galen.owl>
  for Drugs, Joints, Bodyparts
end
```

The semantics of **conservative extension** $N\ c : O_1$ **of** O_2 **for** Σ is defined iff $\Sigma \subseteq \mathbf{Sign}(O_2) \subseteq \mathbf{Sign}(O_1)$ and if c =%mcons and for each $M \in \mathbf{Mod}(O_2)$ there is a model $M' \in \mathbf{Mod}(O_1)$ such that $M'|_\Sigma = M|_\Sigma$, or if c =%ccons and for each $\varphi \in \mathbf{Sen}(\Sigma)$, $O_1 \models \varphi$ implies $O_2 \models \varphi$. Then the graph of N is $(O_2, \mathbf{Inst}(O_2), \mathbf{Sign}(O_2), \mathbf{Mod}(O_2), \mathbf{Ax}(O_2)) \stackrel{\iota}{\to} (O_1, \mathbf{Inst}(O_1), \mathbf{Sign}(O_1), \mathbf{Mod}(O_1), \mathbf{Ax}(O_1))$, with ι being the inclusion.

3.3.6 Alignment Definition

Alignment definitions, written as **alignment** *Id* $card_1\ card_2 : O_1$ **to** $O_2 = c_1, \ldots, c_n$, **assuming** $domain$ where $card_1$ resp. $card_2$ specify constraints on the alignment relation concerning the source resp. target. Each $card_i$ is one of 1, ?, +, * ("1" for injective and total, '+' for total, "?" for injective and "*" for none). The c_j are correspondences of form $sym_1\ rel\ conf\ sym_2$. Here, sym_i is a symbol from O_i, rel is one of the built-in relations $>, <, =, \%, \ni, \in, \mapsto$, or an identifier of a relation specified externally, and *conf* is an (optional) confidence value between 0 and 1. The user can specify the assumption about the universe where the relations in the correspondences are interpreted using the $assuming$ clause, with possible values **SingleDomain** (all ontologies are interpreted over the same universe, which is also the default), **GlobalDomain** (the domains of the ontologies are reconciled w.r.t. a global domain of interpretation) and **ContextualizedDomain** (the domains are connected via relations). This syntax of alignments follows the Alignment API [10].[20] If all correspondences of an alignment have the confidence value 1, the alignment can be given a formal semantics as a network.

```
ontology Onto1 =
  Class: Person
  Class: Woman SubClassOf: Person
  Class: Bank
end

ontology Onto2 =
  Class: HumanBeing
  Class: Woman SubClassOf: HumanBeing
  Class: Bank
end

alignment VAlignment : Onto1 to Onto2 =
  Person = HumanBeing,
  Woman = Woman
end

network N =
 1 : Onto1,
 2 : Onto2,
 VAlignment
```

[20] Note that BioPortal's [40] mappings are correspondences in the sense of the Alignment API and hence of DOL. BioPortal only allows users to collect correspondences, but not to group them into alignments. In a sense, for each pair of ontologies, all BioPortal users contribute to a big alignment between these.

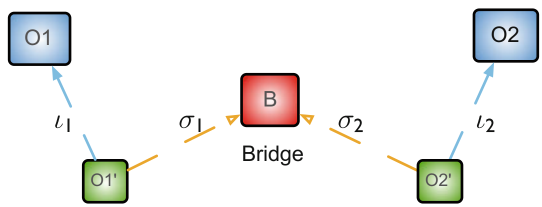


Fig. 2 Semantics of alignments

```
ontology VAlignedOntology =
  combine N
  %% 1:Person is identified with 2:HumanBeing
  %% 1:Woman is identified with 2:Woman
  %% 1:Bank and 2:Bank are kept distinct
end

ontology VAlignedOntologyRenamed =
  VAlignedOntology with 1:Bank ↦ RiverBank, 2:Bank ↦ FinancialBank,
                        Person_HumanBeing ↦ Person
end
```

We sketch the semantics of alignments with the case when the domain of interpretation is assumed to be shared by the ontologies being aligned. In this case, the semantics is given by a W-shaped graph like in Fig. 2 where O_1 and O_2 are the nodes of the ontologies being aligned, O_1' and O_2' collect the symbols of O_1 and O_2, respectively, that appear in the correspondences of the alignment, ι_1 and ι_2 are inclusions and the bridge ontology B together with the morphisms σ_1 and σ_2 is constructed by turning the correspondences into bridge axioms. Details can be found in [8].

3.4 Networks of OMS

OMS networks are introduced by the following grammar:

```
NetworkDefn := network NAME = Network
Network ::= NAME* [ excluding NAME* ]
```

Here, the NAMEs can name OMS, mappings or other networks. A network is specified as a list of network elements (OMS, OMS mappings, and subnetworks), followed by an optional list of excluded network elements. For disambiguating the symbols in the combination of the network, the individual OMS can be prefixed with labels, like $n : O$, which are scoped to the current OMS network. An example has already been presented in the section on alignments. Together with two OMS included in the network, the graph of the network implicitly includes all paths along importations between the two nodes. For example, in the example below, NAT_PLUS imports the specification NAT. Without the implicit inclusion of this import, the combination would duplicate the theory of NAT.

```
spec NAT =
  free type Nat ::= 0 | suc(Nat)
end

spec NAT_PLUS =
 NAT
  then
 op __+__ : Nat * Nat -> Nat
  ...
end

...

network N =
  NAT, NAT_PLUS,...
end

spec N_COMBINED =
  combine N
end
```

Formally, the graph of a network is constructed by taking the union of all graphs of its constituents, provided that we regard the semantics of OMS as a graph with one node and no edges, and removing from it all subparts specified in the **excluding** clause.

3.5 Libraries of OMS

Libraries start with the keyword **library** and the name of the library, followed by a *qualification* choosing the OMS *language*, *logic*, and/or *serialization*. This is followed by a list of imports of other libraries, definitions of OMS, OMS mappings, networks of OMS, or other qualifications which change the current logic. Optionally, a prefix map placed at the beginning of a library may be used to abbreviate IRIs. A library can also be the inclusion of an OMS written in a language corresponding to some institution.

```
%prefix(
         bfo: <https://bfo.googlecode.com/svn/releases/1.1.1/>
        )%

library Parthood

logic CommonLogic

ontology BFOWithAssociatedAxioms =
   bfo:bfo.owl with translation OWL22CommonLogic
  then
   (forall (x y) (if (snap:properTemporalPartOf x y)
                     (exists (z) (and (snap:properTemporalPartOf z y)
                     (not (exists (w)
```

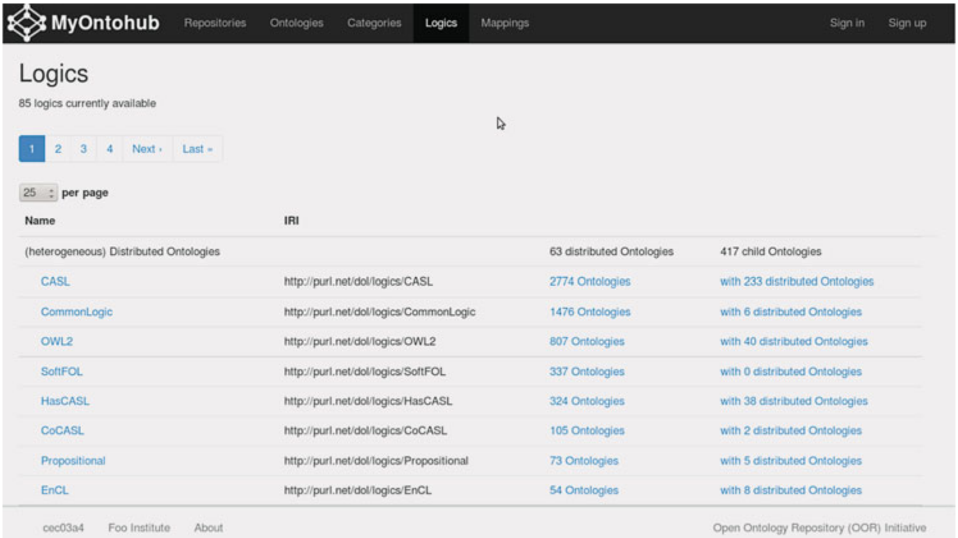


Fig. 3 Overview of logics in Ontohub

```
                                      (and (snap:temporalPartOf w x) (snap:
                                          temporalPartOf w z)
    )))))))
end
```

Note that the prefixes declared in an imported library are available in the imported library, as illustrated in the example above with the prefix `snap:`.

This completes our overview of DOL. The full syntax and semantics of DOL will be available at wiki.ontohub.org and has been submitted to OMG for standardization. The most recent version of the document is available at ontoiop.org.

4 Tool Support for DOL

Currently, DOL is supported by two tools: Ontohub and the Heterogeneous Tool Set (HETS). Ontohub (see http://ontohub.org) is a web-based repository engine for OMS that are written either in DOL or in some specific OMS language.[21]

Ontohub provides means for organizing OMS into repositories (Fig. 4). The distributed nature enables communities to share and exchange their contributions easily. The heterogeneous nature makes it possible to integrate OMS written in various OMS languages. Ontohub supports a wide range of DOL-conforming OMS languages building on DOL and also supports DOL's interpretations, equivalences and alignments. Users of Ontohub

[21] Ontohub's sources are freely available at https://github.com/ontohub/ontohub.

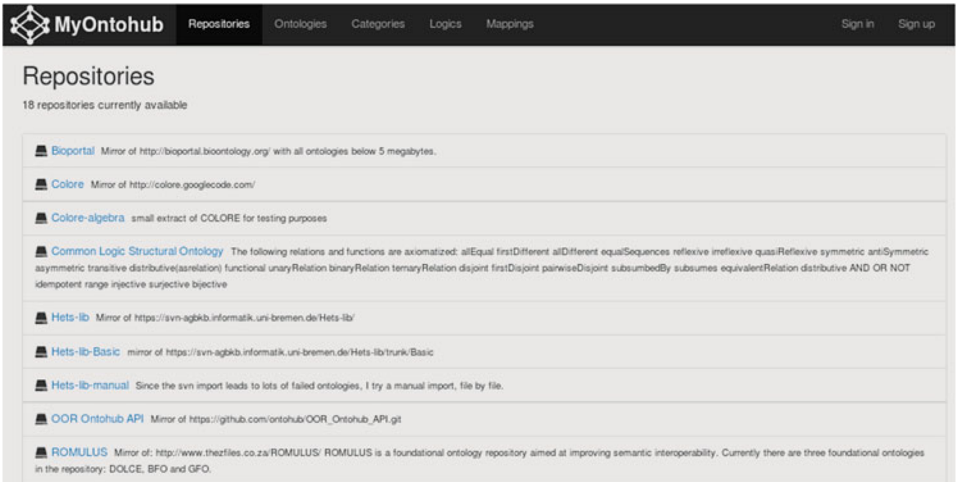


Fig. 4 Some of the repositories hosted on Ontohub

can upload, browse, search and annotate OMS and OMS libraries in various languages via a web front end. Figure 3 shows an excerpt of the 25 logics currently available in Ontohub.

The parsing and inference back end is the Heterogeneous Tool Set (Hets [34, 39], available at hets.eu). Hets supports a large number of basic OMS languages and logics, as well as the DOL metalanguage as described in this chapter.[22]

The structural information extracted from DOL OMS by Hets is stored in the Ontohub database and exposed to human users via a web interface and to machine clients as linked data.[23]

5 Conclusion and Future Work

Interoperability between systems as well as reusability, we argued in the introduction to this chapter, are critical challenges.

We here proposed to address these challenges by introducing two abstractions: firstly, we introduced the notion of OMS, spanning formalized ontologies, models, and specifications; secondly, we introduced the DOL language, an abstraction in the sense that it provides a structuring, module, and mapping language independently of the particular logical formalism used.

The work presented here brings together previous work pursued in a number of communities, including in particular logical pluralism, modular ontologies, algebraic specifi-

[22] Some (but only few) of DOL's features are still being implemented at the time of the writing of this chapter.

[23] "Linked data" is a set of best practises for publishing structured data on the Web in a machine-friendly way [1]. DOL and Ontohub conform with linked data.

cation, and modeling of systems. It therefore combines many isolated logical modeling and specification solutions into one coherent framework with formal semantics.

A number of open problems and challenges, however, remain:

- What is a suitable abstract meta framework for nonmonotonic logics and rule languages such as RIF and RuleML? Are institutions suitable here? Are the modularity questions for these languages different from those for monotonic logics?
- What is a useful abstract notion of OMS query (language)? How to handle answer substitutions in a logic-agnostic way?
- Can the notions of class hierarchy and of satisfiability of a class be generalized from OWL to other languages?
- Can logical frameworks be used for the specification of OMS languages and translations?

Despite these challenges, we hope that the development of DOL will have a profound impact on ontology engineering practices as well as on the way the modeling, ontology, and specification communities interact and how the systems they develop may interoperate. The impact on communities can already be seen e.g. by the use of Ontohub/DOL for the FOIS 2014 ontology competition. We have illustrated the benefits of DOL for a wide range of use cases; including for a framework of heterogeneous modeling in UML [20, 21], in biomedical ontology [26], for the specification of blending diagrams in computational creativity [23], and for the heterogeneous modeling of musical harmonies [6].

We hope that the future will bring many more diverse and interesting use cases for the DOL language.

Acknowledgment The development of DOL is supported by the German Research Foundation (DFG), Project I1-[OntoSpace] of the SFB/TR 8 "Spatial Cognition."

The project COINVENT acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open Grant number: 611553.

The authors would like to thank the OntoIOp working group for their valuable input, particularly Michael Grüninger, Maria Keet, Christoph Lange, and Peter Yim.

We also want to thank Yazmin Angelica Ibañez, Thomas Schneider and Carsten Lutz for valuable input on interpolation and module extraction.

References

1. Berners-Lee, T.: Design Issues: Linked Data. July 27, 2006. http://www.w3.org/DesignIssues/LinkedData.html. Accessed on 2010-01-20
2. Bidoit, M., Mosses, P.D.: CASL User Manual. LNCS (IFIP Series) 2900. Freely available at http://www.cofi.info. Springer, Berlin, Heidelberg (2004)
3. Bonatti, P.A., Lutz, C., Wolter, F.: The Complexity of Circumscription in DLs. J. Artif. Intell. Res. (JAIR) **35**, 717–773 (2009)
4. Borzyszkowski, T.: Logical systems for structured specifications. Theor. Comput. Sci. **286**, 197–245 (2002)

5. Burstall, R.M., Goguen, J.A.: The Semantics of CLEAR, A Specification Language. In Abstract Software Specifications, 1979 Copenhagen Winter School, January 22–February 2, 1979, Proceedings. Ed. by D. Bjørner. Vol. 86. Lecture Notes in Computer Science, pp. 292–332. Springer, Berlin, Heidelberg (1979). http://dx.doi.org/10.1007/3-540-10007-5_41
6. Cambouropoulos, E., Kaliakatsos-Papakostas, M., Kühnberger, K.-U., Kutz, O., Smaill, A.: Concept invention and music: Creating novel harmonies via conceptual blending. Proceedings of the 9th International Conference on Interdisciplinary Musicology (CIM-2014). Berlin (2014)
7. Codescu, M., Mossakowski, T.: Heterogeneous colimits. In: Boulanger, F., Gaston, C., Schobbens, P.-Y.: MoVaH'08 Workshop on Modeling, Validation and Heterogeneity. IEEE Press (2008). http://www.computer.org/portal/web/csdl/abs/proceedings/icstw/2008/3388/00/3388toc.htm
8. Codescu, M., Mossakowski, T., Kutz, O.: A Categorical Approach to Ontology Alignment. In: Proceedings of the 9th International Workshop on Ontology Matching (OM-2014), ISWC-2014, Riva del Garda, Trentino, Italy. CEUR-WS online proceedings (2014)
9. COLORE. An open repository of first-order ontologies represented in Common Logic. http://colore.googlecode.com
10. David, J., Euzenat, J., Scharffe, F., dos Santos, C.T.: The Alignment API 4.0. Semant. Web **2**(1), 3–10 (2011)
11. Diaconescu, R., Goguen, J., Stefaneas, P.: Logical support for modularisation. In: 2nd Workshop on Logical Environments, pp. 83–130. CUP, New York (1993)
12. Goguen, J.A., Burstall, R.M.: Institutions: Abstract Model Theory for Specification and Programming. J. Assoc. Comput. Mach. **39** (1992). Predecessor in: LNCS 164, 221–256 (1984), pp. 95–146
13. Goguen, J., Roşu, G.: Institution morphisms. Form. Asp. Comput. **13**, 274–307 (2002)
14. Goguen, J.A., Rosu, G.: Composing hidden information modules over inclusive institutions. In: Owe, O., Krogdahl, S., Lyche, T. (eds.) From Object-Orientation to Formal Methods, Essays in Memory of Ole-Johan Dahl, vol. 2635. Lecture Notes in Computer Science, pp. 96–123. Springer, Berlin, Heidelberg (2004). http://dx.doi.org/10.1007/978-3-540-39993-3_7
15. Grau, B.C., Honavar, V., Schlicht, A., Wolter, F. (eds.): Proceedings of the 2nd International Workshop on Modular Ontologies. WoMO 2007, Whistler, Canada, October 28, 2007, vol. 315. CEUR Workshop Proceedings. CEUR-WS.org (2008)
16. Haase, P., Honavar, V., Kutz, O., Sure, Y., Tamilin, A. (eds.): Proceedings of the 1st International Workshop on Modular Ontologies. WoMO'06, co-located with the International Semantic Web Conference, ISWC'06 November 5, 2006, Athens, Georgia, USA, vol. 232. CEUR Workshop Proceedings. CEUR-WS.org (2007)
17. The NeOn Ontology Engineering Toolkit. http://www.neon-project.org/ (2008). http://watson.kmi.open.ac.uk/Downloads%20and%20Publications_files/neon-toolkit.pdf
18. Hayes, P.: RDF Semantics. W3C Recommendation. World Wide Web Consortium (W3C), Feb. 10 (2004). http://www.w3.org/TR/2004/REC-rdf-mt-20040210/
19. Horridge, M.: Protégé OWL Tutorial. Version v1.3. Mar. 24 (2011). http://owl.cs.manchester.ac.uk/tutorials/protegeowltutorial/
20. Knapp, A., Mossakowski, T., Roggenbach, M.: Towards an institutional framework for heterogeneous formal development in UML – A position paper. In: Nicola, R.D., Hennicker, R. (eds.) Software, Services and Systems. Essays Dedicated to Martin Wirsing on the Occasion of His Emeritation, vol. 8950. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg (2015)
21. Knapp, A., Mossakowski, T., Roggenbach, M., Glauer, M.: An Institution for Simple UML State Machines. CoRR abs/1411.4495 (2014). http://arxiv.org/abs/1411.4495
22. Kontchakov, R., Wolter, F., Zakharyaschev, M.: Logic-based ontology comparison and module extraction, with an application to DL-Lite. Artif. Intell. **174**(15), 1093–1141 (2010). http://dx.doi.org/10.1016/j.artint.2010.06.003
23. Kutz, O., Bateman, J., Neuhaus, F., Mossakowski, T., Bhatt, M.: E pluribus unum: Formalisation, use-cases, computational support for conceptual blending. In: Besold, T.R., Schorlemmer, M., Smaill, A. (eds.) Computational Creativity Research: Towards Creative Machines. Thinking Machines. Springer, Atlantis (2014)

24. Kutz, O., Hois, J., Bao, J., Cuenca Grau, B. (eds.): Modular Ontologies – Proceedings of the Fourth International Workshop (WoMO 2010), vol. 210. Frontiers in Artificial Intelligence and Applications. IOS Press, Toronto, Canada (2010)
25. Kutz, O., Lücke, D., Mossakowski, T., Normann, I.: The OWL in the CASL – Designing Ontologies Across Logics. In: Dolbear, C., Ruttenberg, A., Sattler, U. (eds.) OWL: Experiences and Directions. 5th International Workshop (OWLED-08), co-located with ISWC-08, Karlsruhe, Germany, October 26–27: CEUR-WS, vol. 432 (2008)
26. Kutz, O., Mossakowski, T., Hastings, J., Castro, A.G., Sojic, A.: Hyperontology for the Biomedical Ontologist: A Sketch and Some Examples. In: Workshop on Working with Multiple Biomedical Ontologies (WoMBO at ICBO 2011). Buffalo, NY, USA (2011)
27. Kutz, O., Mossakowski, T., Lücke, D.: Carnap, Goguen, the Hyperontologies: Logical Pluralism and Heterogeneous Structuring in Ontology Design. Log. Univers. **4**, 2 (2010). Special issue on 'Is Logic Universal?'
28. Kutz, O., Schneider, T. (eds.): Modular Ontologies – Proceedings of the Fifth International Workshop (WoMO 2011), vol. 230. Frontiers in Artificial Intelligence and Applications. IOS Press, Amsterdam (2011)
29. Lifschitz, V.: Circumscription. In: Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 3, pp. 297–352. Oxford University Press, United Kingdom (1994)
30. Lüttich, K., Masolo, C., Borgo, S.: Development of Modular Ontologies in CASL. In: Haase, P., Honavar, V., Kutz, O., Sure, Y., Tamilin, A. (eds.) WoMO, vol. 232. CEUR Workshop Proceedings. CEUR-WS.org (2006)
31. Lutz, C., Seylan, I., Wolter, F.: An automata-theoretic approach to uniform interpolation and approximation in the description logic EL. In: Rewka, G., Eiter, T., McIlraith, S.A. (eds.) KR. AAAI Press (2012)
32. Lutz, C., Wolter, F.: Foundations for uniform interpolation and forgetting in expressive description logics. In: IJCAI, pp. 989–995. (2011)
33. Meseguer, J.: General logics. In: Ebbinghaus, H.J. (ed.) Logic Colloquium '87, pp. 275–329. North Holland (1989)
34. Mossakowski, T.: Hets: the Heterogeneous Tool Set. http://hets.eu. (Accessed 2012-12-10)
35. Mossakowski, T., Haxthausen, A., Sannella, D., Tarlecki, A.: CASL: The common algebraic specification language. In: Bjorner, M.H.D. (ed.) Logics of Formal Specification Languages. Monographs in Theoretical Computer Science, chap. 3, pp. 241–298. Springer, Heidelberg (2008). http://dx.doi.org/10.1007/978-3-540-74107-7_5
36. Mossakowski, T., Kutz, O.: The onto-logical translation graph. In: Kutz, O., Schneider. T. (eds.) Modular Ontologies. IOS, Amsterdam (2011)
37. Mossakowski, T., Kutz, O., Lange, C.: Semantics of the distributed ontology language: Institutes and institutions. In: Martí-Oliet, N., Palomino, M. (eds.) Recent Trends in Algebraic Development Techniques, 21th International Workshop, WADT 2012, vol. 7841. Lecture Notes in Computer Science, pp. 212–230. Springer, Berlin, Heidelberg (2013). http://link.springer.com/chapter/10.1007/978-3-642-37635-1_13
38. Mossakowski, T., Lange, C., Kutz, O.: Three Semantics for the Core of the Distributed Ontology Language. In: Donnelly, M., Guizzardi, G. (eds.) 7th International Conference on Formal Ontology in Information Systems (FOIS), vol. 239. Frontiers in Artificial Intelligence and Applications, pp. 337–352. FOIS Best Paper Award. IOS Press, Amsterdam (2012)
39. Mossakowski, T., Maeder, C., Lüttich, K.: The Heterogeneous Tool Set. In: Grumberg, O., Huth, M. (eds.) TACAS 2007, vol. 4424. Lecture Notes in Computer Science, pp. 519–522. Springer, Heidelberg (2007)
40. Noy, N.F., Shah, N.H., Whetzel, P.L., Dai, B., Dorf, M., Griffith, N., Jonquet, C., Rubin, D.L., Storey, M.-A., Chute, C.G., Musen, M.A.: BioPortal: ontologies and integrated data resources at the click of a mouse. Nucl. Acids Res. **37**, W170–W173 (2009). http://bioportal.bioontology.org
41. Sannella, D., Tarlecki, A.: Specifications in an arbitrary institution. Inf. Comput. **76**, 165–210 (1988)
42. Sannella, D., Tarlecki, A.: Foundations of Algebraic Specification and Formal Software Development. EATCS Monographs in Theoretical Computer Science. Springer, Berlin, Heidelberg (2012)

43. Sattler, U., Tamilin, A. (eds.): Workshop on Ontologies: Reasoning and Modularity (WORM-08), vol. 348. (ESWC) Tenerife, Spain: CEUR Workshop Proceedings (2008). http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-348/
44. Schneider, M., Rudolph, S., Sutcliffe, G.: Modeling in OWL 2 without Restrictions. In: Rodriguez-Muro, M., Jupp, S., Srinivas, K. (eds.) Proceedings of the 10th International Workshop on OWL: Experiences and Directions (OWLED 2013) co-located with 10th Extended Semantic Web Conference (ESWC 2013), Montpellier, France, May 26–27, 2013, vol. 1080. CEUR Workshop Proceedings. CEUR-WS.org (2013). http://ceur-ws.org/Vol-1080/owled2013_14.pdf
45. Schneider, T., Walther, D. (eds.): Proceedings of the 6h International Workshop on Modular Ontologies, vol. 875. CEUR-WS (2012)
46. Wang, Z., Wang, K., Topor, R.W., Pan, J.Z.: Forgetting for knowledge bases in DL-Lite. Ann. Math. Artif. Intell. **58**(1–2), 117–151 (2010)
47. Wirsing, M.: Structured Algebraic Specifications: A Kernel Language. Theor. Comput. Sci. **42**, 123–249 (1986). http://dx.doi.org/10.1016/0304-3975(86)90051-4
48. Zimmermann, A., Krötzsch, M., Euzenat, J., Hitzler, P.: Formalizing ontology alignment and its operations with category theory. In: Proceedings of FOIS-06, pp. 277–288 (2006)

T. Mossakowski (✉) · M. Codescu · F. Neuhaus
Institute of Knowledge and Language Engineering, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany
e-mail: mossakow@iws.cs.uni-magdeburg.de

O. Kutz
KRDB Research Centre for Knowledge and Data, Free University of Bozen-Bolzano, Bozen-Bolzano, Italy

Belnap Constants and Nelson Logic

Sergei Odintsov

Abstract This chapter aims to investigate connections between Belnap's useful four valued logic and Nelson's constructive logic with strong negation and the role of adding logical constants to the language. We consider the paraconsistent Nelson's logic and its expansions obtained by adding to it logical constant's corresponding to truth values of Belnap's useful four-valued logic.

Keywords Nelson's logic · Belnap's logic · Logical constant · Lattice of extensions · Twist-structure

Mathematics Subject Classification (2010) Primary 03F99 · Secondary 03B53

1 Introduction

This work continues the investigation started by Odintsov [18]. It is devoted to an investigation of connections between the well-known useful four-valued logic [2, 3] suggested by N. Belnap and Nelson's constructive logic with strong negation [12] and to the study of effects of adding new logical constants to the language of logic.

Nelson [12] suggested his system of constructive logic with strong negation to overcome the nonconstructive character of the intuitionistic negation. Subsequently, this system, traditionally denoted by **N3**, was axiomatized by Vorob'ev [27, 28] and studied algebraically by Helena Rasiowa [20, 21]. Fidel and Vakarelov [7, 26] independently discovered a convenient representation of algebraic **N3**-models via so-called twist-structures. The paraconsistent Nelson's logic **N4** is obtained by deleting the "explosive" axiom (Duns Scottus law) $\sim p \to (p \to q)$ from the axiomatics of **N3**. From the early 1970s several versions of **N4** were studied independently by Routley (later R. Sylvan) in the propositional case in [22], by López-Escobar in [11] and by Nelson himself in [1], both in the first-order case.

Kripke semantics for **N4** (see, e.g., [19]) is readily obtained from the usual Kripke semantics for intuitionistic logic by assigning to each world, instead of a set of propositional variables, a set of *literals*, i.e., propositional variables or strongly negated propositional variables. Equivalently, we can consider a truth-assignment for the logic **N4** as a mapping sending pairs of propositional variables and worlds into four-element Belnapian matrix

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_22

B4 with its truth values *True*, *False*, *Neither*, and *Both* [2, 3]. If we do not admit the truth value *Both*, we obtain a semantics for explosive Nelson's logic **N3** ([10, 25]). In this way, Nelson's logics **N3** and **N4** can be considered as "intuitionistic approximations" of three- and four-valued Belnap's logics.

Algebraic semantics for **N4** in terms of **N4**-lattices was suggested in [13], and the twist-structure representation of **N4**-lattices was studied in [14]. The logic $\mathbf{N4}^{\perp}$ [15] is obtained from **N4** by adjoining the intuitionistic falsity constant $\perp$. It is a conservative extension of **N4** as well as of the intuitionistic logic, but adding this constant provides the class of $\mathbf{N4}^{\perp}$-extensions with a regular structure (see [17]). In twist-structure semantics [7, 26], we have $\perp = (0, 1)$, where 0 and 1 are zero and unit elements of the underlying Heyting algebra, respectively. In this way $\perp$ is a natural analog of the Belnapian constant *False*, where as the ground term $\sim \perp = (1, 0)$ corresponds to the constant *True* [2, 3].

Another version of paraconsistent Nelson's logic arose in the course of the study of substructural properties of the strong implication connective, $p \Rightarrow q := (p \to q) \wedge (\sim q \to \sim p)$. This implication connective lacks the contraction property. The formulas $p \Rightarrow (p \Rightarrow q)$ and $p \Rightarrow q$ are not equivalent in **N3**. In [23, 24], it was proved syntactically, with the help of the prover OTTER, that the variety of **N3**-lattices (providing algebraic semantics for **N3**) is definitionally equivalent to the variety of Nelson $\mathbf{FL_{ew}}$-algebras (a subvariety of the variety of residuated lattices), and so **N3** can be treated as an axiomatic extension of $\mathbf{FL_{ew}}$, the full Lambeck calculus with exchange and weakening [8]. Busaniche and Cignoli [5] have found more natural (and essentially shorter) semantic proof of this result. Trying to extend this result to **N4** they defined in [4] the logic, which we denote here **bN4**[1], via adding the additional constant b satisfying the equations $b = \sim b$ and $b \to b = b$. The **bN4**-lattices are **N4**-lattices with the interpreted constant b. In twist-structures, the constant b is always interpreted as $(1, 1)$, and, in this way, it corresponds to the Belnapian constant *Both*. In [4], Busaniche and Cignoli defined a variety of NPc-lattices, a subvariety of the variety of residuated lattices with involution, and proved that this variety is equivalent to the variety of **bN4**-lattices. Note that the syntactical translations in this case are essentially different from that of [23].

In [18], we considered the logic $\mathbf{bN4}^{\perp}$, the language of which contains both constants $\perp$ and b and proved that despite the fact that the logic $\mathbf{bN4}^{\perp}$ is paraconsistent, its expressive power is very close to that of **N3**. More exactly, we defined the translation σ, which embed faithfully **N3** into $\mathbf{bN4}^{\perp}$, and the translation β providing the embedding of $\mathbf{bN4}^{\perp}$ into **N3**. Moreover, the translations σ and β will define mutually inverse isomorphisms of the lattice of **N3**-extensions and that of $\mathbf{bN4}^{\perp}$-extensions. This allows us to conclude that the expressive powers of the paraconsistent Nelson's logic $\mathbf{bN4}^{\perp}$ enriched with the Belnapian constant "Both" and of the explosive **N3** are essentially the same.

In this chapter, we consider the result of adding an analog of the Belnapian constant *Neither* to the logic $\mathbf{N4}^{\perp}$, and the result of adding the full set of Belnapian constants to **N4**. It turns out that the logic $\mathbf{nN4}^{\perp}$ obtained in this way is definitially equivalent to the logic $\mathbf{N4}^{\mathsf{N}}$, an important extension of $\mathbf{N4}^{\perp}$, via translations δ and γ. Moreover, the translations δ and γ define mutually inverse isomorpisms between lattices of $\mathbf{nN4}^{\perp}$- and

[1] In [4], this logic was denoted as **eN4**. We change the denotation to emphasize the connection between the constant b and the truth-value *Both*.

$\mathbf{N4}^{\mathsf{N}}$-extensions. The importance of the logic $\mathbf{N4}^{\mathsf{N}}$ is conditioned by the fact (see [17]) that it distinguishes in the lattice $\mathcal{E}\mathbf{N4}^{\perp}$ of $\mathbf{N4}^{\perp}$-extensions the class of so-called *normal* logics $\mathsf{Nor} = \mathcal{E}\mathbf{N4}^{\mathsf{N}}$. The logics from Nor represent in some sense the nontrivial structures of contradictions in paraconsistent extensions of the logic $\mathbf{N4}^{\perp}$ [17]. Studying the class of extensions of a paraconsistent logic it is important to distinguish the subclasses of explosive logics and of logics representing the behaviour of contradictions. In the case of $\mathbf{N4}^{\perp}$-extensions these are the lattices $\mathcal{E}\mathbf{N3}$ and $\mathcal{E}\mathbf{N4}^{\mathsf{N}}$.

Thus, we obtain the following situation. Without the constants $\perp$ and the ground term $\sim\perp$ corresponding to Belnapian truth values *True* and *False* we cannot distinguish the class of normal logics in $\mathcal{E}\mathbf{N4}$, but we can do it in the class of $\mathbf{N4}^{\perp}$-extensions. Adding the constant b corresponding to *Both* we restrict the class of $\mathbf{N4}^{\perp}$-extensions to the class of its explosive extensions, while adding the constant corresponding to *Neither* restricts the class of $\mathbf{N4}^{\perp}$-extensions to the class of normal logics. Finally, adding to the logic $\mathbf{N4}$ constants corresponding to all Belnapian truth values, we obtain the logic $\mathbf{N4}^{\mathrm{Bel}}$ whose class of extensions is syntactically isomorphic to that of $\mathbf{N3}^{\circ} = \mathbf{N3} + \mathbf{N4}^{\mathsf{N}}$, which is isomorphic in turn to the class of intermediate logics.

2 The Nelson Logics

We consider the basic propositional language $\mathcal{L}^{-} = \{\vee, \wedge, \to, \sim\}$ and several languages obtained from $\mathcal{L}$ by adjoining constant symbols:

$$\mathcal{L} = \mathcal{L}^{-} \cup \{\perp\}, \;\; \mathcal{L}^{b} = \mathcal{L} \cup \{b\}, \;\; \mathcal{L}^{n} = \mathcal{L} \cup \{n\}, \;\; \mathcal{L}^{\mathrm{Bel}} = \mathcal{L} \cup \{b, n\}\,.$$

The symbol $\sim$ stands for the *strong negation*, while the constant $\perp$ is the intuitionistic falsity constant and it will be used to define the intuitionistic negation. The set of formulas $\mathrm{For}(\mathcal{L}^{-})$ ($\mathrm{For}(\mathcal{L})$, etc.) is constructed from a fixed set $\mathrm{Prop} = \{p_0, p_1, \ldots\}$ of propositional variables with the help of connectives of the respective language in a standard way. For a formula φ, we denote by $\mathrm{var}(\varphi)$ the set of all propositional variables that occur in φ. We write $\varphi = \varphi(p_0, \ldots, p_n)$ if $\mathrm{var}(\varphi)$ is a subset of $\{p_0, \ldots, p_n\}$.

By a *logic* in the language $\mathcal{L}^{-}$ ($\mathcal{L}$, etc.) we mean a subset of $\mathrm{For}(\mathcal{L}^{-})$ ($\mathrm{For}(\mathcal{L})$, etc.) closed under the substitution rule and the rule of *modus ponens* (MP). A *theory* over a logic L is a set T of formulas of the respective language such that $L \subseteq T$ and T is closed under MP. A theory T is called inconsistent wrt $\sim$ (wrt $\perp$) if $\varphi, \sim\varphi \in T$ for some formula φ (if $\perp \in T$). A theory T over L is *trivial* if T contains all formulas of the respective language. Otherwise, the theory is called *nontrivial*.

Recall that a logic is called *paraconsistent* wrt $\sim$ (wrt $\perp$) if there is a nontrivial inconsistent wrt $\sim$ (wrt $\perp$) theory over that logic. But if such theory does not exist, then the logic is called *explosive* wrt $\sim$ (wrt $\perp$).

If L is a logic and Γ is a set of formulas in the language of L, we denote by $L + \Gamma$ the least logic containing $L \cup \Gamma$. Such logic exists since the intersection of an arbitrary family of logics is again a logic. If L is a logic in the language $\mathcal{L}^{-}$ ($\mathcal{L}$, etc.), we denote by $\mathcal{E}L$ the set of all logics in the language $\mathcal{L}^{-}$ ($\mathcal{L}$, etc.) extending L. The operations $\cap$ and $+$ turn $\mathcal{E}L$ into the lattice.

The paraconsistent Nelson logic **N4** can be defined as the least logic in the language $\mathcal{L}^-$ containing the following list of axioms:

A1. $p \to (q \to p)$
A2. $(p \to (q \to r)) \to ((p \to q) \to (p \to r))$
A3. $(p \wedge q) \to p$
A4. $(p \wedge q) \to q$
A5. $(p \to q) \to ((p \to r) \to (p \to (q \wedge r)))$
A6. $p \to (p \vee q)$
A7. $q \to (p \vee q)$
A8. $(p \to r) \to ((q \to r) \to ((p \vee q) \to r))$
A9. $\sim\sim p \leftrightarrow p$
A10. $\sim(p \vee q) \leftrightarrow (\sim p \wedge \sim q)$
A11. $\sim(p \wedge q) \leftrightarrow (\sim p \vee \sim q)$
A12. $\sim(p \to q) \leftrightarrow (p \wedge \sim q)$.

Recall that a formula φ is a negation normal form (nnf) if the strong negation $\sim$ occurs in φ only in front of propositional variables. The strong negation axioms A9–A12 allow us to reduce any formula to a negative normal form, i.e., to effectively find out for a formula $\varphi \in \mathrm{For}(\mathcal{L}^-)$, a nnf ψ such that $\varphi \leftrightarrow \psi \in$ **N4**. We say in this case that ψ is a negative normal form of φ. One can define the notion of nnf for every extension of the language $\mathcal{L}^-$ via constants. In this case, we assume the strong negation occur in nnf only in front of atomic formulas (propositional variables or constants). The above result on reducing to nnf holds for all logics considered in this chapter. Suppose that we fixed some algorithm reducing a formula φ to its nnf $\bar{\varphi}$.

The explosive Nelson logic **N3** can be defined as

$$\mathbf{N3} = \mathbf{N4} + \{\sim p \to (p \to q)\}\,.$$

The logic $\mathbf{N4}^\perp$ is the least logic in the language $\mathcal{L}$ containing axioms A1–A12 of **N4** and the additional axioms for the constant $\perp$:

A13. $\perp \to p$ and A14. $\sim\perp$

These axioms allow us to define the explosive intuitionistic negation in the usual way: $\neg\varphi := \varphi \to \perp$. More exactly, it can be proved that $\mathbf{N4}^\perp$ is a conservative extension of the intuitionistic logic defined in the language $\{\vee, \wedge, \to, \perp\}$. The logic $\mathbf{N4}^\perp$ is also a conservative extension of **N4**. Notice that $\mathbf{N4}^\perp$ is paraconsistent wrt $\sim$ and explosive wrt $\perp$.

If we define the constant $\perp$ as $\sim (p_0 \to p_0)$, then the axioms A13 and A14 are provable in **N3**. In this way, the logic **N3** can be considered as an extension of $\mathbf{N4}^\perp$. The following logics are important for defining the general structure of the lattice $\mathcal{E}\mathbf{N4}^\perp$:

$$\mathbf{N4}^{\mathsf{N}} = \mathbf{N4}^\perp + \{\neg\neg(p \vee \sim p)\}, \quad \mathbf{N3}^\circ = \mathbf{N4}^{\mathsf{N}} + \{\sim p \to (p \to q)\} = \mathbf{N4}^{\mathsf{N}} + \mathbf{N3}\,.$$

The logics $\mathbf{bN4}^{\perp}$, $\mathbf{nN4}^{\perp}$, and $\mathbf{N4}^{\mathrm{Bel}}$ are defined in the languages $\mathcal{L}^b$, $\mathcal{L}^n$, and respectively $\mathcal{L}^{\mathrm{Bel}}$ as follows:

$$\mathbf{bN4}^{\perp} = \mathbf{N4}^{\perp} + \{b, \sim b\}, \quad \mathbf{nN4}^{\perp} = \mathbf{N4}^{\perp} + \{n \to p, \sim n \to p\},$$
$$\mathbf{N4}^{\mathrm{Bel}} = \mathbf{nN4}^{\perp} + \{b, \sim b\} = \mathbf{bN4}^{\perp} + \mathbf{nN4}^{\perp}.$$

3 Algebraic Semantics for Nelson's Logics

Recall that an *implicative lattice* $\mathcal{A} = \langle A, \wedge, \vee, \to, 1\rangle$ is a distributive lattice with the greatest element 1, and the implication operation $\to$ satisfying the equivalence

$$a \wedge c \leq b \text{ iff } c \leq a \to b,$$

where $\leq$ is the lattice ordering of $\mathcal{A}$ and $a, b, c \in A$. A Heyting algebra $\mathcal{A} = \langle A, \wedge, \vee, \to, 0, 1\rangle$ is an implicative lattice with the least element 0.

The set $F_{\mathrm{d}}(\mathcal{A}) = \{a \vee (a \to b) \mid a, b \in A\}$ is a *filter of dense elements* of an implicative lattice (a Heyting algebra) $\mathcal{A}$. The lattice of filters on $\mathcal{A}$ is denoted by $\mathcal{F}(A)$, and the lattice of ideals is denoted by $\mathcal{I}(A)$. In the case of a Heyting algebra $\mathcal{A}$ we have: $F_{\mathrm{d}}(\mathcal{A}) = \{a \in A \mid \neg\neg a = 1\} = \{a \vee \neg a \mid a \in A\}$.

Definition 3.1 Let $\mathcal{A}$ be an implicative lattice (a Heyting algebra).

1. A full twist-structure over $\mathcal{A}$ is an algebra

$$\mathcal{A}^{\bowtie} = \langle A \times A, \vee, \wedge, \to, \sim\rangle$$
$$(\mathcal{A}^{\bowtie} = \langle A \times A, \vee, \wedge, \to, \sim, \perp, 1\rangle)$$

   with twist-operations defined for $(a, b), (c, d) \in A \times A$ in the following way:

$$(a, b) \vee (c, d) := (a \vee c, b \wedge d), (a, b) \wedge (c, d) := (a \wedge c, b \vee d)$$
$$(a, b) \to (c, d) := (a \to c, a \wedge d), \sim (a, b) := (b, a)$$
$$(\perp := (0, 1), 1 := (1, 0)).$$

2. A twist-structure over $\mathcal{A}$ is an arbitrary subalgebra $\mathcal{B}$ of the full twist-structure $\mathcal{A}^{\bowtie}$ such that $\pi_1(\mathcal{B}) = \mathcal{A}$, where π_i, $i = 1, 2$, denotes the projection of the direct product on the ith coordinate.
3. A class of all twist-structures over the algebra $\mathcal{A}$ is denoted by $S^{\bowtie}(\mathcal{A})$.

Notice that if $\mathcal{B} \in S^{\bowtie}(\mathcal{A})$, then we have $\pi_2(\mathcal{B}) = \mathcal{A}$.

A valuation in a twist-structure $\mathcal{B}$ is defined in a usual way as a homomorphism of the algebra of formulas into $\mathcal{B}$. The relation $\mathcal{B} \models \varphi$ means that $\pi_1 \upsilon(\varphi) = 1$ for every $\mathcal{B}$-valuation υ. We write $\mathcal{B} \models \Gamma$ for a set of formulas Γ if $\mathcal{B} \models \varphi$ for all $\varphi \in \Gamma$. For a formula φ, the notation $\models^{-}_{\bowtie} \varphi$ ($\models_{\bowtie} \varphi$) means that $\mathcal{B} \models \varphi$ for each twist-structure $\mathcal{B}$ over an implicative lattice (a Heyting algebra).

Theorem 3.2 [13, 15] For any $\varphi \in \mathrm{For}(\mathcal{L}^-)$ ($\varphi \in \mathrm{For}(\mathcal{L})$) the following equivalence is true:

$$\varphi \in \mathbf{N4} \;\Leftrightarrow \models^-_{\bowtie} \varphi \quad (\varphi \in \mathbf{N4}^\perp \;\Leftrightarrow \models_{\bowtie} \varphi)\,.$$

Let $\mathcal{A}$ be an implicative lattice (a Heyting algebra). Further, let $\nabla \in \mathcal{F}(\mathcal{A})$, $F_d(\mathcal{A}) \subseteq \nabla$, $\Delta \in \mathcal{I}(\mathcal{A})$. We denote by $Tw(\mathcal{A}, \nabla, \Delta)$ the twist-structure over $\mathcal{A}$ with the universe

$$|Tw(\mathcal{A}, \nabla, \Delta)| = \{(a, b) \mid a, b \in A, a \vee b \in \nabla, a \wedge b \in \Delta\}\,.$$

It is easy to see that this set is closed under twist-operations. Moreover (see [14]), every element $\mathcal{B}$ of $S^{\bowtie}(\mathcal{A})$ can be represented as $\mathcal{B} = Tw(\mathcal{A}, \nabla, \Delta)$, where

$$\nabla = \{a \vee b \mid (a, b) \in \mathcal{B}\} \text{ and } \Delta = \{a \wedge b \mid (a, b) \in \mathcal{B}\}\,.$$

In particular, $\mathcal{A}^{\bowtie} = Tw(\mathcal{A}, \mathcal{A}, \mathcal{A})$. In [15] it was noticed that a twist-structure $\mathcal{B} = Tw(\mathcal{A}, \nabla, \Delta)$ is a model of the logic **N3**, $\mathcal{B} \models \mathbf{N3}$, iff the implicative lattice $\mathcal{A}$ contains the least element 0 and $\Delta = \{0\}$. Further, a twist-structure $\mathcal{B} = Tw(\mathcal{A}, \nabla, \Delta)$ over a Heyting algebra is a model of the logic $\mathbf{N4}^{\mathsf{N}}$ iff $\nabla = F_d(\mathcal{A})$. Thus, the logic **N3** is characterized by the class of twist-structures of the form $Tw(\mathcal{A}, \nabla, \{0\})$, the logic $\mathbf{N4}^{\mathsf{N}}$ is determined by the class of twist-structures of the form $Tw(\mathcal{A}, F_d(\mathcal{A}), \Delta)$. Finally, $\mathbf{N3}^\circ$ is characterized by twist-structures $Tw(\mathcal{A}, F_d(\mathcal{A}), \{0\})$.

An algebra $\mathcal{A} = \langle A, \vee, \wedge, \to, \sim\rangle$ ($\mathcal{A} = \langle A, \vee, \wedge, \to, \sim, \perp, 1\rangle$) isomorphic to a twist-structure over an implicative lattice (a Heyting algebra) is called an **N4**-*lattice* (an $\mathbf{N4}^\perp$-*lattice*). We say that φ is *true on* an $\mathbf{N4}^{(\perp)}$-lattice $\mathcal{A}$, $\mathcal{A} \models \varphi$, if the identity $\varphi \to \varphi = \varphi$ holds on $\mathcal{A}$. This definition agrees with the definition of validity of formulas on twist-structures. Namely, if $\mathcal{B}$ is a twist-structure, then $\mathcal{B} \models \varphi$ iff the identity $\varphi \to \varphi = \varphi$ holds on $\mathcal{B}$. We say that an $\mathbf{N4}^{(\perp)}$-lattice $\mathcal{A}$ is a model of a logic $L \in \mathcal{E}\mathbf{N4}^{(\perp)}$ ($\mathcal{A} \models L$) if $\mathcal{A} \models \varphi$ for all $\varphi \in L$.

In [13, 15], it was proved that the class of all $\mathbf{N4}^{(\perp)}$-lattices forms a variety $\mathcal{N}4^{(\perp)}$, and that the lattice $\mathrm{Sub}(\mathcal{N}4^{(\perp)})$ of subvarieties of $\mathcal{N}4^{(\perp)}$ is dually[2] isomorphic to the lattice $\mathcal{E}\mathbf{N4}^{(\perp)}$ of $\mathbf{N4}^{(\perp)}$-extensions. These mutually inverse dual isomorphisms are given by the rules:

$$V(L) = \{\mathcal{A} \mid \mathcal{A} \in \mathcal{N}4^{(\perp)}, \mathcal{A} \models L\}, \;\; L \in \mathcal{E}\mathbf{N4}^{(\perp)};$$
$$L(V) = \{\varphi \mid \mathcal{A} \models \varphi \text{ for all } \mathcal{A} \in V\}, \;\; V \in \mathrm{Sub}(\mathcal{N}4^{(\perp)})\,.$$

It follows from the above remarks on **N3**-models that for every **N4**-lattice $\mathcal{A}$ the following equivalence holds:

$$\mathcal{A} \in V(\mathbf{N3}) \text{ iff } \mathcal{A} \cong Tw(\mathcal{B}, \nabla, \{0\}) \text{ for some } \mathcal{B} \text{ and } \nabla \supseteq F_d(\mathcal{B})\,. \tag{3.1}$$

[2] By a dual isomorphism of lattices $\mathcal{A}$ and $\mathcal{B}$ we mean a mapping $h : \mathcal{A} \to \mathcal{B}$ such that h is an isomorphism of $\mathcal{A}$ and $\mathcal{B}^{op}$, where $\mathcal{B}^{op}$ is the lattice with the same support as $\mathcal{B}$, but with the inverse ordering.

Similarly, for every $\mathbf{N4}^{\perp}$-lattice, the following holds:

$$\mathcal{A} \in V(\mathbf{N4}^{\mathrm{N}}) \text{ iff } \mathcal{A} \cong Tw(\mathcal{B}, F_{\mathrm{d}}(\mathcal{B}), \Delta) \text{ for some } \mathcal{B} \text{ and } \Delta \in \mathcal{I}(\mathcal{B})\,; \tag{3.2}$$

$$\mathcal{A} \in V(\mathbf{N3}^{\circ}) \text{ iff } \mathcal{A} \cong Tw(\mathcal{B}, F_{\mathrm{d}}(\mathcal{B}), \{0\}) \text{ for some } \mathcal{B}\,. \tag{3.3}$$

An algebra $\mathcal{A} = \langle A, \vee, \wedge, \rightarrow, \sim, b, \perp, 1\rangle$ is called a $\mathbf{bN4}^{\perp}$*-lattice* iff the b-free reduct $\mathcal{A} = \langle A, \vee, \wedge, \rightarrow, \sim, \perp, 1\rangle$ is an $\mathbf{N4}^{\perp}$-lattice and the following equations hold: $b \rightarrow b = b$ and $\sim b = b$.

An algebra $\mathcal{A} = \langle A, \vee, \wedge, \rightarrow, \sim, n, \perp, 1\rangle$ is an $\mathbf{nN4}^{\perp}$*-lattice* iff the n-free reduct $\mathcal{A} = \langle A, \vee, \wedge, \rightarrow, \sim, \perp, 1\rangle$ is an $\mathbf{N4}^{\perp}$-lattice and the following equations hold: $\neg n \rightarrow \neg n = \neg n$ and $\sim n = n$. Finally, an $\mathbf{N4}^{\mathrm{Bel}}$-lattice $\mathcal{A} = \langle A, \vee, \wedge, \rightarrow, \sim, b, n, \perp, 1\rangle$ is an $\mathbf{N4}^{\perp}$-lattice with two additional constants satisfying the above identities.

The classes of $\mathbf{bN4}^{\perp}$-lattices, $\mathbf{nN4}^{\perp}$-lattices, and $\mathbf{N4}^{\mathrm{Bel}}$-lattices form varieties, which are defined by the identities of $\mathbf{bN4}^{\perp}$-lattices and the identities defining the new constants b or (and) n. We denote these varieties as $b\mathcal{N}4^{\perp}$, $n\mathcal{N}4^{\perp}$, and $\mathcal{N}4^{\mathrm{Bel}}$, respectively.

Earlier we have defined mappings V and L establishing mutually inverse dual lattice isomorphisms between the lattice of $\mathbf{N4}^{\perp}$-extensions and the lattice of subvarieties of the variety of $\mathbf{N4}^{\perp}$-lattices. In exactly the same way we define the following mappings:

$$V : \mathrm{Sub}(b\mathcal{N}4^{\perp}) \rightarrow \mathcal{E}\mathbf{bN4}^{\perp} \quad \text{and} \quad L : \mathcal{E}\mathbf{bN4}^{\perp} \rightarrow \mathrm{Sub}(b\mathcal{N}4^{\perp})$$
$$V : \mathrm{Sub}(n\mathcal{N}4^{\perp}) \rightarrow \mathcal{E}\mathbf{nN4}^{\perp} \quad \text{and} \quad L : \mathcal{E}\mathbf{nN4}^{\perp} \rightarrow \mathrm{Sub}(n\mathcal{N}4^{\perp})$$
$$V : \mathrm{Sub}(\mathcal{N}4^{\mathrm{Bel}}) \rightarrow \mathcal{E}\mathbf{N4}^{\mathrm{Bel}} \quad \text{and} \quad L : \mathcal{E}\mathbf{N4}^{\mathrm{Bel}} \rightarrow \mathrm{Sub}(\mathcal{N}4^{\mathrm{Bel}})\,.$$

We have only slightly modify the proofs from [15] to establish that we obtained mutually inverse dual lattice isomorphisms between the lattice $\mathrm{Sub}(b\mathcal{N}4^{\perp})$ of subvarieties of the variety $b\mathcal{N}4^{\perp}$ and the lattice of $\mathbf{bN4}^{\perp}$-extensions, etc.

Proposition 3.3

1. [18] An algebra $\mathcal{A} = \langle A, \vee, \wedge, \rightarrow, \sim, b, \perp, 1\rangle$ is an $\mathbf{bN4}^{\perp}$-lattice iff there are a Heyting algebra $\mathcal{B}$ and a filter ∇ on $\mathcal{B}$ ($F_{\mathrm{d}}(\mathcal{B}) \subseteq \nabla$) such that $\mathcal{A} \cong Tw(\mathcal{B}, \nabla, \mathcal{B})$. The constant b corresponds to $(1, 1)$ under this isomorphism.
2. An algebra $\mathcal{A} = \langle A, \vee, \wedge, \rightarrow, \sim, n, \perp, 1\rangle$ is an $\mathbf{nN4}^{\perp}$-lattice iff there are a Heyting algebra $\mathcal{B}$ and an ideal Δ on $\mathcal{B}$ such that $\mathcal{A} \cong Tw(\mathcal{B}, \mathcal{B}, \Delta)$. The constant n corresponds to $(0, 0)$ under this isomorphism.
3. An algebra $\mathcal{A} = \langle A, \vee, \wedge, \rightarrow, \sim, n, b, \perp, 1\rangle$ is an $\mathbf{N4}^{\mathrm{Bel}}$-lattice iff there is a Heyting algebra $\mathcal{B}$ such that $\mathcal{A} \cong \mathcal{B}^{\bowtie} \cong Tw(\mathcal{B}, \mathcal{B}, \mathcal{B})$. The constants b and n correspond under this isomorphism to $(1, 1)$ and $(0, 0)$ respectively.

Proof Item 1 was established in [18].

2. Let $\mathcal{A}$ be an $\mathbf{nN4}^{\perp}$-lattice. In particular, $\mathcal{A}$ is an $\mathbf{N4}^{\perp}$-lattice. Consequently, there are a Heyting algebra $\mathcal{B}$, a filter ∇ and an ideal Δ on $\mathcal{B}$ such that $\mathcal{A} \cong Tw(\mathcal{B}, \nabla, \Delta)$. Let n be interpreted in $\mathcal{B}$ as (a, b). The equality $\sim n = n$ implies $(b, a) = (a, b)$, i.e., $a = b$. The equality $\neg n \rightarrow \neg n = \neg n$ means that $\neg(a, b) \rightarrow \neg(a, b) = (\neg a, a) \rightarrow (\neg a, a) =$

$(\neg a \to \neg a, \neg a \wedge a) = (1,0) = (\neg a, a) = \neg(a,b)$, i.e., $a = 0$. Thus, n is interpreted in $\mathcal{B}$ as $(0,0)$, in particular, $(0,0) \in \mathcal{B}$. Recall that $\nabla = \{a \vee b \mid (a,b) \in Tw(\mathcal{B}, \nabla, \Delta)\}$. From $0 = 0 \vee 0$ we obtain $0 \in \nabla$, which means that $\nabla = \mathcal{B}$. On the other hand, if $\mathcal{A} \cong Tw(\mathcal{B}, \mathcal{B}, \Delta)$, then $(0,0) \in Tw(\mathcal{B}, \mathcal{B}, \Delta)$ since $0 = 0 \wedge 0$ belongs to every ideal Δ on C. Thus, interpreting n as $(0,0)$ turns $\mathcal{B}$ into an $\mathbf{nN4}^{(\perp)}$-lattice.

3. This item follows from 1 and 2. □

Now we know how the constants b and n are interpreted in models of logics $\mathbf{bN4}^{\perp}$, $\mathbf{nN4}^{\perp}$, and $\mathbf{N4}^{\mathrm{Bel}}$, and this is the right place to emphasize the connection of these constats with the truth-values *Both* and *Neither* in Belnap's four-valued logic.

N. Belnap [2, 3, 6] suggested a logical system based on the following four-valued matrix

$$\mathbf{B4} := \langle \{True, False, Neither, Both\}, \wedge, \vee, \sim, \{True, Both\} \rangle .$$

Let $\mathbf{4} = \{True, False, Neither, Both\}$. The intended interpretations of elements of $\mathbf{4}$ are subsets of the set $\{0,1\}$ of classical truth values:

$$True = \{1\},\ False = \{0\},\ Neither = \emptyset,\ Both = \{0,1\} .$$

The matrix operations are then operations on the subsets of the set of classical truth values, e.g.,

$$\{0,1\} \vee \{0\} = \{0,1\}, \{1\} \wedge \emptyset = \emptyset, \sim \{0,1\} = \{0,1\},$$

where it is possible to calculate by classical considerations; otherwise the value is determined by monotonicity wrt the *truth ordering* (see, e.g., [2]).

We can represent the truth values as characteristic functions of subsets of $\{0,1\}$, i.e., as pairs (a,b), where $a,b \in \{0,1\}$, $a = 1$ iff 1 belongs to the subset, and $b = 1$ iff 0 belongs to the subset:

$$True = (1,0),\ False = (0,1),\ Neither = (0,0),\ Both = (1,1).$$

Under such a presentation, the matrix operations of $\mathbf{B4}$ have the form of twist-operations:

$$(a,b) \vee (c,d) = (a \vee c, b \wedge d),\ \ (a,b) \wedge (c,d) = (a \wedge c, b \vee d),\ \ \sim (a,b) = (b,a).$$

This presentation suggests a natural way of defining an implication operation on $\mathbf{B4}$:

$$(a,b) \to (c,d) = (a \to c, a \wedge d).$$

We also add to the language the constant $\perp$ interpreted as *False* and will consider Belnap's matrix in this extended language:

$$\mathbf{B4} := \langle \mathbf{4}, \wedge, \vee, \to, \perp, \sim, \{\mathbf{T}, \mathbf{B}\} \rangle.$$

Obviously, the algebra $\langle \mathbf{4}, \wedge, \vee, \to, \bot, \sim\rangle$ coincides with $\mathbf{2}^{\bowtie}$, where $\mathbf{2}$ is a two-element Boolean algebra. The set of distinguished truth values of **B4** contains exactly all elements a of $\mathbf{2}^{\bowtie}$ with $\pi_1 a = 1$. Thus, the validity of a formula on the matrix **B4** is equivalent to its validity on the twist-structure $\mathbf{2}^{\bowtie}$.

The twist-structure $\mathbf{2}^{\bowtie}$ is not necessarily a subalgebra of an $\mathbf{N4}^{\bot}$-lattice $\mathcal{A}$. A twist-structure presentation of $\mathcal{A}$ must contain $(0,1) = \bot$ and $(1,0) =\sim \bot$, where as *Neither* $= (0,0)$ and *Both* $= (1,1)$ are optional. Only if we consider $\mathbf{2}^{\bowtie}$ as an $\mathbf{N4}^{\mathrm{Bel}}$-lattice with interpreted constants b and n, it becomes a subalgebra of any $\mathbf{N4}^{\mathrm{Bel}}$-lattice. Due to this reason we consider $\mathbf{N4}^{\mathrm{Bel}}$ as Nelson's logic with the full set of Belnapian constants. The logic $\mathbf{bN4}^{\bot}$ is a Nelson's logic with *Both*, etc.

Concluding this section, we make a few remarks on quotients of $\mathbf{N4}^{\bot}$-lattices. Let $\mathcal{A}$ be an $\mathbf{N4}^{\bot}$-lattice. A nonempty subset $\nabla \subseteq \mathcal{A}$ is called a *special filter of the first kind* (sffk) on $\mathcal{A}$ if: 1) $a \in \nabla$ and $b \in \nabla$ imply $a \wedge b \in \nabla$; 2) $a \in \nabla$ and $(a \to b) \to (a \to b) = a \to b$ imply $b \in \nabla$. It is obvious that the set of all sffk on $\mathcal{A}$ forms a lattice, which we denote $\mathcal{F}^1(\mathcal{A})$.

Denote by $\mathrm{Con}(\mathcal{A})$ the lattice of congruences on $\mathcal{A}$. It was proved in [14, 15] that the mappings $\nabla \mapsto \theta_\nabla$, $\nabla \in \mathcal{F}^1(\mathcal{A})$, and $\theta \mapsto \nabla_\theta$, $\theta \in \mathrm{Con}(\mathcal{A})$,where

$$\theta_\nabla = \{(a,b) \mid a, b \in \mathcal{A},\ a \leftrightarrow b, \sim a \leftrightarrow \sim b \in \nabla\}\,,$$
$$\nabla_\theta = \{a \mid a \in \mathcal{A},\ a\theta(a \to a)\}\,,$$

establish mutually inverse isomorphisms between lattices $\mathcal{F}^1(\mathcal{A})$ and $\mathrm{Con}(\mathcal{A})$.

Let $\mathcal{A} \in S^{\bowtie}(\mathcal{B})$. For $F \subseteq \mathcal{B}$, put $F^{\bowtie} = \{(a,b) \mid (a,b) \in \mathcal{B}, a \in F\}$. It was proved in [15] that $F \in \mathcal{F}(\mathcal{B})$ iff $F^{\bowtie} \in \mathcal{F}^1(\mathcal{A})$. Moreover, if $F \in \mathcal{F}(\mathcal{B})$, then for every $\nabla \in \mathcal{F}(\mathcal{B})$ with $F_{\mathrm{d}}(\mathcal{B}) \subseteq \nabla$ and $\Delta \in \mathcal{I}(\mathcal{B})$, we have the following isomorphism:

$$Tw(\mathcal{B}, \nabla, \Delta)/F^{\bowtie} \cong Tw(\mathcal{B}/F, \nabla/F, \Delta/F)\,. \tag{3.4}$$

The result concerning sffk can be immediately transferred to $\mathbf{bN4}^{\bot}$-, $\mathbf{nN4}^{\bot}$-, and $\mathbf{N4}^{\mathrm{Bel}}$-lattices.

4 The Lattices $\mathcal{E}\mathbf{bN4}^{\bot}$, $\mathcal{E}\mathbf{nN4}^{\bot}$, and $\mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$

First we survey the results from [18] concerning the lattice of $\mathbf{bN4}^{\bot}$-extensions.

The translation $\sigma : \mathrm{For}(\mathcal{L}) \to \mathrm{For}(\mathcal{L}^b)$ is given by the rule

$$\sigma(\varphi) := \left(\bigwedge_{p \in \mathrm{var}(\varphi)} \neg(p \wedge \sim p) \right) \to \varphi\,.$$

For arbitrary $L \in \mathcal{E}\mathbf{N3}$, put

$$\sigma(L) = \mathbf{bN4}^{\bot} + \{\sigma(\varphi) \mid \varphi \in L\}.$$

Recall that a logic L_1 is *faithfully embedded into* L_2 *via* a translation τ if for every formula φ in the language of L_1, we have the equivalence: $\varphi \in L_1$ iff $\tau(\varphi) \in L_2$.

Theorem 4.1 [18] For every logic $L \in \mathcal{E}\mathbf{N3}$ and formula $\varphi \in \mathrm{For}(\mathcal{L})$, the following equivalence holds:

$$\varphi \in L \iff \sigma(\varphi) \in \sigma(L).$$

In other words, every **N3**-extension L is faithfully embedded into a $\mathbf{bN4}^{\perp}$-extension via the same syntactic translation. In particular, we have

Corollary 4.2 [18] The logic **N3** is faithfully embedded into $\mathbf{bN4}^{\perp}$ via σ.

The proof of Theorem 4.1 is based on the following observation.

Lemma 4.3 [18] For every formula $\varphi \in \mathrm{For}(\mathcal{L})$, Heyting algebra $\mathcal{A}$, and filter ∇ on $\mathcal{A}$ such that $F_{\mathrm{d}}(\mathcal{A}) \subseteq \nabla$, the following equivalence holds:

$$Tw(\mathcal{A}, \nabla, \mathcal{A}) \vDash \sigma(\varphi) \text{ iff } Tw(\mathcal{A}, \nabla, \{0\}) \vDash \varphi.$$

In fact, the mapping $L \mapsto \sigma(L)$ establishes an isomorphism between lattices $\mathcal{E}\mathbf{N3}$ and $\mathcal{E}\mathbf{bN4}^{\perp}$. The proof of this fact is based on the following translation β: $\mathrm{For}(\mathcal{L}^b) \to \mathrm{For}(\mathcal{L})$.

First we define the translation $\beta : \mathrm{For}(\mathcal{L}^b) \to \mathrm{For}(\mathcal{L})$ for formulas that are nnfs. If $\varphi = \varphi(p_0, \ldots, p_n)$ is a nnf, then there is a uniquely defined positive formula φ', i.e., a formula constructed from propositional variables with the help of positive connectives $\vee$, $\wedge$, and $\rightarrow$, such that

$$\varphi(p_0, \ldots, p_n) = \varphi'(\bot, \sim\bot, b, \sim b, p_0, \sim p_0, \ldots, p_n, \sim p_n).$$

We put

$$\beta(\varphi(p_0, \ldots, p_n)) = \varphi'(\bot, \sim\bot, \sim\bot, \sim\bot, p_0, p_1 \vee \sim p_0, \ldots, p_{2n}, p_{2n+1} \vee \sim p_{2n}).$$

In other words, for nnfs the translation β is defined as follows:

$$\beta(\bot) = \bot, \quad \beta(\sim\bot) = \sim\bot, \quad \beta(b) = \sim\bot, \quad \beta(\sim b) = \sim\bot,$$
$$\beta(p_i) := p_{2i}, \quad \beta(\sim p_i) = p_{2i+1} \vee \sim p_{2i},$$
$$\beta(\varphi \diamond \psi) := \beta(\varphi) \diamond \beta(\psi), \quad \diamond \in \{\vee, \wedge, \rightarrow\}.$$

For a formula φ, which is not a nnf, we put

$$\beta(\varphi) = \beta(\bar{\varphi}).$$

For a logic $L \in \mathcal{E}\mathbf{bN4}^{\perp}$, we define $\beta(L) = \mathbf{N3} + \{\beta(\varphi) \mid \varphi \in L\}$.

Again, every $\mathbf{bN4}^{\perp}$-extension L is faitfully embedded into $\beta(L)$ via β.

Theorem 4.4 [18] For every logic $L \in \mathcal{E}\mathbf{bN4}^{\perp}$ and formula $\varphi \in \mathrm{For}(\mathcal{L}^b)$, the following equivalence holds:

$$\varphi \in L \iff \beta(\varphi) \in \beta(L).$$

In particular, the logic $\mathbf{bN4}^{\perp}$ is faithfully embedded into **N3** via β.

From this we can easily infer

Corollary 4.5 The mapping $L \mapsto \beta(L)$, $L \in \mathcal{E}\mathbf{bN4}^{\perp}$, is an order preserving embedding of the lattice $\mathcal{E}\mathbf{bN4}^{\perp}$ into $\mathcal{E}\mathbf{N3}$.

Now it remains to prove that the lattice embeddings defined by the translations σ and β are mutually inverse, and since the lattices that are isomorphic as orders are isomorpic as lattices also, we obtain.

Theorem 4.6 [18] The mappings $L \mapsto \sigma(L)$, $L \in \mathcal{E}\mathbf{N3}$, and $L \mapsto \beta(L)$, $L \in \mathcal{E}\mathbf{bN4}^{\perp}$, are mutually inverse lattice isomorphisms between $\mathcal{E}\mathbf{N3}$ and $\mathcal{E}\mathbf{bN4}^{\perp}$.

We pass to the investigation of $\mathbf{nN4}^{\perp}$-extensions.

Consider an embedding of $\mathbf{N4}^{\mathsf{N}}$ into $\mathbf{nN4}^{\perp}$.

The translation δ: For$(\mathcal{L}) \to$ For$(\mathcal{L}^n)$ is given by the rule

$$\delta(\varphi) := \left(\bigwedge_{p \in \mathrm{var}(\varphi)} \neg\neg(p \vee \sim p) \right) \to \varphi\,.$$

For arbitrary $L \in \mathcal{E}\mathbf{N4}^{\mathsf{N}}$, put

$$\delta(L) = \mathbf{nN4}^{\perp} + \{\delta(\varphi) \mid \varphi \in L\}\,.$$

Lemma 4.7 For every formula $\varphi \in$ For$(\mathcal{L})$, Heyting algebra $\mathcal{A}$, and an ideal Δ on $\mathcal{A}$, the following equivalence holds:

$$Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \delta(\varphi) \ \text{ iff } \ Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \models \varphi\,.$$

Proof Assume that $Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \delta(\varphi)$ and consider a $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)$-valuation v. Due to the inclusion $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \subseteq Tw(\mathcal{A}, \mathcal{A}, \Delta)$ we can consider v as a $Tw(\mathcal{A}, \mathcal{A}, \Delta)$-valuation too.

By assumption $\pi_1 v(\delta(\varphi)) = \pi_1 v(\bigwedge_{p \in \mathrm{var}(\varphi)} \neg\neg(p \vee \sim p)) \to \pi_1 v(\varphi) = 1$, whence $\pi_1 v(\bigwedge_{p \in \mathrm{var}(\varphi)} \neg\neg(p \vee \sim p)) \leq \pi_1 v(\varphi)$. Since v is a $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)$-valuation, we have $\pi_1 v(p) \vee \pi_2 v(p) = \pi_1 v(p \vee \sim p) \in F_{\mathrm{d}}(\mathcal{A})$ for every $p \in$ Prop. The latter is equivalent to $\pi_1 v(\neg\neg(p \vee \sim p)) = 1$. Consequently, $\pi_1 v(\bigwedge_{p \in \mathrm{var}(\varphi)} \neg\neg(p \vee \sim p)) = 1$ and $\pi_1 v(\varphi) = 1$. We have thus proved that $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \models \varphi$.

Now we assume that $Tw(\mathcal{A}, \mathcal{A}, \Delta) \not\models \delta(\varphi)$. Let a $Tw(\mathcal{A}, \mathcal{A}, \Delta)$-valuation v be such that $\pi_1 v(\delta(\varphi)) \neq 1$. The latter means that

$$\pi_1 v \left(\bigwedge_{p \in \mathrm{var}(\varphi)} \neg\neg(p \vee \sim p) \right) \not\leq \pi_1 v(\varphi)\,.$$

Denote $a = \pi_1 v(\bigwedge_{p \in \mathrm{var}(\varphi)} \neg\neg(p \vee \sim p))$ and $b = \pi_1 v(\varphi)$, then $a \not\leq b$. Let $\langle a \rangle$ be a filter on $\mathcal{A}$ generated by a, $\langle a \rangle = \{c \mid a \leq c\}$. Then we have $b \notin \langle a \rangle$. Consider the twist-structure $Tw(\mathcal{A}/\langle a \rangle, \mathcal{A}/\langle a \rangle, \Delta/\langle a \rangle)$ and the valuation v' given by the rule $v'(p) =$

$(\pi_1 v(p)/\langle a\rangle, \pi_2 v(p)/\langle a\rangle)$. An easy induction on the structure of formulas allows us to prove that for any formula $\psi \in \mathrm{For}(\mathcal{L}^n)$, we have $v'(\psi) = (\pi_1 v(\psi)/\langle a\rangle, \pi_2 v(\psi)/\langle a\rangle)$. In particular, we have $\pi_1 v'(\varphi) = b/\langle a\rangle$. Since $b \notin \langle a\rangle$, $b/\langle a\rangle \neq 1_{\mathcal{A}/\langle a\rangle}$, whence $Tw(\mathcal{A}/\langle a\rangle, \mathcal{A}/\langle a\rangle, \Delta/\langle a\rangle) \not\models \varphi$. At the same time, $\pi_1 v'(\bigwedge_{p\in \mathrm{var}(\varphi)} \neg\neg(p\vee \sim p)) = a/\langle a\rangle = 1_{A/\langle a\rangle}$, which implies for all $p \in \mathrm{var}(\varphi)$ that $\pi_1 v'(\neg\neg(p\vee \sim p)) = 1_{\mathcal{A}/\langle a\rangle}$, i.e., $\pi_1 v'(p\vee \sim p) = \pi_1 v'(p) \vee \pi_2 v'(p) \in F_{\mathrm{d}}(\mathcal{A}/\langle a\rangle)$ for all $p \in \mathrm{var}(\varphi)$. It is true that $F_{\mathrm{d}}(\mathcal{A}/\langle a\rangle) = F_{\mathrm{d}}(\mathcal{A})/\langle a\rangle$. Indeed

$$F_{\mathrm{d}}(\mathcal{A}/\langle a\rangle) = \{c/\langle a\rangle \vee \neg c/\langle a\rangle \mid c \in \mathcal{A}\},\ \ F_{\mathrm{d}}(\mathcal{A})/\langle a\rangle = \{(c \vee \neg c)/\langle a\rangle \mid c \in \mathcal{A}\}\,.$$

It remains to notice that $c/\langle a\rangle \vee \neg c/\langle a\rangle = (c \vee \neg c)/\langle a\rangle$ for all $c \in \mathcal{A}$. It follows from the above considerations that on propositional variables from $\mathrm{var}(\varphi)$, the valuation v' acts as a $Tw(\mathcal{A}/\langle a\rangle, F_{\mathrm{d}}(\mathcal{A})/\langle a\rangle, \Delta/\langle a\rangle)$-valuation. This fact and $\pi_1 v'(\varphi) \neq 1_{\mathcal{A}/\langle a\rangle}$ imply $Tw(\mathcal{A}/\langle a\rangle, F_{\mathrm{d}}(\mathcal{A})/\langle a\rangle, \Delta/\langle a\rangle) \not\models \varphi$, i.e., the identity $\varphi \to \varphi = \varphi$ does not hold on $Tw(\mathcal{A}/\langle a\rangle, F_{\mathrm{d}}(\mathcal{A})/\langle a\rangle, \Delta/\langle a\rangle)$. By (3.4) the twist-structure $Tw(\mathcal{A}/\langle a\rangle, F_{\mathrm{d}}(\mathcal{A})/\langle a\rangle, \Delta/\langle a\rangle)$ is isomorphic to $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)/\langle a\rangle^{\bowtie}$, the quotient algebra of $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)$, which means that $\varphi \to \varphi = \varphi$ is not an identity of $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)$ too. Consequently, $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \not\models \varphi$. □

Theorem 4.8 For every logic $L \in \mathcal{E}\mathbf{N4}^{\mathsf{N}}$ and formula $\varphi \in \mathrm{For}(\mathcal{L})$, the following equivalence holds:

$$\varphi \in L \iff \delta(\varphi) \in \delta(L)\,.$$

Proof That $\varphi \in L$ implies $\delta(\varphi) \in \delta(L)$ follows from the definition of the logic $\delta(L)$. Take a formula $\varphi \notin L$ and an algebra $\mathcal{A} \in V(L)$ such that $\mathcal{A} \not\models \varphi$. Since L extends $\mathbf{N4}^{\mathsf{N}}$, we have $\mathcal{A} \models \mathbf{N4}^{\mathsf{N}}$. Consequently, $\mathcal{A} \cong Tw(\mathcal{B}, F_{\mathrm{d}}(\mathcal{B}), \Delta)$ for some $\mathcal{B}$ and Δ. By the previous lemma we have $Tw(\mathcal{B}, \mathcal{B}, \Delta) \models \delta(\psi)$ for all $\psi \in L$, whence $Tw(\mathcal{B}, \mathcal{B}, \Delta) \in V(\delta(L))$. At the same time, $Tw(\mathcal{B}, \mathcal{B}, \Delta) \not\models \delta(\varphi)$ in view of $\mathcal{A} \not\models \varphi$. We have thus proved that $\delta(\varphi) \notin \delta(L)$. □

Corollary 4.9 The logic $\mathbf{N4}^{\mathsf{N}}$ is faithfully embedded into $\mathbf{nN4}^{\perp}$ via δ.

Proof It remains to notice that $\delta(\mathbf{N4}^{\mathsf{N}}) = \mathbf{nN4}^{\perp}$, i.e., $\delta(\varphi) \in \mathbf{nN4}^{\perp}$ for all $\varphi \in \mathbf{N4}^{\mathsf{N}}$. If $\varphi \in \mathbf{N4}^{\mathsf{N}}$, then $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \models \varphi$ for all $\mathcal{A}$ and Δ. By Lemma 4.7 we obtain $Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \delta(\varphi)$ for all $\mathcal{A}$ and Δ. Every $\mathbf{nN4}^{\perp}$-lattice is isomorphic to a twist-structure of the form $Tw(\mathcal{A}, \mathcal{A}, \Delta)$, consequently, $\delta(\varphi) \in \mathbf{nN4}^{\perp}$. □

Corollary 4.10 The mapping $L \mapsto \delta(L)$, $L \in \mathcal{E}\mathbf{N4}^{\mathsf{N}}$, is an order preserving embedding of the lattice $\mathcal{E}\mathbf{N4}^{\mathsf{N}}$ into $\mathcal{E}\mathbf{nN4}^{\perp}$.

Proof It follows by definition that this mapping preserves the inclusion relation. Let $L_1, L_2 \in \mathcal{E}\mathbf{N4}^{\mathsf{N}}$ and $\varphi \in L_1 \setminus L_2$. By Theorem 4.8, we obtain that $\delta(\varphi) \in \delta(L_1) \setminus \delta(L_2)$. Thus, the mapping $L \mapsto \delta(L)$ is one-to-one. □

In fact, the mapping $L \mapsto \sigma(L)$ establishes an isomorphism between lattices $\mathcal{E}\mathbf{N4}^{\mathsf{N}}$ and $\mathcal{E}\mathbf{nN4}^{\perp}$. Again we define a translation $\gamma : \mathrm{For}(\mathcal{L}^n) \to \mathrm{For}(\mathcal{L})$ and prove that it allows us to define an embedding of $\mathcal{E}\mathbf{nN4}^{\perp}$ into $\mathcal{E}\mathbf{N4}^{\mathsf{N}}$, which is inverse to the embedding defined by δ.

For formulas that are nnfs the translation γ: For($\mathcal{L}^n$) $\to$ For($\mathcal{L}$) is defined as follows:

$$\gamma(\perp) = \perp,\ \ \gamma(\sim\perp) = \sim\perp,\ \ \gamma(n) = \perp,\ \ \gamma(\sim n) = \perp\,,$$
$$\gamma(p_i) := p_{2i},\ \ \gamma(\sim p_i) = p_{2i+1}\wedge \sim p_{2i}\,,$$
$$\gamma(\varphi \diamond \psi) := \gamma(\varphi) \diamond \gamma(\psi),\ \ \diamond \in \{\vee, \wedge, \to\}\,.$$

For a formula φ, which is not a nnf, we put

$$\gamma(\varphi) = \gamma(\bar{\varphi})\,.$$

For a logic $L \in \mathcal{E}\mathbf{nN4}^{\perp}$, we define $\gamma(L) = \mathbf{N4}^{\mathsf{N}} + \{\gamma(\varphi) \mid \varphi \in L\}$.

Lemma 4.11 For every formula $\varphi \in$ For($\mathcal{L}^n$), Heyting algebra $\mathcal{A}$, and ideal Δ on $\mathcal{A}$, the following equivalence holds:

$$Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \models \gamma(\varphi)\ \text{ iff }\ Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \varphi\,.$$

Proof Since $Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \varphi$ iff $Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \bar{\varphi}$, it will be enough to consider the case of a nnf $\varphi(p_0, \dots, p_n)$.

Assume that $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \models \gamma(\varphi)$ and take arbitrary $Tw(\mathcal{A}, \mathcal{A}, \Delta)$-valuation v. Consider a $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)$-valuation v' such that:

$$v'(p_{2i}) = (\pi_1 v(p_i), \neg\pi_1 v(p_i) \vee \pi_2 v(p_i)),\ \ v'(p_{2i+1}) = (\pi_2 v(p_i), \neg\pi_2 v(p_i))\,,$$

where $i = 0, \dots, n$. To check the correctness of this definition we calculate

$$\pi_1 v'(p_{2i}) \vee \pi_2 v'(p_{2i}) = \pi_1 v(p_i) \vee \neg\pi_1 v(p_i) \vee \pi_2 v(p_i) \in F_{\mathrm{d}}(\mathcal{A})\,.$$

At the same time

$$\pi_1 v'(p_{2i}) \wedge \pi_2 v'(p_{2i}) = (\pi_1 v(p_i) \wedge \neg\pi_1 v(p_i)) \vee (\pi_1 v(p_i) \wedge \pi_2 v(p_i)) \in \Delta\,.$$

Consequently, $v'(p_{2i}) \in Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)$. For $v'(p_{2i+1})$, we have

$$\pi_1 v'(p_{2i+1}) \vee \pi_2 v'(p_{2i+1}) = \pi_2 v(p_i) \vee \neg\pi_2 v(p_i) \in F_{\mathrm{d}}(\mathcal{A})\,,$$
$$\pi_1 v'(p_{2i+1}) \wedge \pi_2 v'(p_{2i+1}) = \pi_2 v(p_i) \wedge \neg\pi_2 v(p_i) = 0 \in \Delta\,.$$

We have thus proved $v'(p_{2i+1}) \in Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)$.

Further, $\pi_1 v'(\gamma(p_i)) = \pi_1 v'(p_{2i}) = \pi_1 v(p_i)$ and

$$\begin{aligned}\pi_1 v'(\gamma(\sim p_i)) &= \pi_1 v'(p_{2i+1}\wedge \sim p_{2i}) = \pi_2 v(p_i) \wedge (\neg\pi_1 v(p_i) \vee \pi_2 v(p_i)) \\ &= \pi_2 v(p_i) = \pi_1 v(\sim p_i)\,.\end{aligned}$$

Using these facts by induction on the structure of formula we can prove that

$$\pi_1 v'(\gamma(\varphi)) = \pi_1 v(\varphi)\,.$$

By assumption $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \models \gamma(\varphi)$, consequently $\pi_1 v(\varphi) = \pi_1 v'(\gamma(\varphi)) = 1$. We have thus proved $Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \varphi$.

Now we assume that $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta) \not\models \gamma(\varphi)$. Let $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \Delta)$-valuation v be such that $\pi_1 v(\gamma(\varphi)) \neq 1$. Consider a valuation v' such that

$$v'(p_i) = (\pi_1 v(p_{2i}), \pi_2 v(p_{2i}) \wedge \pi_1 v(p_{2i+1})),\ i = 0, \ldots, n\,.$$

Check that v is a $Tw(\mathcal{A}, \mathcal{A}, \Delta)$-valuation. Indeed,

$$\pi_1 v(p_{2i}) \wedge \pi_2 v(p_{2i}) \wedge \pi_1 v(p_{2i+1}) \leq \pi_1 v(p_{2i}) \wedge \pi_2 v(p_{2i}) \in \Delta\,.$$

We have $\pi_1 v(\gamma(p_i)) = \pi_1 v(p_{2i}) = \pi_1 v'(p_i)$ and

$$\begin{aligned}\pi_1 v(\gamma(\sim p_i)) &= \pi_1 v(p_{2i+1} \wedge \sim p_{2i}) = \pi_1 v(p_{2i+1}) \wedge \pi_2 v(p_{2i})\\ &= \pi_2 v'(p_i) = \pi_1 v'(\sim p_i)\,.\end{aligned}$$

Again by induction on the structure of formulas we obtain

$$\pi_1 v(\gamma(\varphi)) = \pi_1 v'(\varphi)\,.$$

Consequently, $\pi_1 v'(\varphi) \neq 1$ and $Tw(\mathcal{A}, \mathcal{A}, \Delta) \not\models \varphi$. □

Using this lemma and following the line of proof of Theorem 4.8 we can prove

Theorem 4.12 For every logic $L \in \mathcal{E}\mathbf{nN4}^{\perp}$ and formula $\varphi \in \mathrm{For}(\mathcal{L}^n)$, the following equivalence holds:

$$\varphi \in L \iff \gamma(\varphi) \in \beta(L)\,.$$

The next two statements are obtained similarly to Corollaries 4.9 and 4.10.

Corollary 4.13 The logic $\mathbf{nN4}^{\perp}$ is faithfully embedded into $\mathbf{N4}^{\mathsf{N}}$ via γ.

Corollary 4.14 The mapping $L \mapsto \gamma(L)$, $L \in \mathcal{E}\mathbf{nN4}^{\perp}$, is an order preserving embedding of the lattice $\mathcal{E}\mathbf{nN4}^{\perp}$ into $\mathcal{E}\mathbf{N4}^{\mathsf{N}}$.

It remains to prove that the lattice embeddings defined by the translations σ and γ are mutually inverse.

Theorem 4.15 The mappings $L \mapsto \delta(L)$, $L \in \mathcal{E}\mathbf{N4}^{\mathsf{N}}$, and $L \mapsto \gamma(L)$, $L \in \mathcal{E}\mathbf{nN4}^{\perp}$, are mutually inverse lattice isomorphisms between $\mathcal{E}\mathbf{N4}^{\mathsf{N}}$ and $\mathcal{E}\mathbf{nN4}^{\perp}$.

Proof Let $\mathcal{A}$ be a Heyting algebra and Δ an ideal on $\mathcal{A}$. It follows from Lemma 4.7 and the definition of the logic $\delta(L)$ that

$$Tw(\mathcal{A}, F_d(\mathcal{A}), \Delta) \models L \iff Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \delta(L) \tag{4.1}$$

for all $L \in \mathcal{E}\mathbf{N4}^{\mathsf{N}}$.

Similarly, from Lemma 4.11 and the definition of the logic $\gamma(L)$ we obtain

$$Tw(\mathcal{A}, \mathcal{A}, \Delta) \models \delta(L) \iff Tw(\mathcal{A}, F_d(\mathcal{A}), \Delta) \models \gamma\delta(L) \tag{4.2}$$

for all $L \in \mathcal{E}\mathbf{nN4}^{\perp}$.

It follows from (4.1) and (4.2) that the logics L and $\gamma(\delta(L))$ have the same models of the form $Tw(\mathcal{A}, F_d(\mathcal{A}), \Delta)$. Consequently, $L = \gamma(\delta(L))$ for all $L \in \mathcal{E}\mathbf{N4}^{\mathsf{N}}$.

In the same way, the logics L and $\delta(\gamma(L))$ have the same models of the form $Tw(\mathcal{A}, \mathcal{A}, \Delta)$, which implies that $L = \delta(\gamma(L))$ for all $L \in \mathcal{E}\mathbf{nN4}^{\perp}$.

Consequently, the mappings determined by δ and γ are mutually inverse. By Corollaries 4.10 and 4.14, $\mathcal{E}\mathbf{N4}^{\mathsf{N}}$ and $\mathcal{E}\mathbf{nN4}^{\perp}$ are isomorphic as orders. Consequently, they are isomorphic as lattices. □

Finally, we consider the connections between extensions of the logic $\mathbf{N4}^{\mathrm{Bel}}$ and extensions of $\mathbf{N3}^{\circ}$. We can use the composition of translations σ and δ to embed the lattice of $\mathbf{N3}^{\circ}$ into $\mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$.

For $L \in \mathcal{E}\mathbf{N3}^{\circ}$, put

$$\delta\sigma(L) = \mathbf{N4}^{\mathrm{Bel}} + \{\delta(\sigma(\varphi)) \mid \varphi \in L\}\,.$$

Lemma 4.16 For every formula $\varphi \in \mathrm{For}(\mathcal{L})$ and Heyting algebra $\mathcal{A}$, the following equivalence holds:

$$\mathcal{A}^{\bowtie} = Tw(\mathcal{A}, \mathcal{A}, \mathcal{A}) \models \delta(\sigma(\varphi)) \;\text{ iff }\; Tw(\mathcal{A}, F_d(\mathcal{A}), \{0\}) \models \varphi\,.$$

Proof By Lemma 4.7 the relation $Tw(\mathcal{A}, \mathcal{A}, \mathcal{A}) \models \delta(\sigma(\varphi))$ is equivalent to $Tw(\mathcal{A}, F_d(\mathcal{A}), \mathcal{A}) \models \sigma(\varphi)$. The last statement is equivalent in turn to $Tw(\mathcal{A}, F_d(\mathcal{A}), \{0\}) \models \varphi$ by Lemma 4.3. □

From this lemma arguing as above we infer

Theorem 4.17 For every logic $L \in \mathcal{E}\mathbf{N3}^{\circ}$ and formula $\varphi \in \mathrm{For}(\mathcal{L})$, the following equivalence holds:

$$\varphi \in L \iff \delta(\sigma(\varphi)) \in \delta\sigma(L)\,.$$

In particular, $\mathbf{N3}^{\circ}$ is faithfully embedded into $\mathbf{N4}^{\mathrm{Bel}}$ via the composition $\delta \circ \sigma$.

Corollary 4.18 The mapping $L \mapsto \delta\sigma(L)$, $L \in \mathcal{E}\mathbf{N3}^{\circ}$, is an order preserving embedding of the lattice $\mathcal{E}\mathbf{N3}^{\circ}$ into $\mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$.

The inverse translation $\varepsilon : \mathrm{For}(\mathcal{L}^{\mathrm{Bel}}) \to \mathrm{For}(\mathcal{L})$ is defined for nnfs as follows:

$$\begin{gathered}\varepsilon(\bot) = \bot,\ \ \varepsilon(\sim\bot) = \sim\bot\,,\\ \varepsilon(b) = \sim\bot,\ \ \varepsilon(\sim b) = \sim\bot\,,\\ \varepsilon(n) = \bot,\ \ \varepsilon(\sim n) = \bot\,,\\ \varepsilon(p_i) := p_{2i}\,,\ \ \varepsilon(\sim p_i) = p_{2i+1}\,,\\ \varepsilon(\varphi \diamond \psi) := \varepsilon(\varphi) \diamond \varepsilon(\psi),\ \ \diamond \in \{\vee, \wedge, \to\}\,.\end{gathered}$$

For a formula φ, which is not a nnf, we put

$$\varepsilon(\varphi) = \varepsilon(\bar{\varphi})\,.$$

For a logic $L \in \mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$, we define $\varepsilon(L) = \mathbf{N3}^{\circ} + \{\varepsilon(\varphi) \mid \varphi \in L\}$.

Lemma 4.19 For every formula $\varphi \in \mathrm{For}(\mathcal{L}^{n})$ and Heyting algebra $\mathcal{A}$, the following equivalence holds:

$$Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \{0\}) \models \varepsilon(\varphi) \ \text{ iff } \ \mathcal{A}^{\bowtie} = Tw(\mathcal{A}, \mathcal{A}, \mathcal{A}) \models \varphi\,.$$

Proof We consider only the case of nnf s.

Assume that $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \{0\}) \models \varepsilon(\varphi)$ and take arbitrary $Tw(\mathcal{A}, \mathcal{A}, \mathcal{A})$-valuation v. Consider a $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \{0\})$-valuation v' such that

$$v'(p_{2i}) = (\pi_1 v(p_i), \neg\pi_1 v(p_i)),\ v'(p_{2i+1}) = (\pi_2 v(p_i), \neg\pi_2 v(p_i))\,,$$

where $i = 0, \ldots, n$. An easy induction on the structure of formulas allows us to prove that

$$\pi_1 v'(\varepsilon(\varphi)) = \pi_1 v(\varphi)\,.$$

By assumption $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \{0\}) \models \varepsilon(\varphi)$, consequently $\pi_1 v(\varphi) = \pi_1 v'(\varepsilon(\varphi)) = 1$. We have thus proved $Tw(\mathcal{A}, \mathcal{A}, \mathcal{A}) \models \varphi$.

Now let $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \{0\}) \not\models \varepsilon(\varphi)$, and let v be a $Tw(\mathcal{A}, F_{\mathrm{d}}(\mathcal{A}), \{0\})$-valuation such that $\pi_1 v(\varepsilon(\varphi)) \neq 1$. Consider a $Tw(\mathcal{A}, \mathcal{A}, \Delta)$-valuation v' such that

$$v'(p_i) = (\pi_1 v(p_{2i}), \pi_1 v(p_{2i+1})),\ i = 0, \ldots, n\,.$$

By induction on the structure of formulas we prove $\pi_1 v(\varepsilon(\varphi)) = \pi_1 v'(\varphi)$. Consequently $\pi_1 v'(\varphi) \neq 1$ and $Tw(\mathcal{A}, \mathcal{A}, \mathcal{A}) \not\models \varphi$. □

Using this lemma, we can prove

Theorem 4.20 For every logic $L \in \mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$ and formula $\varphi \in \mathrm{For}(\mathcal{L}^{\mathrm{Bel}})$, the following equivalence holds:

$$\varphi \in L \iff \varepsilon(\varphi) \in \varepsilon(L)\,.$$

Corollary 4.21 The mapping $L \mapsto \varepsilon(L)$, $L \in \mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$, is an order preserving embedding of the lattice $\mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$ into $\mathcal{E}\mathbf{N3}^{\circ}$.

The proof that the lattice embeddings defined by the translations $\delta\circ\sigma$ and ε are mutually inverse is completely analogous to that of Theorem 4.15.

Theorem 4.22 The mappings $L \mapsto \delta\sigma(L)$, $L \in \mathcal{E}\mathbf{N3}^{\circ}$, and $L \mapsto \varepsilon(L)$, $L \in \mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$, are mutually inverse lattice isomorphisms between $\mathcal{E}\mathbf{N3}^{\circ}$ and $\mathcal{E}\mathbf{N4}^{\mathrm{Bel}}$.

Acknowledgment This work was supported by Russian Foundation for Basic Research, project No. 12-01-00168-a, and by the Grants Council (under RF President) for State Aid of Leading Scientific Schools (grant NSh-860.2014.1).

References

1. Almukdad, A., Nelson, D.: Constructible falsity and inexact predicates. J. Symb. Log. **49**, 231–233 (1984)
2. Belnap, N.D.: How a computer should think. In: Ryle, G. (ed.): Contemporary Aspects of Philosophy, pp. 30–56. Oriel Press, Stocksfield (1976)
3. Belnap, N.D.: A useful four-valued logic. In: Dunn, J.M., Epstein, G. (eds.) Modern Uses of Multiple-Valued Logics, pp. 8–37. D. Reidel Publishing, Dordrecht (1977)
4. Busaniche, M., Cignoli, R.: Residuated lattices as algebraic semantics for paraconsitent Nelson's logic. J. Log. Comput. **19**, 1019–1029 (2009)
5. Busaniche, M., Cignoli, R.: Constructive logic with strong negation as a substructural logic. J. Log. Comput. **20**, 761–793 (2010)
6. Dunn, J.M.: Intuitive semantics for first degree entailments and coupled trees. Phil. Stud. **29**, 149–168 (1976)
7. Fidel, M.: An algebraic study of a propositional system of Nelson. In: Arruda, A., da Costa, N., Chuaqui, R. (eds.) Mathematical Logic. Proceedings of the First Brasilian Conference, Lectures in Pure and Applied Mathematics, vol. 39, pp. 99–117. Marcel Dekker, New York (1978)
8. Galatos, N., Jipsen, P., Kowalski, T., Ono, H.: Residuated Lattices: An Algebraic Glimpse at Substructural Logics. Elsevier, Amsterdam (2007)
9. Goranko, V.: The Craig interpolation theorem for propositional logics with strong negation. Stud. Log. **44**(3), 291–317 (1985)
10. Gurevich, Y.: Intuitionistic logic with strong negation. Stud. Log. **36**, 49–59 (1977)
11. López-Escobar, E.G.K.: Refutability and elementary number theory. Indag. Math. **34**, 362–374 (1972)
12. Nelson, D.: Constructible falsity. J. Symb. Log. **14**, 16–26 (1949)
13. Odintsov, S.: Algebraic semantics for paraconsistent Nelson's Logic. J. Log. Comput. **13**, 453–468 (2003)
14. Odintsov, S.: On representation of N4-lattices. Stud. Log. **76**, 385–405 (2004)

15. Odintsov, S.: The Class of Extensions of Nelson's paraconsistent logic. Stud. Log. **80**, 293–322 (2005)
16. Odintsov, S.: On extensions of Nelson's logic satisfying Dummett's axiom. Sib. Math. J. **48**, 112–125 (2007)
17. Odintsov, S.: Constructive Negations and Paraconsistency. Springer, Dordrecht (2008)
18. Odintsov, S.: On the equivalence of paraconsistent and explosive versions of Nelson logic. In: Brattka, V., Diener, H., Spreen, D. (eds.) Logic, Computation, Hierarchies, pp. 227–241. Ontos Publishing, Berlin (2014)
19. Odintsov, S.P., Wansing, H.: Modal logics with Belnapian truth values. J. Appl. Non-Class. Log. **20**, 279–301 (2010)
20. Rasiowa, H.: N-lattices and constructive logic with strong negation. Fund. Math. **46**, 61–80 (1958)
21. Rasiowa, H.: An Algebraic Approach to Non-Classical Logic. PWN, Warsaw and North-Holland, Amsterdam (1974)
22. Routley, R.: Semantical analyses of propositional systems of Fitch and Nelson. Stud. Log. **33**(3), 283–298 (1974)
23. Spinks, M., Veroff, R.: Constructive logic with strong negation is a substructural logic I. Stud. Log. **88**, 325–348 (2008)
24. Spinks, M., Veroff, R.: Constructive logic with strong negation is a substructural logic II. Stud. Log. **89**, 401–425 (2008)
25. Thomason, R.H.: A semantical study of constructible falsity. Z. Math. Log. Grundl. Math. **15**, 247–257 (1969)
26. Vakarelov, D.: Notes on N-lattices and constructive logic with strong negation. Stud. Log. **34**, 109–125 (1977)
27. Vorob'ev, N.N.: A constructive propositional calculus with strong negation (in russian). Doklady AN SSSR **85**, 465–468 (1952)
28. Vorob'ev, N.N.: The problem of deducibility in constructive propositional calculus with strong negation (in russian). Doklady AN SSSR **85**, 689–692 (1952)

S. Odintsov (✉)
Sobolev Institute of Mathematics, Novosibirsk, Russian Federation
e-mail: odintsov@math.nsc.ru

Negative Modalities in the Light of Paraconsistency

Hitoshi Omori and Toshiharu Waragai

Abstract Modality and non-classical negation have some interesting connections. One of the most famous connections is the relation between **S4**-modality and intuitionistic negation. In this chapter, we focus on the negative modalities in the perspective of paraconsistency. The basic idea here is to consider the negative modality defined as 'not necessarily' or equivalently 'possibly not' where 'not' is classical negation, and 'necessarily' and 'possibly' are modalities in modal logics. This chapter offers a solution to the problem of axiomatizing systems of modal logic such as **D** and **S4** in terms of negative modalities. One of the upshots of this solution is that we may consider the semantics of paraconsistency with the help of various considerations known in the literature of modal logics related to **D** and **S4**.

Keywords Modal logic · Negative modality · Paraconsistency

Mathematics Subject Classification (2000) Primary 03B45 · Secondary 03B53

1 Introduction

We first give the background of this chapter and then clarify our aim and outline. These are then followed by some preliminaries.

1.1 Background

In developing systems of modal logic, the usual strategy is to enrich the language of classical logic by affirmative modalities which are read to be 'necessarily', 'possibly', etc. However, there is another possible strategy for this purpose. That is to add negative modalities instead of affirmative modalities. Here, by negative modalities, we mean connectives to be obtained by combinations of affirmative modalities and classical negation. Thus, negative modalities may be read as 'not necessarily', 'necessarily not', 'not possibly', 'possibly not', etc. where 'not' is to be taken as classical negation.

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_23

Now, the negative modalities are sometimes connected to negations in non-classical logics. The most famous example must be the intuitionistic negation which is interpreted as 'necessarily not' where necessity here is known to be the modality in the system **S4** together with the heredity condition. Another example can be found in the literature of paraconsistent logic. Indeed, in [3], Jean-Yves Béziau observed explicitly that negative modality 'not necessarily' (or equivalently 'possibly not'), where the modalities being that of **S5**, may be regarded as a paraconsistent negation. But then, there must be a rather natural objection when we talk about paraconsistent *negation*. Indeed, the only condition accepted by paraconsistent logicians is the failure of the rule known as *ex falso quodlibet*, i.e., the following:

$$A, \neg A \not\vdash B\ .$$

Here, the unary operation $\neg$ is intended to be a negation, though the definition of negation is not given in general, or at least there is no agreement amongst paraconsistent logicians, and this has been one of the problems for those who raise some objections against paraconsistency. For example, Hartley Slater in [24] argues that paraconsistent negation is not a negation according to a traditional account. On the other hand, there is an attempt of defining negation. For example, Wolfgang Lenzen in [14, 15] offers some examinations of negation through three kinds of conditions, i.e., unacceptable conditions, dispensable conditions and indispensable conditions. And based on these conditions, he classifies various negations including those in intuitionistic logic and paraconsistent logics. In this chapter, we will not go into this philosophical debate.[1] Instead, we will simply understand negation as a unary operation which meets only some of the theses of classical negation.[2] In particular, paraconsistent negations are those which satisfies some of the theses of classical negation but not *ex falso quodlibet*.

Regardless of the motivation, some people turned their attention to the observation of Béziau. Indeed, after the formulation, given in [4], of the system **HZ**, which may be regarded as an axiomatization of **S5** in terms of a negative modality, those such as João Marcos [16], and Krystyna Mruczek-Nasieniewska and Marek Nasieniewski, in a series of papers [17–19], made some effort to extend the results of Béziau. More concretely, they discovered some axioms for formulating modal logics different from **S5** by a negative modality 'not necessarily' (or equivalently 'possibly not'). However, their attempts are not entirely conclusive. In particular, they have not discovered the appropriate axioms for systems such as **D** and **S4** for which many interesting semantic considerations are known in the literature.

[1] We only note that as Michael De is pointing out in [7, Chap. 2], a semantic approach rather than a syntactic approach seems to be promising in this context.

[2] Strictly speaking, this will exclude the connexive negation (cf. [28]), but our purpose here is to clarify our usage of the word 'negation', not to claim something about negation in general. Thus we do not necessarily claim that connexive negation is not a negation.

1.2 Aim and Outline

Based on the above observations, the aim of this chapter is to add some further results following the observation of Béziau. The main result is the axiomatizations of **D** and **S4** through a negative modality, which remained open for a while. In view of our result, we may expect some new perspectives on paraconsistent negation. If paraconsistency is not of interest, then one can place the result in the context of modal logic, and regard it as a result on the formulation of modal logics that contain axioms for **D** and **S4** in terms of a negative modality.

This chapter is organized as follows. After some preliminaries in the next subsection, we quickly review some of the preceding results related to the present work in the second section. Then, in the third section, we refine some of the existing results which simplifies some of the axiomatization considerably. We will meet the new and elegant axiomatization of **S5** in terms of a negative modality in this section. The fourth section will be devoted to answering some of the open problems. Namely, we provide axiomatizations of **D** and **S4**, and also some of the extensions of **S4**. We also make some remarks on the topological interpretation of negation induced by the topological interpretation of the **S4**-modality. Finally, we conclude this chapter by summarizing our results and indicating some future directions.

1.3 Preliminaries

First, in the present work, we need the negation-less fragment of the intuitionistic propositional calculus and the classical propositional calculus. For this purpose, we fix the former as follows:

(A1) $A \supset (B \supset A)$ (A2) $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$
(A3) $A \supset (A \vee B)$ (A4) $B \supset (A \vee B)$ (A5) $(A \supset C) \supset ((B \supset C) \supset ((A \vee B) \supset C))$
(A6) $(A \wedge B) \supset A$ (A7) $(A \wedge B) \supset B$ (A8) $(C \supset A) \supset ((C \supset B) \supset (C \supset (A \wedge B)))$
(MP) If A and $A \supset B$ are theorems, then so is B.

We refer to this system as $\mathbf{IL}^+$. The extension of $\mathbf{IL}^+$ which can be obtained by adding either Peirce's law or Dummett's law (i.e., $A \vee (A \supset B)$) to $\mathbf{IL}^+$ will be the classical positive propositional calculus and we refer to this system as $\mathbf{CL}^+$. We also refer to the classical propositional calculus as **CL**.

Second, our concern lies in some relation between modal logics, expansions of **CL** and paraconsistent logics, subsystems of **CL**. Therefore, we need some tool to bridge these two logics living in the different languages. For this purpose, we introduce some translations. Let us refer to the languages of classical logic and modal logic as $\mathcal{L}_{\mathbf{CL}}$ and $\mathcal{L}_{\mathbf{ML}}$, respectively, which consist of set of logical connectives $\{\sim, \supset, \wedge, \vee\}$ and $\{\neg, \Box, \supset, \wedge, \vee\}$, respectively, together with a set of propositional variables Prop, and also refer to the sets of formulas defined in the usual way within those languages as $\mathsf{For}_{\mathbf{CL}}$ and $\mathsf{For}_{\mathbf{ML}}$, respectively.[3]

Definition 1.1 Let τ_1 be a translation from $\mathsf{For}_{\mathbf{ML}}$ to $\mathsf{For}_{\mathbf{CL}}$ that satisfies the following conditions:

$$\begin{array}{lll} \tau_1(p_i) = p_i & \tau_1(\Box A) = \neg^* {\sim} \tau_1(A) & \tau_1(A \supset B) = \tau_1(A) \supset \tau_1(B) \\ \tau_1(\neg A) = \neg^* \tau_1(A) & \tau_1(A \wedge B) = \tau_1(A) \wedge \tau_1(B) & \tau_1(A \vee B) = \tau_1(A) \vee \tau_1(B), \end{array}$$

where $\neg^* A$ is defined as $A \supset {\sim}(A \supset A)$. On the other hand, let τ_2 be a translation from $\mathsf{For}_{\mathbf{CL}}$ to $\mathsf{For}_{\mathbf{ML}}$ that satisfies the following conditions:

$$\begin{array}{lll} \tau_2(p_i) = p_i & \tau_2(A \supset B) = \tau_2(A) \supset \tau_2(B) & \tau_2(A \vee B) = \tau_2(A) \vee \tau_2(B) \\ \tau_2({\sim} A) = \neg \Box \tau_2(A) & \tau_2(A \wedge B) = \tau_2(A) \wedge \tau_2(B)\,. & \end{array}$$

Now, let $\mathbf{S_{ML}}$ be a system of modal logic. Then our main goal is to find or reformulate the subsystem $\mathbf{S_{CL}}$ of **CL** that satisfies the following conditions:

$$\vdash_{\mathbf{S_{ML}}} A \text{ iff } \vdash_{\mathbf{S_{CL}}} \tau_1(A) \text{ and } \vdash_{\mathbf{S_{CL}}} B \text{ iff } \vdash_{\mathbf{S_{ML}}} \tau_2(B)\,.$$

When $\mathbf{S_{CL}}$ satisfies the above two conditions with respect to $\mathbf{S_{ML}}$, then we shall call such $\mathbf{S_{CL}}$ as $\mathbf{nS_{ML}}$ where '**n**' is intended to be read 'negative'. For example, if we take **S5** for $\mathbf{S_{ML}}$, then the corresponding subsystem of **CL** will be referred to as **nS5**.

Finally, we list some formulas and rules which we make use of in the following sections.

(AT) $A \vee {\sim}A$
(AB) ${\sim}{\sim}A \supset A$
(AS5) ${\sim}A \supset ({\sim}{\sim}A \supset B)$
(AS4) ${\sim}({\sim}A \supset B) \supset ({\sim}A \wedge {\sim}B)$
(AK1) ${\sim}(A \wedge B) \supset ({\sim}A \vee {\sim}B)$
(AK2) ${\sim}(A \supset A) \supset (A \wedge {\sim}A)$
(AD) $(A \supset A) \supset {\sim}{\sim}(A \supset A)$
(AS4.4) $A^I \supset (({\sim}A \vee B) \supset (A \supset B))$
(Rep) If $A \supset B$ is a theorem, then so is ${\sim}B \supset {\sim}A$.
(RMar) If A is a theorem, then so is ${\sim}A \supset B$.

Note here that $A^I =_{\text{def}} A \supset {\sim}{\sim}A$.

2 Revisiting the Previous Results

We now recall some of the previous results related to the present work. We will present them in chronological order of publication. Also, the proofs for the results will be omitted.

[3] For the purpose of emphasizing that negation in the language $\mathcal{L}_{\mathbf{CL}}$ is *not* necessarily intended to be classical negation whereas $\mathcal{L}_{\mathbf{ML}}$ *is* intended to be classical negation, we distinguished the negations by writing $\sim$ and $\neg$, respectively.

2.1 Work of Béziau

The relation between modal logic and paraconsistent logic is rather old. Indeed, Stanisław Jaśkowski, one of the founding fathers of paraconsistent logic, already made use of modal logic in developing his system of paraconsistent logic $\mathbf{D}_2$ (cf. [10, 11]). However, it seems that Béziau was the first who pointed out another kind of tight connection between paraconsistent negation and modality. The system he developed in [4] was axiomatized as follows.

Definition 2.1 (Béziau) The system **HZ** consists of the following formulas and a rule of inference in addition to $\mathbf{CL}^+$:

(AZ1)$A \vee {\sim}A$ (AZ2)$(A \wedge {\sim}B) \wedge {\sim}(A \wedge {\sim}B) \supset (A \wedge {\sim}A)$
(AZ3)${\sim}(A \wedge B) \supset ({\sim}A \vee {\sim}B)$ (AZ4)${\sim}{\sim}A \supset A$
(RZ) If $A \supset B$ is a theorem, then so is ${\sim}(A \wedge {\sim}B)$.

Remark 2.2 Note that the above axiomatization is not independent as is proved in [21]. According to the result presented in [21], we obtain an independent axiomatization by eliminating (AZ1) and (AZ3), and replacing (AZ2) by the formula ${\sim}(A \wedge {\sim}B) \supset ({\sim}B \supset {\sim}A)$. However, this axiomatization is still far from being intuitive. We will present an elegant axiomatization for **HZ** in the next section. Note also that Diderik Batens [2] developed a system which is also equivalent to **HZ**. There is a difference though that Batens employed a language with two primitive negations whereas we only have one negation here, just as in the language of classical logic.

Then, the provability and deducibility are defined as follows:

Definition 2.3 A formula A is *provable* in **HZ** (notation $\vdash_{\mathbf{Z}} A$) iff there exists a finite sequence of formulas $B_1, \ldots, B_n$ such that $B_n = A$ and B_i $(1 \leq i \leq n)$ is an axiom of **HZ** or is the conclusion of a rule whose premises are among $B_1, \ldots, B_m$ $(m < i)$.

Definition 2.4 A formula A is *deducible* in **HZ** from a theory Γ (notation $\Gamma \vdash_{\mathbf{Z}} A$) iff there are $n \geq 0$, $B_1, \ldots, B_n \in \Gamma$ such that $\vdash_{\mathbf{Z}} B_1 \wedge \cdots \wedge B_n \supset A$.

Remark 2.5 Note that the deducibility relation is defined in terms of provable formulas, as is usually done in modal logic (e.g. [9, p. 211]). However, as is well-known, there is at least another definition of deducibility *not* relying on provable formulas as follows: $\Gamma \vdash_{\mathbf{S}} A$ if there is a sequence of formulas $B_1, \ldots, B_n, A$, $n \geq 0$, such that every formula in the sequence $B_1, \ldots, B_n, A$ either (i) belongs to Γ; (ii) is an axiom of **S**; (iii) is obtained by one of the rules of **S** from formulas preceding it in sequence. And in the context of paraconsistency, even if we have the same set of axioms and rules, the paraconsistent behavior may be lost depending on the definition of deducibility. In this chapter, we will stick to the convention of modal logic to keep the connection with modal logic.

Now, the interesting relation proved by Béziau is the following.

Theorem 2.6 (Béziau) Let τ_2 be the translation introduced in Definition 1.1. Then,

$$\Gamma \vdash_{\mathbf{Z}} A \text{ iff } \tau_2(\Gamma) \vdash_{\mathbf{S5}} \tau_2(A)\,.$$

Based on the above theorem, a natural question is to ask if we can generalize this kind of relation to other modal logics different from **S5**. And it is this line of investigation that we will focus in the present chapter. As we shall see in the following, there are mainly two works preceding the present chapter. One is by Marcos, and the other is by Mruczek-Nasieniewska and Nasieniewski.

2.2 Works of Marcos and Mruczek-Nasieniewska and Nasieniewski

Motivated by the work of Béziau, Marcos made effort to generalize the results so that the translation results hold for other modal logics different from **S5**. And in [16], Marcos observed that system **nK** can be obtained by adding (AK1), (RMar) and (Rep.) to $\mathbf{CL}^+$. He also found that systems **nKT**, **nKB**, **nK5** can be obtained by adding formulas (AT), (AB) and (AS5) to **nK**, respectively.

On the other hand, in their series of papers [17–19], Mruczek-Nasieniewska and Nasieniewski also extended the work of Béziau. Some of the results of Marcos were rediscovered independently, but they also explored further where Marcos had not reached. Indeed, they deal with non-normal modal logics, known as regular modal logics, in [18]. (The base system they consider consists of (AK1) and (Rep), and they consider its extensions by (AT), (AS5) and $\sim(A \supset A)$, in our notation.) And finally in [19], they consider the case of quasi-regular modal logics.

2.3 Work of Waragai and Shidori

Another work in this direction was presented in [30] by Toshiharu Waragai and Tomoki Shidori. The intuition of their work is quite different from that of Béziau. The starting point is not in modal logic, but rather in paraconsistent logic in the tradition of Newton da Costa (cf. [5]). Indeed, the system they introduced was axiomatized with the help of the notion of 'behaving classically' which was the characteristic notion of da Costa's systems $\mathbf{C}_n$. With a slight modification presented in [29], the system can be formulated as follows.

Definition 2.7 The system **PCL1** consists of the following formulas and a rule of inference in addition to $\mathbf{IL}^+$:

$\mathrm{A}_{\mathbf{PCL1}}10$	$\sim\sim A \supset A$	$\mathrm{A}_{\mathbf{PCL1}}13$	$(A^I \wedge B^I) \supset (\sim(A \wedge B) \supset (\sim A \vee \sim B))$
$\mathrm{A}_{\mathbf{PCL1}}11$	$A \vee \sim A$	$\mathrm{A}_{\mathbf{PCL1}}14$	$(A^I \wedge B^I) \supset (\sim(A \vee B) \supset (\sim A \wedge \sim B))$
$\mathrm{A}_{\mathbf{PCL1}}12.1$	$A^I \supset ((\sim A \vee B) \supset (A \supset B))$	$\mathrm{A}_{\mathbf{PCL1}}15$	$A^I \supset ((A \wedge \sim B) \supset \sim(A \supset B))$

RA15 $$\frac{A^I}{\sim(A \supset B) \supset (A \wedge \sim B)}\,.$$

Note here that $A^I =_{\text{def}} A \supset \sim\sim A$.

As observed in [29], there are several differences between the systems $\mathbf{C}_n$ of da Costa and **PCL1** of Waragai and Shidori, though the main difference lies in the relation with modal logic. We quote a theorem from [30] which highlights the difference.

Theorem 2.8 (Waragai and Shidori) Let us consider an extension of **PCL1** enriched by Peirce's law and the following rule of inference:

$$\text{If } A \text{ is a theorem, then so is } {\sim}{\sim}A\,. \tag{RN}$$

This system, which we refer to as **PCL1CN**, satisfies the following :

$$\vdash_{\mathbf{PCL1CN}} A \;\text{ iff }\; \vdash_{S5} \tau(A)\,.$$

Remark 2.9 Instead of adding (RN), we can add (Rep) in the above theorem. This shows one of the differences between $\mathbf{C}_n$ and **PCL1**, since it is proved in [26] that we obtain classical logic if we add (Rep) to $\mathbf{C}_n$. Note also that we can prove the equivalence of **HZ** and **PCL1CN**. Moreover, the extension of **PCL1** by (Rep) is equivalent to $\mathbf{TCC}_\omega$ of Gordienko introduced in [8].[4]

3 Reexamination of the Results of Marcos

We now turn to examine the result of Marcos.

Definition 3.1 Let $\mathbf{nK}'$ be a system which consists of (AK1), (AK2) and (Rep.) in addition to $\mathbf{CL}^+$.

Proposition 3.2 Two systems **nK** and $\mathbf{nK}'$ are equivalent.

Proof In order to prove that $\mathbf{nK}'$ is a subsystem of **nK**, we have to establish that (AK2) is derivable in **nK**. But this follows immediately by (RMar). Indeed, we only have to apply (RMar) to $A \supset A$. On the other hand, in order to prove that **nK** is a subsystem of $\mathbf{nK}'$, we have to show that the rule (RMar) is derivable in $\mathbf{nK}'$. For this purpose, suppose A. By (A1), we get $(B \supset B) \supset A$. Furthermore, by applying (Rep.), we have ${\sim}A \supset {\sim}(B \supset B)$. Hence this together with (AK2), we obtain ${\sim}A \supset B$, which is the desired conclusion. This completes the proof. □

Remark 3.3 As is well known, normal modal logics are formulated by adding a single rule, called necessitation, together with various axioms. And the above result shows that similar result holds in the case of formulating normal modal logics in terms of a negative modality.

[4] These results are proved in [22].

Definition 3.4 Let **nKB**′ be the extension of $\mathbf{CL}^+$ enriched with (AB) and (Rep).

Proposition 3.5 Two systems **nKB** and **nKB**′ are equivalent.

Proof Since it is obvious that **nKB**′ is a subsystem of **nKB**, we only have to show that **nKB** is a subsystem of **nKB**′. For this purpose, it will be sufficient to prove two formulas (AK1) and (AK2) in **nKB**′. The details can be found in the Appendix. □

Remark 3.6 Note that Ray Jennings observed a similar result in [12] that the system **KTB** can be axiomatized by two axioms $\Box A \supset A$ and $A \supset \Box \Diamond A$ and the following rule together with classical propositional logic[5]:

$$\text{if } A \supset B \quad \text{is a theorem, then so is } \Box A \supset \Box B\,.$$

Thus our result is being a generalization of Jennings' result as we can even establish a simple axiomatization for **KB** which is weaker than **KTB**.

Theorem 3.7 **nKT5** can be axiomatized as (AT), (AS5) and (Rep) with $\mathbf{CL}^+$.

Proof Let **nS5**′ be the system which consists of (AT), (AS5), (Rep) with $\mathbf{CL}^+$. Then, our goal is to prove **nKT5** is equivalent to **nS5**′. The key here is that if we have (AT) and (AS5) together with $\mathbf{CL}^+$, then we can prove (AB). Indeed, by (AS5), we have $\sim\sim A \supset (\sim A \supset A)$, and by (AT), we get $(\sim A \supset A) \supset A$. Therefore, we obtain $\sim\sim A \supset A$, as desired. Thus, both systems **nS5**′ and **nKT5** turn out to be equivalent to **nKB**′ enriched with (AT) and (AS5) in view of Proposition 3.5. This completes the proof. □

Remark 3.8 **nS5'** seems to be an elegant axiomatization of **nS5**. It is interesting to compare with the fact that classical propositional logic can be obtained by adding (AT) and $(A \wedge \sim A) \supset B$ to $\mathbf{CL}^+$.

4 Some Formulation of Systems D, S4, etc.

We now proceed to consider some of the open problems, and draw some remarks from the results.

4.1 Answers to Some Open Problems

We will first deal with the formulation of **KD** in terms of a negative modality.

Definition 4.1 Let **nKD** be a system obtained by adding (AD) to **nK**′.

[5] This is also mentioned in [9, p. 69, Exercise 3.13].

Lemma 4.2 Let τ_1 and τ_2 be translations defined in Definition 1.1. Then,

(i)	If $\vdash_{\mathbf{KD}} A$ then $\vdash_{\mathbf{nKD}} \tau_1(A)$	(iii)	$\vdash_{\mathbf{KD}} A \equiv \tau_2(\tau_1(A))$
(ii)	If $\vdash_{\mathbf{nKD}} B$ then $\vdash_{\mathbf{KD}} \tau_2(B)$	(iv)	$\vdash_{\mathbf{nKD}} B \equiv \tau_1(\tau_2(B))$.

Proof As for (i), it would be sufficient to prove $\vdash_{\mathbf{nKD}} \tau_1(\Box A \supset \Diamond A)$, whereas it would be sufficient to prove $\vdash_{\mathbf{KD}} \tau_2({\sim}{\sim}(A \supset A))$ for (ii). (iii) and (iv) can be proved by induction on the complexity of A. Details can be found in the Appendix. □

Theorem 4.3 $\vdash_{\mathbf{KD}} A$ iff $\vdash_{\mathbf{nKD}} \tau_1(A)$, and $\vdash_{\mathbf{nKD}} B$ iff $\vdash_{\mathbf{KD}} \tau_2(B)$.

Proof We only consider the former equivalence since the latter can be proved in a similar manner. Now, the left-to-right direction is already proved in (i) of the previous lemma. The other way around can be proved by making use of (ii) and (iii). Indeed, by (ii), we obtain that if $\vdash_{\mathbf{nKD}} \tau_1(A)$ then $\vdash_{\mathbf{KD}} \tau_2(\tau_1(A))$. Therefore, by applying (iii) to the consequent, we obtain that if $\vdash_{\mathbf{nKD}} \tau_1(A)$ then $\vdash_{\mathbf{KD}} A$, as desired. □

Let us now turn to the formulation of **K4**. The outline of the proof is exactly the same as the case for **KD**.

Definition 4.4 Let **nK4** be a system obtained by adding (AS4) to **nK**′.

Lemma 4.5 Let τ_1 and τ_2 be translations defined in Definition 1.1. Then,

(i)	If $\vdash_{\mathbf{K4}} A$ then $\vdash_{\mathbf{nK4}} \tau_1(A)$	(iii)	$\vdash_{\mathbf{K4}} A \equiv \tau_2(\tau_1(A))$
(ii)	If $\vdash_{\mathbf{nK4}} B$ then $\vdash_{\mathbf{K4}} \tau_2(B)$	(iv)	$\vdash_{\mathbf{nK4}} B \equiv \tau_1(\tau_2(B))$.

Proof As for (i), it would be sufficient to prove $\vdash_{\mathbf{nK4}} \tau_1(\Box A \supset \Box\Box A)$, whereas it would be sufficient to prove $\vdash_{\mathbf{K4}} \tau_2({\sim}({\sim}A \supset B) \supset ({\sim}A \wedge {\sim}B))$ for (ii). (iii) and (iv) can be proved by induction on the complexity of A. Again, details can be found in the Appendix. □

Theorem 4.6 $\vdash_{\mathbf{K4}} A$ iff $\vdash_{\mathbf{nK4}} \tau_1(A)$, and $\vdash_{\mathbf{nK4}} B$ iff $\vdash_{\mathbf{K4}} \tau_2(B)$.

Proof Apply the same procedure given in the proof of Theorem 4.3. □

Remark 4.7 Based on this result, we obtain a paraconsistent negation with a topological semantics, as we will observe in the next subsection. Another point to be noted is the following. As is well known, there are many systems of modal logics that are extensions of modal logic **K4**. Therefore, once we get **nK4**, then it is natural to ask if those extensions of modal logic can be axiomatized through negative modalities. And, for many cases this is possible. However, for the purpose of keeping the paper compact, we will leave the details for another paper. Finally, a slightly modified version of the system **KT4**, obtained by deploying the deducibility relation mentioned in Remark 2.5, offers us a new perspective on the problem related to the law of double negation first considered by da Costa and

Béziau in [6], and later discussed by Urbas in [27]. We will keep the details of this too for another occasion.

Finally, let us consider the following system which is closely related to the axiom included in **PCL1**.

Definition 4.8 Let **nS4.4** be the extension of **nKT4** obtained by adding (AS4.4).

Theorem 4.9 $\vdash_{\mathbf{S4.4}} A$ iff $\vdash_{\mathbf{nS4.4}} \tau_1(A)$, and $\vdash_{\mathbf{nS4.4}} B$ iff $\vdash_{\mathbf{S4.4}} \tau_2(B)$.

Proof Apply the same procedure given in the proof of Theorem 4.3. □

Remark 4.10 Note first that (AS4.4) is essentially the axiom which is employed in the formulation of **PCL1** (cf. ($A_{\mathbf{PCL1}}$12.1) of Definition 2.7.). And, together with the axiom (AB), we obtain the relation to **S5**, but the result here shows that without the axiom (AB), then we obtain a connection to **S4.4**. Note further that (AS4.4) is equivalent to the following formula within $\mathbf{IPC}^+$:

$$(A \wedge {\sim}A \wedge {\sim}{\sim}A) \supset B\,.$$

In fact, this formula has been known for a long time in the literature, and has been criticized to be included in a system of paraconsistent logic. Indeed, Jaśkowski refers to this thesis in [10, p. 41] as the thesis already realized by Jan Łukasiewicz as a result of study in three-valued logic. And the system having this formula as a thesis is not accepted by Jaśkowski as a solution to the problem of finding systems of paraconsistent logic.

4.2 Topological Semantics

Since we obtained a formulation of **S4** in terms of a negative modality, it is natural to consider the topological semantics of **nS4** which is induced by the topological semantics for **S4**. We will devote this subsection to spell out that semantics, and make some related remarks. We begin by revisiting the classic result of Mackinsey and Tarski.

Definition 4.11 (topological validity) Let $\langle X, \mathcal{O}, i\rangle$ be a topological model, i.e., $\langle X, \mathcal{O}\rangle$ is a topological space and i is a mapping from $\mathsf{For}_{\mathbf{ML}}$ to X which naturally extends to a valuation v from Prop to X by the following clauses.

- $i(p_i) = v(p_i)$
- $i(\neg A) = X \backslash i(A)$
- $i(\Box A) = i(A)^o$ where o is the interior operator.
- $i(A \wedge B) = i(A) \cap i(B)$
- $i(A \vee B) = i(A) \cup i(B)$
- $i(A \supset B) = (X \backslash i(A)) \cup i(B)$.

Then, A is *topologically valid* if $i(A) = X$ for any topological model $\langle X, \mathcal{O}, i\rangle$.

Theorem 4.12 (Mackinsey and Tarski) A is topologically valid iff $\vdash_{\mathbf{S4}} A$.

Now, in view of this result together with our result, we define n-topological validity as follows.[6]

Definition 4.13 (n-topological validity) Let $\langle X, \mathcal{O}, j\rangle$ be an n-topological model, i.e., $\langle X, \mathcal{O}\rangle$ is a topological space and j is a mapping from $\mathsf{For}_{\mathbf{CL}}$ to X which naturally extends to a valuation v from Prop to X by the following clauses:

- $j(p_i) = v(p_i)$
- $j(\sim A) = X \backslash (j(A)^o)$, where o is the interior operator.
- $j(A \wedge B) = j(A) \cap j(B)$
- $j(A \vee B) = j(A) \cup j(B)$
- $j(A \supset B) = (X \backslash j(A)) \cup j(B)$.

Then, A is *n-topologically valid* if $j(A) = X$ for any n-topological model $\langle X, \mathcal{O}, j\rangle$.

Lemma 4.14 Let $\langle X, \mathcal{O}, i\rangle$ and $\langle X, \mathcal{O}, j\rangle$ be topological and n-topological model induced by a topological space $\langle X, \mathcal{O}\rangle$ and a valuation v. Then, for any $A \in \mathsf{For}_{\mathbf{CL}}$, we have $j(A) = i(\tau_2(A))$.

Proof By an induction on the complexity of the formula. For the base case, if $A = p_i$, then $j(p_i) = v(p_i) = i(p_i) = i(\tau_2(p_i))$. For the induction step, we only consider the case when $A = \sim B$ since others are straightforward. The proof for the case in question runs as follows:

$$\begin{aligned} j(\sim B) &= X \backslash (j(B)^o) && \text{definition of } j \\ &= X \backslash (i(\tau_2(B))^o) && \text{Induction hypothesis} \\ &= i(\neg \Box \tau_2(B)) && \text{definition of } i \\ &= i(\tau_2(\sim B)) && \text{definition of } \tau_2\,. \end{aligned}$$

This completes the proof. □

Corollary 4.15 For any $A \in \mathsf{For}_{\mathbf{CL}}$, A is n-topologically valid iff $\tau_2(A)$ is topologically valid.

Proof Let $\langle X, \mathcal{O}\rangle$ be any topological space and v be any valuation. Then, these naturally induce topological model $\langle X, \mathcal{O}, i\rangle$ and n-topological model $\langle X, \mathcal{O}, j\rangle$. By the above lemma, we have $j(A) = X$ iff $i(\tau_2(A)) = X$ for any $A \in \mathsf{For}_{\mathbf{CL}}$. And this is sufficient for the desired result. □

Based on these, we reach the following result.

Theorem 4.16 For any $A \in \mathsf{For}_{\mathbf{CL}}$, A is n-topologically valid iff $\vdash_{\mathbf{nS4}} A$.

---

[6] Where 'n' is for 'negative'.

Proof By the combination of the above results as follows.

$\vdash_{\mathbf{nS4}} A$ iff $\vdash_{\mathbf{S4}} \tau_2(A)$	Theorem 4.6
iff $\tau_2(A)$ is topologically valid	Theorem 4.12
iff A is n-topologically valid.	Corollary 4.15 .

This completes the proof. □

Remark 4.17 In view of the above result, it is interesting to examine the semantics for paraconsistent logics with the help of topology. For some recent discussions toward this direction, see [1].

5 Conclusion

Our main result is a discovery of some axiomatizations of well-known modal logics through negative modalities. In view of our results, we may conclude that many of the well-known modal logics can be developed in terms of a negative modality instead of an affirmative modality. Then, it might be interesting to discuss whether any differences will be observed from a philosophical viewpoint between the two formulations, one based on an affirmative modality and the other based on a negative modality. Seen in this way, the result we presented may be dealt within the story of modal logic.

However, it is true that we were motivated by a result within paraconsistency given by Béziau. But then, what kind of conclusions can we draw from that perspective? For optimistic paraconsistent logicians, the upshot will be that we get more semantic tools for paraconsistency. However, we find ourselves not being relaxed enough. The most interesting and important question seems to be to ask if the negative modalities such as 'not necessarily' or 'possibly not' are negations or not. In other words, the question is to ask if we can draw a line between negative modalities and negations. This certainly requires some reasonable accounts of negation and modality, which seems to be not available in the literature. Thus, we wish to pursue this question in the subsequent paper.

More technically speaking, there seems to be at least three directions to proceed. Let us illustrate them briefly.

First, we should consider the first-order expansions of the modal logics. As is known, there are some interesting problems such as the way to deal with Barcan formula. We here note briefly that if we apply the translation, then the formula is reduced to be de Morgan's law for the universal quantifier with respect to the paraconsistent negation. In this way, things might look somewhat different from the case employing the usual affirmative modality.

Second, we may add more modalities such as with bi-modal logic. This will then give us systems equipped with several negations. In the case of bi-modal logics, we will have two negations. Then one of the topics that might be interesting is to see the connection between logics with two negations such as systems of Nelson (cf. [13, 20]) and connexive logic (cf. [28]).

Finally, we may consider the case for intuitionistic modal logics. For example, as is observed by Hiroakira Ono in [23], there are several systems for **S5** in the intuitionistic context. Furthermore, necessity and possibility will not be inter-definable in the usual way anymore. However, at the same time, there are some related results in the literature. Indeed, Richard Sylvan [25], considers extensions of $\mathbf{C}_{\omega}$ of da Costa to recover the inter-substitutability of provable equivalents and the resulting systems are intuitionic versions of **KB** in the following sense. Namely, from the semantic viewpoint, there is another relation between worlds besides the relation for modality, and from the proof theoretic perspective, the system can be obtained by simply employing $\mathbf{IL}^{+}$ instead of $\mathbf{CL}^{+}$. In this connection, the result of this chapter offers some intuitionistic versions for **KD** and **K4**, though these are only a small part of the bigger picture, and thus there is much work to be done to fill the gap.

Appendix

We give the full proofs for those we omitted in the body. For this purpose, let us raise the theses and a rule of **CL** which we make use of in the following:

$B \supset (A \supset A)$	(T1)	$(A \supset B) \supset (\neg B \supset \neg A)$	(T5)
$\neg(A \supset A) \supset B$	(T2)	$(A \supset B) \equiv (\neg B \supset \neg A)$	(T6)
$(A \wedge \neg A) \supset B$	(T3)	$A \equiv \neg\neg A$	(T7)
$(A \vee B) \supset (\neg A \supset B)$	(T4)	$\dfrac{A \supset B \quad B \supset C}{A \supset C}\,.$	(Syll.)

Note also that $\neg^{*}$ in extensions of **nK** is classical negation, and thus we can apply any of the above theses in extensions of **nK**, as well as to the negation $\neg$ in extensions of modal logic **K**.

Proof of Proposition 3.5

Ad$\sim(A \wedge B) \supset (\sim A \vee \sim B)$

1	$\sim(\sim A \vee \sim B) \supset \sim\sim A$	[(A3), (Rep)]
2	$\sim(\sim A \vee \sim B) \supset A$	[1, (AB), (Syll.)]
3	$\sim(\sim A \vee \sim B) \supset \sim\sim B$	[(A4), (Rep)]
4	$\sim(\sim A \vee \sim B) \supset B$	[3, (AB), (Syll.)]
5	$\sim(\sim A \vee \sim B) \supset (A \wedge B)$	[2, 4, (A8), (MP)]
6	$\sim(A \wedge B) \supset \sim\sim(\sim A \vee \sim B)$	[5, (Rep)]
7	$\sim(A \wedge B) \supset (\sim A \vee \sim B)\,.$	[6, (AB), (Syll.)]

Ad$\sim(A\supset A)\supset(A \wedge \sim A)$

1	$\sim(A\supset A)\supset\sim\sim A$	[(T1), (Rep)]
2	$\sim(A\supset A)\supset A$	[1, (AB), (Syll.)]
3	$\sim(A\supset A)\supset\sim A$	[(A1), (Rep)]
4	$\sim(A\supset A)\supset(A \wedge \sim A)$.	[2, 3, (A8), (MP)]

This completes the proof. □

Proof of Lemma 4.2

Ad (i) It would be sufficient to prove $\vdash_{\mathbf{nKD}} \tau_1(\Box A\supset\Diamond A)$, but this is equivalent to proving $\vdash_{\mathbf{nKD}} \neg^*\sim\tau_1(A)\supset\neg^*\neg^*\sim\neg^*\tau_1(A))$ by the definition of τ_1. Therefore, it is sufficient to prove $\vdash_{\mathbf{nKD}} \neg^*\sim A\supset\sim\neg^* A$. The proof runs as follows:

1	$\sim\sim(A\supset A)\supset\sim\neg^*(A\supset A)$	[(T2), (Rep.)]
2	$\sim\neg^*(A\supset A)$	[2, (AD), (MP)]
3	$\sim\neg^*(A\supset A)\supset\sim(A \wedge \neg^* A)$	[(T3), (Rep.)]
4	$\sim(A \wedge \neg^* A)$	[2, 3, (MP)]
5	$\sim(A \wedge \neg^* A)\supset(\neg^*\sim A\supset\sim\neg^* A)$	[(AK1), (T4), (Syll.)]
6	$\neg^*\sim A\supset\sim\neg^* A$.	[4, 5, (MP)]

Ad (ii) It would be sufficient to prove $\vdash_{\mathbf{KD}} \tau_2(\sim\sim(A\supset A))$, but this is equivalent to proving $\vdash_{\mathbf{KD}}$ $(\neg\Box\neg\Box(\tau_2(A)\supset\tau_2(A)))$ by the definition of τ_2. Thus we need to prove $\vdash_{\mathbf{KD}}$ $(\Diamond\Box(\tau_2(A)\supset\tau_2(A)))$, but this is easy.
Ad (iii) We proceed by induction on the number n of connectives.
(Base) If $n = 0$, then what we need to prove is $\vdash_{\mathbf{KD}} A \equiv A$ which is provable.
(Induction Step) We split the case depending on the main connective.
Case 1: If $A = \neg B$, then

$$\begin{aligned}
\tau_2(\tau_1(A)) &= \tau_2(\tau_1(\neg B)) \\
&= \tau_2(\tau_1(B) \supset \sim(\tau_1(B) \supset \tau_1(B))) \\
&= \tau_2(\tau_1(B)) \supset \tau_2(\sim(\tau_1(B) \supset \tau_1(B))) \\
&= \tau_2(\tau_1(B)) \supset \neg\Box\tau_2((\tau_1(B) \supset \tau_1(B))) \\
&= \tau_2(\tau_1(B)) \supset \neg\Box(\tau_2(\tau_1(B)) \supset \tau_2(\tau_1(B)))\,.
\end{aligned}$$

By (IH), this is equivalent to $B \supset \neg\Box(B \supset B)$. Furthermore, by (T6) we obtain $\Box(B \supset B) \supset \neg B$, and in view of axiom for **KD**, this is equivalent to $\neg B$, as desired.

Case 2: If $A = \Box B$, then

$$\begin{aligned}\tau_2(\tau_1(A)) &= \tau_2(\tau_1(\Box B))\\ &= \tau_2({\sim}\tau_1(B) \supset {\sim}(\tau_1(B) \supset \tau_1(B)))\\ &= \tau_2({\sim}\tau_1(B)) \supset \tau_2({\sim}(\tau_1(B) \supset \tau_1(B)))\\ &= \neg\Box\tau_2(\tau_1(B)) \supset \neg\Box\tau_2((\tau_1(B) \supset \tau_1(B)))\\ &= \neg\Box\tau_2(\tau_1(B)) \supset \neg\Box(\tau_2(\tau_1(B)) \supset \tau_2(\tau_1(B)))\,.\end{aligned}$$

By (IH), this is equivalent to $\neg\Box B \supset \neg\Box(B \supset B)$. Furthermore, by (T6) we obtain $\Box(B \supset B) \supset \Box B$, and in view of axiom for **KD**, this is equivalent to $\Box B$, as desired.
Case 3: If $A = B * C$ where $* \in \{\wedge, \vee, \supset\}$, then

$$\begin{aligned}\tau_2(\tau_1(A)) &= \tau_2(\tau_1(B * C))\\ &= \tau_2(\tau_1(B) * \tau_1(C))\\ &= \tau_2(\tau_1(B)) * \tau_2(\tau_1(C))\,.\end{aligned}$$

By (IH), this is equivalent to $B * C$, as desired.
Ad (iv) We proceed by induction on the number n of connectives.
(Base) If $n = 0$, then what we need to prove is $\vdash_{\mathbf{nKD}} A \equiv A$ which is provable.
(Induction Step) We split the case depending on the main connective.
Case 1: If $A = {\sim}B$, then

$$\begin{aligned}\tau_1(\tau_2(A)) &= \tau_1(\tau_2({\sim}B))\\ &= \tau_1(\neg\Box\tau_2(B))\\ &= \neg^*\tau_1(\Box\tau_2(B))\\ &= \neg^*\neg^*{\sim}\tau_1(\tau_2(B))\,.\end{aligned}$$

By (IH), this is equivalent to $\neg^*\neg^*{\sim}B$, and by (T7) we obtain ${\sim}B$, as desired.
Case 2: If $A = B * C$ where $* \in \{\wedge, \vee, \supset\}$, then

$$\begin{aligned}\tau_1(\tau_2(A)) &= \tau_1(\tau_2(B * C))\\ &= \tau_1(\tau_2(B) * \tau_2(C))\\ &= \tau_1(\tau_2(B)) * \tau_1(\tau_2(C))\,.\end{aligned}$$

By (IH), this is equivalent to $B * C$, as desired. □

Proof of Lemma 4.5

Ad (i) It would be sufficient to prove $\vdash_{\mathbf{nK4}} \tau_1(\Box A \supset \Box\Box A)$, but this is equivalent to prove $\vdash_{\mathbf{nK4}} \neg^*{\sim}\tau_1(A) \supset \neg^*{\sim}\neg^*{\sim}\tau_1(A)$ by the definition of τ_1. Therefore, it is sufficient to prove $\vdash_{\mathbf{nK4}} {\sim}\neg^*{\sim}A \supset {\sim}A$. Now, by recalling the definition of $\neg^*$, the formula concerned is equivalent to ${\sim}({\sim}A \supset {\sim}(A \supset A)) \supset {\sim}A$. So, it would be sufficient to prove ${\sim}({\sim}A \supset B) \supset {\sim}A$ for our purpose. But this can be obtained immediately by (AS4), (A6) and (Syll.).

Ad (ii) It would be sufficient to prove $\vdash_{\mathbf{K4}} \tau_2(\sim(\sim A \supset B) \supset (\sim A \wedge \sim B))$, but this is equivalent to proving $\vdash_{\mathbf{K4}} \neg\Box(\neg\Box\tau_2(A) \supset \tau_2(B)) \supset (\neg\Box\tau_2(A) \wedge \neg\Box\tau_2(B))$ by the definition of τ_2. Thus it would be sufficient to prove $\vdash_{\mathbf{K4}} \neg\Box(\neg\Box A \supset B) \supset (\neg\Box A \wedge \neg\Box B)$, and this can be proved as follows:

1	$\Box\Box A \supset \Box(\neg\Box A \supset B)$	[(T3), (R$\Box$)]
2	$\Box A \supset \Box(\neg\Box A \supset B)$	[1, Axiom **4**, (Syll.)]
3	$\neg\Box(\neg\Box A \supset B) \supset \neg\Box A$	[2, (T5), (MP)]
4	$\Box B \supset \Box(\neg\Box A \supset B)$	[(A1), (R$\Box$)]
5	$\neg\Box(\neg\Box A \supset B) \supset \neg\Box B$	[4, (T5), (MP)]
6	$\neg\Box(\neg\Box A \supset B) \supset (\neg\Box A \wedge \neg\Box B)$.	[3, 5, (A9), (MP)]

Note here that Axiom **4** and (R$\Box$) are the following axiom and rule, respectively:

Axiom **4** : $\Box A \supset \Box\Box A$, (R$\Box$): If $A \supset B$ is a theorem, then so is $\Box A \supset \Box B$.

Ad (iii) and (iv): Similar to the proof of Lemma 4.2. □

Acknowledgment Hitoshi Omori is a JSPS Postdoctoral Fellow for Research Abroad. An earlier version of the chapter was presented at a special session 'Negation' of UNILOG III in Lisbon. We would like to thank Michael De for his careful reading and suggestions to improve both content and English of the chapter. We would also like to thank Katsuhiko Sano and Makoto Kikuchi for their helpful discussions on the chapter.

References

1. Başkent, C.: Some topological properties of paraconsistent models. Synthese **190**, 4023–4040 (2013)
2. Batens, D.: On some remarkable relations between paraconsistent logics, modal logics, and ambiguity logics. In: Carnielli, W.A., Coniglio, M.E., D'Ottaviano, I.M.L. (eds.) Paraconsistency: The Logical Way to the Inconsistent, Proceedings of the II World Congress on Paraconsistency, pp. 275–293. Marcel Dekker, New York (2002)
3. Béziau, J.Y.: S5 is a paraconsistent logic and so is first-order classical logic. Log. Investig. **9**, 301–309 (2002)
4. Béziau, J.Y.: The Paraconsistent Logic Z – A possible solution to Jaśkowski's problem. Log. Log. Philos. **15**, 99–111 (2006)
5. da Costa, N.C.A.: On the theory of inconsistent formal systems. Notre Dame J. Form. Log. **15**, 497–510 (1974)
6. da Costa, N.C.A., Béziau, J.Y.: Carnot's logic. Bull. Sect. Log. **22**, 98–105 (1993)
7. De, M.: Negation in context. PhD thesis, University of St. Andrews, Scotland (2011)
8. Gordienko, A.B.: A paraconsistent extension of Sylvan's logic. Algebra Log. **46**(5), 289–296 (2007)
9. Hughes, G.E., Cresswell, M.J.: A New Introduction to Modal Logic. Routledge, London, New York (1996)
10. Jaśkowski, S.: A propositional calculus for inconsistent deductive systems. Log. Log. Philos. **7**, 35–56 (1999)
11. Jaśkowski, S.: On the discussive conjunction in the propositional calculus for inconsistent deductive systems. Log. Log. Philos. **7**, 57–59 (1999)

12. Jennings, R.: A note on the axiomatisation of Brouweresche modal logic. J. Philos. Log. **10**, 341–343 (1981)
13. Kamide, N., Wansing, H.: Proof theory of Nelson's paraconsistent logic: A uniform perspective. Theor. Comput. Sci. **415**, 1–38 (2012)
14. Lenzen; W.: Necessary conditions for negation operators. In: Wansing, H. (ed.) Negation: A Notion in Focus, pp. 37–58. Walter de Gruyter, Berlin, New York (1996)
15. Lenzen, W.: Necessary conditions for negation operators (with particular applications to paraconsistent negation). In: Besnard, P., Hunter, A. (eds.) Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. 2. Reasoning with Actual and Potential Contradictions, pp. 211–239. Kluwer, Dordrecht (1998)
16. Marcos, J.: Nearly every normal modal logics is paranormal. Log. Anal. **48**, 279–300 (2005)
17. Mruczek-Nasieniewska, K., Nasieniewski, M.: Syntactical and semantical characterization of a class of paraconsistent logics. Bull. Sect. Log. **34**(4), 118–125 (2005)
18. Mruczek-Nasieniewska, K., Nasieniewski, M.: Paraconsitent logics obtained by J.Y. Beziau's method by means of some non-normal modal logics. Bull. Sect. Log. **37**(3/4), 185–196 (2008)
19. Mruczek-Nasieniewska, K., Nasieniewski, M.: Beziau's logics obtained by means of quasi-regular logics. Bull. Sect. Log. **38**(3/4), 189–204 (2009)
20. Odintsov, S.: Constructive Negations and Paraconsistency. Springer, Dordrecht (2008)
21. Omori, H., Waragai, T.: On Béziau's logic Z. Log. Log. Philos. **17**(4), 305–320 (2008)
22. Omori, H., Waragai, T.: On extensions of a system of paraconsistent logic PCL1 (in Japanese). J. Jpn. Assoc. Philos. Sci. **39**(2), 1–18 (2012)
23. Ono, H.: On some intuitionistic modal logics. Publ. RIMS, Kyoto Univ. **13**, 687–722 (1977)
24. Slater, B.H.: Paraconsistent logics? J. Philos. Log. **24**(4), 451–454 (1995)
25. Sylvan, R.: Variations on da Costa C systems and dual-intuitionistic logics I. Analyses of C_ω and CC_ω. Stud. Log. **49**(1), 47–65 (1990)
26. Urbas, I.: Paraconsistency and the C-systems of da Costa. Notre Dame J. Form. Log. **30**, 583–597 (1989)
27. Urbas, I.: A Note on 'Carnot's Logic'. Bull. Sect. Log. **23**, 118–125 (1994)
28. Wansing, H.: Connexive logic. Stanford Encyclopedia of Philosophy (Fall 2014 Edition) (2014). http://plato.stanford.edu/entries/logic-connexive/
29. Waragai, T., Omori, H.: Some new results on PCL1 and its related systems. Log. Log. Philos. **19**(1–2), 129–158 (2010)
30. Waragai, T. Shidori, T.: A system of paraconsistent logic that has the notion of 'behaving classically' in terms of the law of double negation and its relation to S5. In: Béziau, J.Y., Carnielli, W.A., Gabbay, D. (eds.) Handbook of Paraconsistency, pp. 177–187. College Publications, London (2007)

H. Omori (✉)
The Graduate Center, City University of New York, New York, USA
e-mail: hitoshiomori@gmail.com

T. Waragai
Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Tokyo, Japan

Operativity and Representativity of the Sign in Leibniz

Olga Pombo

Abstract In opposition to, or in the line of, philosophy cannot grow except inside a tradition. This is precisely the issue with contemporary research on diagrammatic thinking. We all agree that it is essential to go back to Peirce or Husserl not only to recognize important roots, to honour significant predecessors, but to recover their insights, to make use of their hypothesis, to recuperate the conceptual instruments that they have been able to put forward. What I propose here is to go back even more in time, to go back to Leibniz. Not to claim for Leibniz as a predecessor of cognitive semantics or of present-day diagrammatic research, but to recall what Leibniz thought out in this respect, to bring into play the conceptual devices he put up, to bear in mind the distinctions he was able to constitute.

Keywords Leibniz · Descartes · Semiology · Theory of blind thought · Operativity of the sign · Representativity of the sign

Mathematics Subject Classification (2010) Primary 03A05 · Secondary 03B65

This paper is a homage to Jean-Yves Béziau and to his marvellous capacity of combining the strictness and rigor of logics with the care and compassion of meaning.

1 Leibniz' Cognitive Conception of Language

My first claim is that the historical importance of Leibniz' philosophy of language is mostly due to his cognitive conception of language, that is, his recognition of the constitutive symbolic nature of human thought. As he says in the celebrated *Dialogus de connexione inter res et verba* (1677):

> I will never be able to know, to discover, to prove without using words or without the presence in my mind of other signs. [2, 7.191]

This cognitive conception of language was mainly formulated in the scope of the debate with Descartes and of their different mathematical experiences. Descartes tends to star in geometry, in which figural representation is a merely auxiliary, imaginative support for reasoning. Leibniz stars in algebra where operations are fully symbolically performed,

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_24

where symbols constitute the reasoning process itself, completely replacing the supposed direct experience of mathematical objects.

In fact, the Cartesian principle of evidence entails a merely instrumental and communicative conception of language. For Descartes, language has above all a communicative function. The most language may be asked to do is to operate as a mnemonic support for the recall of the long chain of reasons. Leibniz refuses the intuitionism of Descartes. For Leibniz, rigour cannot be dependent either on subjectively based certainties or on the confidence in the intuitive infallibility of natural light. As Leibniz says in a *Letter to Gallois* (1677),

> [The Cartesian methodological rules] give surely beautiful decrees but not the guideline for developing those decrees. [2, 7.21]

According to Leibniz, rigour should be achieved by the use of a symbolic system which would render visible the more abstract ideas and would constitute a material support for thought and reasoning, a "*filum Ariadnes*" [2, 7.22][1], a "*filum palpabile*" [2, 7.57, 59, 125], a *filum cogitandi* [1, 420], a *filum meditandi* [2, 7: 14], a *filum mechanico* [1, 351], that is, a symbolic criterion, a manipulatory device. Something which mathematics has already developed and which, as Leibniz used to say, needs nothing but paper and ink. Let us look for what Leibniz writes in the well-known *Preface à la Science Générale* (1677)

> Now the reason why the art of demonstrating has been until now found only in mathematics [...] is this: mathematics carries its own test with it. For when I am presented with a false theorem, I do not need to examine or even to know the demonstration, since I shall discover its falsity a posteriori by means of an easy experiment that is, by a calculation, costing no more than paper and ink.
>
> The tests or experiments made in mathematics to guard against mistakes in reasoning [...] are not made on a thing itself, but on the characters which we have substituted in place of the thing.
>
> Take for example a numerical calculation: if 1677 times 365 are 612,105, we should hardly ever have reached this result if it were necessary to make 365 piles of 1677 pebbles each and then finally to count them all in order to know whether the aforementioned number is found. This test is performed only on paper, and consequently on the characters which represent the thing, and not on the thing itself. [1, 154]

Symbolism is for Leibniz the proper, the necessary and essential means of human reason. Leibniz even goes a step further by claiming that only symbolic signs allow us to operate with ideal significations, which can only be established by the sign and that only signs enable us to think. The example of the polygon of a thousand faces presented by Leibniz in his celebrated *Meditationes de Cognitione, Veritate et Ideis* (1684) is eloquent:

> When I think of a polygon of thousand faces, I do not always consider what is a face, an equality or the number thousand but I use this words (whose meaning is present on my spirit only in a very confuse and imperfect way) in order that they (the words) take the place of the ideas which I have of them [...] I call this knowledge as blind or symbolic. We make use of it in algebra and arithmetic and in almost all domains. [2, 4.423]

Again, Leibniz' mathematical experience lies in the basis of this important theory of blind thought (*cogitatio caeca*). In infinitesimal calculus, sign does not evoke but fully substitutes notions which the human mind cannot otherwise completely reach. So, I believe, we

[1] "La veritable méthode nous doit fournir un *Filum Ariadnes*, c'est à dire, un certain moyen sensible et grossier, qui conduise l'esprit, comme sont les lignes tracées en geometrie et les formes des operations qu'on prescrit aux apprentifs en Arithmetique" [2, 7.22].

here face two main correlated Leibnizian theses which deserve to be recalled by our current research on diagrammatic thinking: a) the not merely communicative but cognitive conception of language, b) the theory of blind thought.

2 Specificity of Leibniz' Semiology

My second claim is that it is necessary to be aware of the specificity of Leibniz' semiology, namely the prospective, heuristic potentialities that Leibniz attributes to symbolic systems. In the *Letter to the German Mathematician Walter von Tschirnhaus*, dated May 1678, Leibniz writes:

> No one should fear that the contemplation of characters will lead us away from the things themselves; on the contrary, it leads us into the interior of things. For we often have confused notions today because the characters we use are badly arranged; but then, with the aid of characters, we will easily have the most distinct notions, for we will have at hand a mechanical thread of meditation, as it were, with whose aid we can very easily resolve any idea whatever into those of which it is composed. In fact, if the character expressing any concept is considered attentively, the simpler concepts into which it is resolvable will at once come to mind. Since the analysis of concepts thus corresponds exactly to the analysis of a character, we need merely to see the characters in order to have adequate notions brought to our mind freely and without effort. We can hope for no greater aid than this in the perfection of the mind. [3, 4.461]

We find here some of the main thesis of the Leibniz semiology:

1. Characters do not disturb the knowledge of reality.
2. On the contrary, characters constitute a powerful means for the development of knowledge.
3. Characters not only reflect but also promote human knowledge allowing us to go from confused to clear, distinct notions.
4. This happens because of the written, graphic (iconic) nature of characters.
5. Characters are mostly valuable because they can be contemplated, because they can be spatially arranged, because they can be handled, manipulated and used as mechanical threads.
6. Characters could then become a mechanical thread of meditation.
7. However, for that to be possible, it is necessary that characters may express concepts.
8. With such expressive characters, the analysis of characters will correspond exactly to the analysis of concepts.
9. If this happens, then seeing the characters would be the easy way of getting adequate knowledge.
10. The result should be of the greatest importance for the perfection of the human mind.

Leibniz is always very clear as concerns the recognition of the qualities and potentialities coming from the spatial nature of characters. He extends that recognition, not only to the characters of the new artificial philosophical language to be constructed, but also to natural languages, namely to writing. The most illuminative passages in this respect appear in the *Nouveaux Essais*. While he recognizes the critical arguments presented by Lock's Philalethe against the difficulties, insufficiencies and disturbing effects put forward by natural languages[2], Leibniz position consists invariably in stressing the virtues of language and most of all the merits of writing for a more rigorous and regulated utilization of the various dimensions of signification. As he writes:

> Mais pour revenir à vos quatre défauts de la dénomination, je vous dirai, Monsieur, qu'on peut remédier à tous, surtout depuis que l'écriture est inventée (Nouveaux Essais, III, IX, §9)

Leibniz' project is thus double. He aims: 1) to overcome the difficulties of natural languages by the exploration of their capacities in terms of definition and writing, 2) to construct a new language, more exactly, a writing system or *Characteristica Universalis* of which mathematics would be just an example. As Leibniz says in a *Letter to Gallois* (December, 1678)

> But to make it easier and, so to speak, more tangible, I intend to make use of the characteristic, of which I have spoken with you on occasion, and of which algebra and mathematics are merely examples. This characteristic consists of a certain writing or language (since he who has the one can have the other) which perfectly corresponds to the relations of our thoughts. This characteristic would be completely different from any that has been envisaged until now, since the most important thing has been overlooked, which is that the characters of this writing have to serve for discovery and for judgement, like in algebra and in arithmetic. [4, II 1.669]

Again, it is clear that Leibniz praises the properties of writing. Only its two-dimensional spatiality makes possible a more clear and differentiated expression of relations.

We are here facing the enormous and very much ambitious project of a *Characteristica Universalis*. In all his life, Leibniz will be fully committed with this program which he claimed to be very much innovative.[3] In fact, Leibniz does not avoid criticism towards his predecessors in terms of the construction of new philosophical language, the line that goes from Lull to Kircher and that which goes from the English pasigraphers to Dalgarno and Wilkins.[4] However, according to Leibniz, none of his predecessors managed to guarantee the heuristic virtualities that symbolism may offer. Moreover, and this for two main reasons: the insufficient analysis of thoughts which, in Leibniz' opinion, underlies the set of *predicamenta* or *suma genera* in the basis of which their systems were constructed[5] and the arbitrary nature of the characters they have established[6].

---

[2] In fact, the 17th century discusses deeply the role that language performs in the process of knowledge. Does language help to promote knowledge? Or, on the contrary, is language a disturbing factor for the acquisition of knowledge? Two great positions can be signalized. A critical position, which emphasizes language's insufficiencies and ambiguities: Bacon, Locke, Descartes, Arnauld, Melabranche and, in general, all those who look for the construction of new artificial languages, and a positive position which, although recognizing some limits and imperfection of human languages, nevertheless stresses its constitutive character. From my point of view, just two names in modern times: Thomas Hobbes and Leibniz.

[3] Something which is more significant since, as well known, Leibniz was in general very much interested and attentive to all the achievements of his predecessors.

[4] Leibniz knows all their works well. He praised above all Wilkins's project, *An Essay Towards a Real Character and a Philosophical Language, with an Alphabetical Dictionary*, published in 1668, with the support of the Royal Society. As Leibniz writes: "J'ai consideré avec attention le grand ouvrage du Caractere reel et Langage Philosophique de Mons. Wilkins; je trouve qu'il y a mis une infinité de belles choses, et nous n'avons jamais eu une Table des predicateurs plus accomplie" (*Letter to Burnet*, 24 August, 1697 [2, 3.216]).

[5] See, for instance, [2, 3.216].

[6] See for instance, [1, 177 ff].

That is to say, contrary to what it may appear, Leibniz is not a full formalist. He does not fall into the illusion that the automatic functioning of a set of operational rules can permit the development of science. On the contrary, Leibniz claims that the new symbolic system should be semantically opened to the reality that it must permit to say. According to Leibniz, beyond responding to a logical objective of faithfully and rigorously expressed thought and its articulations, characters should open the road to the progress of knowledge. That heuristic capacity can only be achieved if the system of characters to be constructed could be directly open to the reality that they are supposed to say. Above and beyond operativity and functional capacity of the signs within the formal system, characters should also be "natural", that is representative of the world which is to be known. Leibniz' heuristics involves a semantic exigency.

3 Leibniz' Theory of the Representativity of the Sign

My third claim intends precisely to call attention to the representativity of the sign that Leibniz wanted to conquer for the characters of the philosophical language to be constructed.

With this requirement of the representativity of the sign, Leibniz looks for the isomorphism between characteristic signs and the reality that they signify. But how could it be done? How is it possible for a sign to represent reality? And which reality should be represented? Its sensory, particular traits as they appear to our perception? Its essence? The composition of its elements? The inner relation of its parts? Given the extreme difficulty of such a task, it is easy to understand the heterogeneity of this Leibnizian theory of the representativity of the sign. I will try to briefly summarise several models that I believe Leibniz pursued.[7]

The first model is the extreme proposal of a figurative representativity as the pictorial representation of the sensible, imagetic traits of the signified reality. As Leibniz says in the *Nouveaux Essais*:

> Et on pourrait introduire un caractère universel fort populaire et meilleur que le leur, si on employait de petites figures à la place des mots, qui représentassent les choses visibles par leurs traits, et les invisibles par des visibles qui les accompagnent, y joignant de certaines marques additionnelles, convenables pour faire entendre les flexions et les particules. [2, 5.379]

A procedure that Leibniz believes to be applicable to visible and also to invisible realities.

> Those (realities) which cannot be figured [pingi], like the intelligible, should however be represented by some hieroglyphic method at some time uniform and philosophical. This can be done if we do not pursue particular similarities, as the painters, the mystic and the Chinese do, but if we follow the idea of the very thing. [3, 5.216]

In second and weaker model, representativity is conceived, not as the figuration of the particular traits of the signified reality as they appear to our perception, but as its essence, that is, as the direct representation of the essential basis of those particularities. A good

[7] For a developed presentation of the various models of representativity in Leibniz and which we present here in a condensed, abridged way, see [5, pp. 174–190].

example is the project that Leibniz calls a *Characteristica Realis* [2, 7.12–13], in which the character would represent the unifying principle or "key" (*clavis*) of the multiple properties composing the signified idea.

> The name of each thing will be the key of all we should say, think and reason about that thing [. . .]. The name we will give the gold will be the key of all we can humanly know about gold, that is, by reason, and according to such an order that, by the examination of that name, we can discover which experiences should be rationally realized with that name. [2, 7.13]

There is, finally, a third model, which I propose to call expressive representativity, which aims to discover neither the figuration of sensory particularities of the signified reality (figurative representativity), nor the essential basis of those particularities (essentialist representativity), but rather the analogical and structural reproduction of the network of relations constituting the idea and its articulations.

Even if rigorously speaking all representativity is expressive for Leibniz (since expression is for Leibniz the supreme instance of any kind of relation), it is now geometry that provides the privilege model for the relation of expression. In fact, geometry is characterized by the immediate establishment of an isomorphism between the idea represented, as an ideal structure, and its schematic figuration. However, this isomorphism does not imply similarity, as the imitation of the original.[8] That which expresses does not have to be similar to the thing expressed provided that some kind of analogy can be discerned between them. The reaction of expression can even accept dissimilarity. As Leibniz says in *Essais de Théodicée*:

> The same circle can be represented by an ellipse, a parabola, a hyperbola, another circle and even a straight line or a point. Nothing seems so different, neither so dissimilar, than those pictures; however, there is an exact relation of each point to each point. [2, 6.327]

Now, it is within this expressive model that Leibniz points to a purely diagrammatic form of representativity. The most interesting indication is present in the fragment *Essais d'Analyse Grammaticale* (1683/4) where Leibniz considers it to be licit to connect the components of characters by different lines since, in this way, he says, it would be possible "to see" all those components in a simultaneous way. It is worthwhile giving the word to Leibniz:

> It would be licit to connect by different lines the parts of the character since, in this way, it would be possible to see the in paper in a simultaneous form while the sound speech vanishes and, thus, the first sound cannot refer the posterior unless the second contains something which corresponds to what it was in the first. [1, 285]

It should be noted that the recognition of the value of writing in contrast to speech which Leibniz presents here it is not limited to the stressing of writing's advantages in terms of fixing and registering or as a support for the failures of attention or of reasoning. What Leibniz emphasises here is the two-dimensionality of writing, the possibility of simultaneous grasp of the multiple relations that characters (and the ideas they signify) establish among them. That is, diagrammatic symbolism, to which the passages cited above tend, has the merit of permitting the simultaneous apprehension of the relations among the entities represented.

[8] In fact, this isomorphism implies a search, not for similarities, but rather, as Leibniz puts it, for *"un rapport constant et reglé entre ce qui se peut dire de l'une et de l'autre" [2, 2.112].*

In another passage, Leibniz introduces a slight distinction within expressive representativity. As he writes in the celebrated *Dialogus de connexione inter res et verba*, of August 1677:

> Even if characters are arbitrary, their use and connection have nonetheless something which is not arbitrary, that is to say, a certain proportion between characters and things, and, at the same time, between the diverse characters which express the relations of things among themselves. [2, 7.192]

Here, Leibniz is looking for the possibility of displacing the ideal of representativity from the character (without, however, completely abandoning it) on to the plane of syntactic relations. The meaning of a character now comes to depend either on its integration in a formal structure – its *usus et connexio* – or on the correspondence (*proportio*) between this structure and the structure of the reality which it aims to represent. Now, it is the very form of relations between characters that is seen analogically with things in the form of their relations.

Ultimately, there is – I will further argue – another last possibility, which Leibniz explores in the fragment *Characteristica universalis* of 10 August, 1679:

> The more exact are the characters, that is, the more they represent the relation of things, the bigger is their utility and if characters are able to exhibit all the relations of the things, as do the arithmetic characters that I use, then there will be nothing in the things which cannot be deduced from the characters. [3, 5.141], our emphasis

Representativity is, again, not between structures (syntactic and natural) but between elements, between the individual (containing relations as its predicates) and the character as a sign which, through its own flexion, would "exhibit" the multiple relations of the thing. It is now again each character, and not the relations between characters, that is to express the relations which individual entities contain as their predicates.

Then, within the expressive model of representativity, Leibniz pursues three approaches. In the first, the structure is purely formal; in the second, the structure becomes a source of meaning because, in itself, it imitates the real conceived as a structure; in the third the structure is absorbed by the sign and it is through the sign that it becomes expressive. If, in the first tendency (clearly diagrammatic) the semantical level tends to be entirely reduced to the syntactic, and in the second tendency (strictly structural) priority is given to analogy between the formal structures of language and the structures of the real, the third and last tendency (from the figurative and the essentialist models) recovers the requirement of the semanticity of the sign itself, that is, of a space that should be symbolically differentiated and, as such, indicative of the signified reality, open to the world that is to be known.

As I have tried to argue, it is precisely this indicative power of the sign, this openness of the symbolic system to the signified reality that Leibniz wants to safeguard at all costs. That is why he pursued so many models, attempted so many different ways of, besides operativity, gaining such a representativity.

This is a thesis that Leibniz recuperates and pulls through a long tradition coming from Plato's *Cratylus*, the speculations on Adamic language all through the Middle Ages, Renaissance and modern times.

It is a thesis that will be advanced by Peirce's conception of the iconic nature of diagrams and of the revelatory, heuristic power resulting from the relational analogy that lives in their heart.

* * *

We know that Leibniz never achieved any new language system nor did he succeed in establishing a system of characters with the openness with which he wanted them to be endowed. He just left projects, drafts, sketches, numerous fragments, *specimina and échantillons*. However, his unfinished work must be regarded, not so much as the root of recent research in the area, but as a source for clarification of the difficulties as well as of virtualities involved, as a font for the formulation of powerful hypothesis and of delicate distinctions.

References

1. Couturat, L. (ed.): Opuscules et fragments inédits de Leibniz – Extraits des manuscrits de la Bibliothèque royale de Hannover. Alcan, F., Paris (1903)
2. Gerhardt, C.I. (ed.): Die philosophischen Schriften von Gottfried Wilhelm Leibniz, 1–7. Olms, Hildesheim (1960)
3. Gerhardt, C.I. (ed.): G.W. Leibniz – Mathematische Schriften, vol. 1–7. Olms, Hildesheim (1962)
4. Leibniz, G.W.: Sämtliche Schriften und Briefe, Reihe I–VI. Akademie der Wissenschaften, Berlin, Darmstadt (1923)
5. Pombo, O.: Leibniz and the Problem of a Universal Language. Münster, Nodus Publikationen (1987)

O. Pombo (✉)
Faculdade de Ciências, Centre for Philosophy of Science of the University of Lisbon, Universidade de Lisboa, Lisboa, Portugal
e-mail: opombo@fc.ul.pt

Being Consistent About Inconsistency: Toward the Rational Fusing of Inconsistent Propositional Logic Bases

Didier Dubois and Henri Prade

Abstract This short note discusses preliminary steps in the way to a general picture of inconsistency handling in an information fusion perspective, by relating some natural approaches.

Keywords Inconsistency · Information fusion · Knowledge representation

Mathematics Subject Classification (2000) Primary 68T27, 68T30 · Secondary 03B42, 03B53, 03B60

1 Introduction

The fusion of pieces of information coming from different sources and the handling of inconsistency have motivated works in logic [2, 11] as well as in artificial intelligence (see, e.g. [7]). The issues are then the appropriate representation of the information provided by each source, and the handling of conflicts between sources, and possibly inside sources. However, it is usually assumed that the information provided by each source is consistent. In the following short discussion, we allow for the inconsistency of information provided by individual sources in a propositional format, and show that it leads to interesting pending questions.

2 Background on the Handling of Possibly Inconsistent Propositional Knowledge Bases

By a knowledge base K here, we mean a finite set of propositional formulas $K = \{\varphi_i \mid i = 1, \ldots, m\}$. Then, given two knowledge bases K and $K' = \{\psi_j \mid j = 1, \ldots, n\}$, one can define their conjunctive and disjunctive combinations, respectively, in a syntactic manner as

$$K \otimes K' = \{\varphi_i \wedge \psi_j \mid i = 1, \ldots, m, \text{ and } j = 1, \ldots, n\},$$
$$K \oplus K' = \{\varphi_i \vee \psi_j \mid i = 1, \ldots, m, \text{ and } j = 1, \ldots, n\}.$$

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_25

Note that $K \otimes K$ and $K \oplus K$ have the same semantic content as K. Moreover, $K \otimes K'$ has the same set of models (possibly empty) as the mere union $K \cup K'$ of K and K'. When K and K' are consistent, but not $K \otimes K'$, classical inference from $K \otimes K'$, or $K \cup K'$, becomes trivial, and the disjunctive combination $K \oplus K'$ is of interest. It tacitly assumes that one source is right but we do not know which source [8]. This idea may be refined by considering the maximal consistent bases of $K \cup K'$ and then combining them disjunctively. Or yet it has been proposed for a long time to compute the intersection of the sets of consequences of each maximal consistent bases of $K \cup K'$ (it was first proposed by Rescher and Manor already in 1970 [12]). In that case, we are back to the problem of non-trivial inference from a single inconsistent knowledge base.

So consider the case when a single base K is inconsistent. Let $S_1, \ldots, S_k$ be the maximal consistent subbases of K, and let $C(S)$ denote the deductive closure of a consistent propositional base S. It is easy to see that $C(S_1 \oplus \cdots \oplus S_k) \subseteq C(S_1) \cap \cdots \cap C(S_k)$. Note that in order to obtain an equality instead of this inclusion, one should replace $\oplus$ by a richer disjunction, where we not only perform disjunctions of the form $\psi^1 \vee \cdots \vee \psi^k$ with $\psi^i \in S_i$ for $i = 1, \ldots, k$, but also *all* the disjunctions of this kind where now ψ^i is any conjunction $\psi_1^i \wedge \cdots \wedge \psi_j^i$ of propositions taken in S_i (where j ranges between 1 and $\text{card}(S_i)$).

Several other approaches exist to obviate the explosive nature of classical inference.

In [4, 5], it has been proposed to associate an inconsistent propositional logic base $K = \{\varphi_i \mid i = 1, \ldots, m\}$ with its so-called paraconsistent completion K^o, defined as $K^o = \{(\varphi_i, 1, x_i) \mid i = 1, \ldots, m\}$ where $x_i = 1$ if $\exists A \subseteq K, A$ consistent, $A \vdash \neg\varphi_i$, and where $x_i = 0$ otherwise.

This method partitions K^o into the inconsistency-free part K^o_{fr} and the paraconsistent part K^o_{par} of K^o, respectively, defined by

$$K^o_{\text{fr}} = \{\varphi_i \mid (\varphi_i, 1, 0) \in K^o, i = 1, \ldots, m\}$$
$$K^o_{\text{par}} = \{\varphi_i \mid (\varphi_i, 1, 1) \in K^o, i = 1, \ldots, m\}.$$

However, reducing K to K^o_{fr}, which is consistent, is quite conservative, since while every consequence from K^o_{fr} is also a consequence of all maximal consistent subbases of K, the converse is false, as shown by the following counter-example [4]:

Example 2.1 Let $K = \{\alpha, \neg\alpha \vee \neg\beta, \beta, \neg\alpha \vee \gamma, \neg\beta \vee \gamma\}$. Then it can be verified that

- K has three maximal consistent subbases $\{\neg\alpha \vee \neg\beta, \beta, \neg\alpha \vee \gamma, \neg\beta \vee \gamma\}$, $\{\alpha, \beta, \neg\alpha \vee \gamma, \neg\beta \vee \gamma\}$, $\{\alpha, \neg\alpha \vee \neg\beta, \neg\alpha \vee \gamma, \neg\beta \vee \gamma\}$;
- $K^o_{\text{fr}} = \{\neg\alpha \vee \gamma, \neg\beta \vee \gamma\}$ is the intersection of these three bases (this is always true);
- K^o_{fr} does not entail γ;
- γ is entailed by each maximal consistent subbase of K.

Besides, the "paraconsistent completion" can be extended to layered knowledge bases in the setting of possibilistic logic [5, 9], taking into account the levels of certainty of the pieces of information.

A bolder inference mechanism that produces a set of consequences larger than $C(S_1) \cap \cdots \cap C(S_k)$ is provided by the so-called *argued inference*, which is defined as follows: ψ is an argued consequence of a (possibly inconsistent) base K if there is a consistent subset of K that entails ψ and none that entails $\neg\psi$. A set of two argued consequences is always consistent, but it is no longer the case for larger sets of argued consequences [4]. This contrasts with the fact that $C(S_1) \cap \cdots \cap C(S_k)$ is clearly consistent.

These approaches can be encompassed by extending the paraconsistent completion $\mathcal{L}^o$ to the whole language $\mathcal{L}$ of K, in the spirit of the proposal by Arieli [1]: $\forall \phi \in \mathcal{L}$:

- $\phi \in \mathcal{L}_T$ if and only if there is a consistent subset of K that entails ϕ and none that entails $\neg\phi$; and we can write $(\phi, 1, 0) \in \mathcal{L}^o$.
- $\phi \in \mathcal{L}_F$ if and only if there is a consistent subset of K that entails $\neg\phi$ and none that entails ϕ; and we can write $(\phi, 0, 1) \in \mathcal{L}^o$.
- $\phi \in \mathcal{L}_U$ if and only if there is no consistent subset of K that entails ϕ nor any that entails $\neg\phi$; and we can write $(\phi, 0, 0) \in \mathcal{L}^o$.
- $\phi \in \mathcal{L}_I$ if and only if there is a consistent subset of K that entails ϕ and another one that entails $\neg\phi$; and we can write $(\phi, 1, 1) \in \mathcal{L}^o$.

In the above definition, one can restrict to maximal consistent subbases of K. These four sets of formulas $\mathcal{L}_T, \mathcal{L}_F, \mathcal{L}_U, \mathcal{L}_I$ partition the language. It can be checked that $K^o \subset \mathcal{L}^o$ (in particular, $K^o_{\text{fr}} \subset \mathcal{L}^o_T$ and $K^o_{\text{par}} \subset \mathcal{L}^o_I$), and that $\mathcal{L}_T$ is the set of argued consequences. One can view the four annotations by pairs of Boolean values as akin to Belnap [2] epistemic truth-values, TRUE, FALSE, NONE, and BOTH, respectively. However, Belnap logic comes down to computing epistemic statuses of atomic propositions based on information from various sources, then obtaining the epistemic status of other formulas via truth-tables extending the usual ones to four values.

The paraconsistent completion and its extension also stem from possibilistic logic [9], since it comes down to computing degrees of necessity $N_i(\phi)$, and $N_i(\neg\phi)$ whereby $N_i(\phi) = 1$ if and only if $S_i \vdash \phi$; otherwise $N_i(\phi) = 0$. Necessity measures N_i are $\wedge$-decomposable functions ($N_i(\phi \wedge \psi) = \min(N_i(\phi), N_i(\psi))$) akin to KD modalities. The extended paraconsistent completion can be formally defined as

$$\mathcal{L}^o = \left\{ \left(\phi, \max_{i=1}^{k} N_i(\phi), \max_{i=1}^{k} N_i(\neg\phi) \right) : \phi \in \mathcal{L} \right\} .$$

Note that $\max_{i=1}^k N_i$ is just a general monotonic Boolean set-function (and conversely any such set function is of this form for some integer k), and the annotated formulas can be encoded in a non-regular modal logic [10].

It is easy to see that for the cautious approach using intersection, it holds that $C(S_1) \cap \cdots \cap C(S_k) = \{\phi : \min_{i=1}^k N_i(\phi) = 1\}$. The companion completion of $\mathcal{L}^o$ of the form $\mathcal{L}^{\text{cautious}} = \{(\phi, \min_{i=1}^k N_i(\phi), \min_{i=1}^k N_i(\neg\phi)) : \phi \in \mathcal{L}\}$ only contains triples of the form $(\phi, 1, 0), (\phi, 0, 1), (\phi, 0, 0)$, with $(\phi, 1, 0) \in \mathcal{L}^{\text{cautious}}$ if and only if $(\neg\phi, 0, 1) \in \mathcal{L}^{\text{cautious}}$. The set $\mathcal{L}_T^{\text{cautious}} = \{\phi : (\phi, 1, 0) \in \mathcal{L}^{\text{cautious}}\}$ is consistent, deductively closed, and equal to $C(S_1) \cap \cdots \cap C(S_k)$. Indeed, $\min_{i=1}^k N_i$ is again a necessity measure.

3 Toward Fusing Inconsistent Knowledge Bases

In this section, we turn to the case when the source bases are already inconsistent.

3.1 A Misleadingly Simple Problem

Let us consider the simple case of the fusion of the two atomic knowledge bases $K = \{p, \neg q, q\}$ and $K' = \{p, \neg p, q\}$. There are at least four ways of envisaging their fusion:

- A first idea is to merely perform the set-union $K \cup K' = \{p, \neg p, q, \neg q\}$; we get a state of contradiction both about p and about q, which is not appealing, as one may feel that some information has been lost in the process (the same result would obtain by merging $\{p, \neg p\}$ and $\{q, \neg q\}$).
- However, we can observe that all the maximal consistent subbases of K (resp. K') entail p (resp. q). So if we apply an inconsistency-tolerant inference to each of K and K' prior to merging, we may feel to be entitled to derive $p \wedge q$ in the end.
- Alternatively, consider $K \oplus K' = \{p, p \vee \neg p, p \vee q, p \vee \neg q, \neg p \vee \neg q, \neg q \vee q, \neg p \vee q, q\}$. It may be reduced to $(K \oplus K')^d = \{p, \neg p \vee \neg q, q\}$ if we delete each formula which is subsumed by another one. Although $(K \oplus K')^d$ is inconsistent, it is syntactically different from $K \cup K'$, and now all consistent subbases of $(K \oplus K')^d$ only entail $p \vee q$.
- Finally, note that p (resp. q) is inconsistency-free in K (resp. K'). This may suggest to view K as a tentative representation of a situation where the associated source asserts p, and states that it has contradictory information about q, in the sense that $(q, 1, 1) \in K^o_{\text{par}}$, which we shall denote by $\perp_q$. Similarly, K' corresponds to a situation where the associated source asserts q, and states that it has contradictory information about p, namely $\perp_p$. It may be considered reasonable to assume that $p \vee \perp_p = p$ (since $p \vee (\neg p \wedge p) = p$, and it agrees with the truth-table of disjunction for Belnap logic), we get

$$\{p, \perp_q\} \oplus \{q, \perp_p\} = \{p, q, p \vee q, \perp_p \vee \perp_q\},$$

  which may be reduced to $\{p, q, \perp_p \vee \perp_q\}$. It expresses that p and q can be asserted, and that we have contradictory information about p or about q.

The latter approach that uses a reification of local inconsistencies does not sound counterintuitive; it is just an optimistic way of fusing the information coming from the two sources stating, respectively, that p is true and information about q is conflicting (paraconsistent) and that q is true and information about p is conflicting. It is less pessimistic than the second approach which throws away much information to restore consistency of source information prior to merging.

The basic elements of the calculus implicitly underlying this latter view may be nicely pictured on Fig. 1, where three squares in solid lines (which are not squares of opposition, at least in the classical sense) are displayed. The middle of the side of a square is always

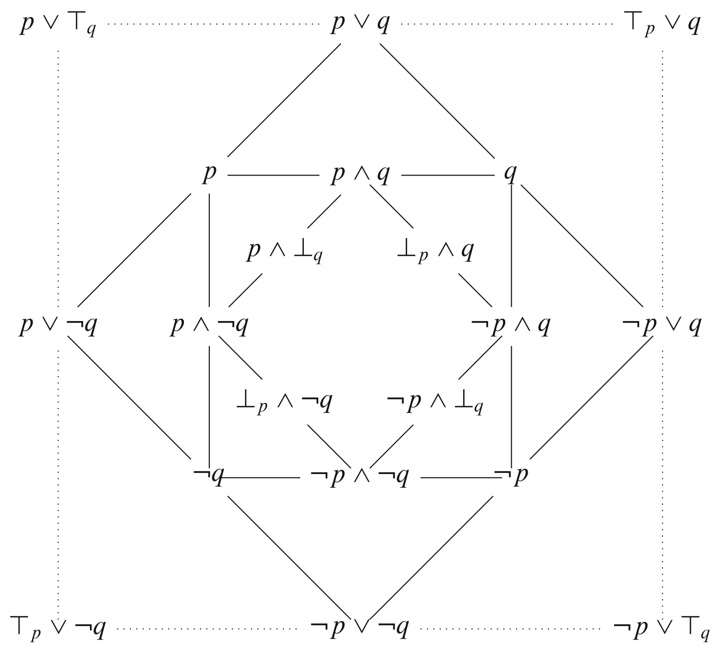


Fig. 1 Conjunctions and disjunctions in squares

the conjunct of its edges, and the vertices of the isosceles triangles are the disjunction of the edges of their opposite side.

3.2 Elements for a Discussion

Two basic issues, in order to make sense of the problem from an artificial intelligence point of view, are certainly (i) to have a proper representation of what has to be fused (which requires a clear understanding of what is supposed to be stated); and (ii) to lay bare what are the principles underlying the merging process or the conjoint exploitation of the knowledge bases reflecting the sources of information.

As it may be expected, one observes that if information is modeled by an inconsistent set of atomic formulas, one may be in one of the four states of information p, $\neg p$, $\bot_p$, and $\top_p$ regarding a propositional variable p. Here the symbol $\top_p$ would encode a local tautology $p \vee \neg p$ in a symmetric way as does $\bot_p$ for the contradiction. One may then be tempted to complete the three squares by vertices of the form $\top_p \vee q$, etc. (see the external square in dotted lines). But it is hard to defend the idea that $(\top_p \vee q) \wedge (\top_q \vee p) = p \vee q$, as, strictly speaking, it is just a tautology. However, just as the inconsistency-tolerant approaches try to confine the contradiction to a local feature not destroying knowledge bases, one could similarly, as a topic of further thought, try to have a local approach to tautologies that would avoid the opposite form of trivialization (nothing follows from a tautology).

The possibility of defining four statuses for formulas with respect to an inconsistent knowledge base may suggest to try either a multiple-valued logic approach (as in [2]), or a modal logic approach (as in [3]), or an argumentative approach (as in [1]).

The multiple-valued logic view may be considered misleading as one may argue that truth-tables are a clumsy way of handling states of information attached to atomic formulas [6]. In fact, the symbol p in the above attempt is ambiguous when put along with $\bot_p$, as writing p in the knowledge base here refers to *the fact of consistently asserting* p, without its being contradicted by other assertions: in the annotation framework outlined in the previous section, it really stands for $(p, 1, 0)$, while $\bot_p$ stands for $(p, 1, 1)$.

If we understand the four states of formulas induced by an inconsistent knowledge base K as encoded by the paraconsistent completion of the language, we see that it may lead us to questioning the edges of the second inner square (even though they look safe), namely we can doubt whether, in an inconsistent framework, the conjunction of p and q is $p \wedge q$ at all. Indeed if instead of p, $\neg p$, $\bot_p$, and $\top_p$, we write $(p, 1, 0)$, $(p, 0, 1)$, $(p, 1, 1)$, and $(p, 0, 0) \in \mathcal{L}^o$, then it is easy to check that $(p, 1, 0) \in \mathcal{L}^o, (q, 1, 0) \in \mathcal{L}^o$ do not imply $(p \wedge q, 1, 0) \in \mathcal{L}^o$, if we work with an inconsistent knowledge base [6]. This is consistent with the fact that the modal approach to capacities [10], set-functions that provide a general approach to uncertainty, is not regular (axioms K and D do not hold). This is as well at odds with Belnap logic, which assumes that the four-valued logic truth-tables extend the ones of Boolean logic. So the proposed set of squares should not be taken for granted under all approaches to inconsistency.

4 Conclusion

The aim of the above discussion (tentative) is only to draw attention on basic difficulties in the problem of handling, and more particularly fusing and inferring from propositional bases in the presence of inconsistency, whereby the syntactic framework of propositional logic seems to be ambiguous and insufficient to express what is going on.

Clearly, even if many works have been published on inconsistency and what to do with it, a number of questions remain, and one of the ambitions of this note is to make preliminary steps in the way to a general picture, relating some natural attempts. The only other purpose is to seize the occasion of this Festschrift for Jean-Yves Béziau to heartily wish him a happy birthday.

References

1. Arieli, O.: Conflict-tolerant semantics for argumentation frameworks. In: Fariñas del Cerro, L., Herzig, A., Mengin, J. (eds.) Proc. 13th Eur. Conf. Log. Artif. Intell. (JELIA'12), Toulouse, Sept. 26–28, LNCS, vol. 7519, pp. 28–40. Springer, Berlin (2012)

2. Belnap, N.D.: A useful four-valued logic. In: Dunn, J.M., Epstein, G. (eds.) Modern Uses of Multiple-Valued Logic, pp. 7–37. D. Reidel, Dordrecht (1977)
3. Béziau, J.-Y.: A new four-valued approach to modal logic. Log. Anal. **213**, 109–121 (2011)
4. Benferhat, S., Dubois, D., Prade, H.: Some syntactic approaches to the handling of inconsistent knowledge bases: A comparative study – Part 1: The flat case. Stud. Log. **58**, 17–45 (1997)
5. Benferhat, S., Dubois, D., Prade, H.: Some syntactic approaches to the handling of inconsistent knowledge bases: A comparative study – Part 2: The prioritized case. In: Orlowska, E. (ed.) Logic at Work, pp. 473–511. Physica-Verlag, Heidelberg (1998)
6. Dubois, D.: On ignorance and contradiction considered as truth-values. Log. J. IGPL **16**(2), 195–216 (2008)
7. Dubois, D., Everaere, P., Konieczny, S., Papini, O.: Approches de la révision et de la fusion d'informations. In: Marquis, P., Papini, O., Prade, H. (eds.) Panorama de l'Intelligence Artificielle. Volume 1: Représentation des Connaissances et Formalisation des Raisonnements, pp. 321–361. Cépaduès Éditions, Toulouse (2014)
8. Dubois, D., Liu, W., Ma, J., Prade, H.: Toward a general framework for information fusion. In: Torra, V., Narukawa, Y., Navarro-Arribas, G., Megías, D. (eds.) Proc. 10th Int. Conf. Model. Decis. Artif. Intell. (MDAI'13), Barcelona, Nov. 20–22. LNCS 8234, pp. 37–48. Springer, Berlin (2013)
9. Dubois, D., Prade, H.: Possibilistic logic – An overview. In: Gabbay, D.M., Siekmann, J., Woods, J. (eds.) Handbook of the History of Logic, vol. 9: Computational Logic, pp. 197–255. Elsevier, Amsterdam (2014)
10. Dubois, D., Prade, H., Rico, A.: Qualitative capacities as imprecise possibilities. In: van der Gaag, L.C. (ed.) Proc. 12th Eur. Conf. Symbol. Quant. Approaches Reason. Uncertain. (ECSQARU'13), Utrecht, July 8–10. LNCS 7958, pp. 169–180. Springer, Berlin (2013)
11. Rescher, N., Brandom, R.: The Logic of Inconsistency. Basil Blackwell, Oxford (1980)
12. Rescher, N., Manor, R.: On inference from inconsistent premises. Theory Decis. **1**, 179–219 (1970)

D. Dubois · H. Prade (✉)
Institut de Recherche en Informatique de Toulouse (IRIT), Université Paul Sabatier, Toulouse Cedex, France
e-mail: prade@irit.fr

Realist Consequence, Epistemic Inference, Computational Correctness

Giuseppe Primiero

Abstract Standard views on logical consequence stem historically from the propositions as truth-bearers tradition on the one hand, and from the assertoric standpoint on truth for propositions by proof-objects, on the other. A further step in the evolution of the notion of logical validity is represented by the formulation of correctness on computational processes, as suggested by the proofs-as-programs interpretation. We analyse this fairly recent computational interpretation of logic in view of the new principles it offers to characterize the notion of validity: execution conditions; resources accessibility; local validity; error-handling. In this new and extended sense, logical validity significantly improves the simple assertoric interpretation of correctness of non-realistic philosophies of logic. We set explicitly the connection to the notion of eventual consistency that holds for computational systems in a distributed setting.

Keywords History of logic · Philosophy of computer science · Eventual consistency

Mathematics Subject Classification (2010) Primary 03A10 · Secondary 68Q01

1 Introduction

The scope of logic is defined through the development of logical notions, relations and (meta-theoretical) results. In particular, the main relations of validity and correctness play a crucial role in defining the meaning of a logical system. These notions are historically unstable and their explanation for current logical systems in view of their historical background is most needed.

Over the centuries, the notion of logical consequence has been considered the pivotal turnaround for understanding the philosophy behind our logical systems. Currently, and for the greatest part of the contemporary approaches, the explanation of the notion of consequence is based on propositions as truth-bearers, expressing (the obtaining of) corresponding states of affairs. This tradition stems from the Bolzano–Frege–Quine–Tarski school and it has been more recently took over by the Russellian inspired theory of

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_26

truth-makers.[1] In this framework, the central notion at stake is the one of consequence characterized by validity.

Definition 1.1 (Valid consequence) A consequence from antecedents to conclusion is valid if the latter is true in the same state of affairs or models where the former are.

A different understanding of the notion of logical consequence comes from the assertoric perspective on formulas of logical systems, typical of non-realistic semantics; under this reading, the propositional content of a formula is declared true by establishing conditions for knowing it. Contents and their assertion conditions have been treated as mathematical objects in Brouwer's philosophy and given a full semantic translation in Heyting's interpretation of intuitionistic logic from around 1930. What we now call the Brouwer–Heyting–Kolmogorov semantics defines the truth of a proposition as the existence of a proof, in the sense of a proof-object of it, translated by Kolmogorov in terms of problems or tasks and their solutions.[2] Possession of a demonstration for a judgement 'A is true' expresses thus, in anti-realistic terms, the condition for being entitled to its assertion. In this framework, holding of a consequence is properly understood by reducing truth to satisfaction of assertion conditions, and so to the obtaining of an inference between judgements (the premises and the conclusion) validly asserted under the same conditions.

Definition 1.2 (Correct Inference) An inference from premises to conclusion is correct if one is entitled to assert the latter under the same conditions of the former ones.

One of the last heirs of this tradition is represented by Martin-Löf's Type Theory, based on intuitionistic logic.[3] Notoriously, intuitionistic-type theory fully endorses the formulas-as-types correspondence, known as the Curry–Howard isomorphism:[4] propositions (and equivalently sets) are expressed by types; proof-objects for the truth of a proposition (equivalent to an element in a set) are interpreted by terms in types. Expressions of the calculus are judgements of the form $A\ true$ (read: 'A is true', with A a proposition) justified by judgements of the form $a : A$ (read: 'a is a proof-object for A'). Standard intuitionistic connectives apply to propositions, defined by inference rules on proof-objects. The relation of correctness for such formulas is then generalized to truth valid under assumptions: $[A_1\ true, \dots, A_n\ true] \vdash A\ true$ (read: 'A is true, given that propositions A_1 up to A_n are true'). Type correctness (validity) is expressed by type inhabitation, as the existence of elements (proofs) defining types, and logical consequence is formulated in terms of dependent types, as the inhabitation of type A given that types $A_1, \dots, A_n$ are also inhabited.[5]

After the realist–anti-realist debate, the historical development of logical systems has generated no unique framework: along with the well-established realist and anti-realist

[1] For the notion of truth maker, see [24] and [21]; this notion was anticipated, among others, by *facts* in Moore; and as *complex* in Russell.
[2] See [19].
[3] This story is beautifully told in [39].
[4] See [38].
[5] See also [31, Cap. 1].

approaches, also a number of other paradigms have been developed, each with a slightly specific understanding of correctness and validity, like in the case of the sub-structural, dynamic, and non-monotonic frameworks. One such framework directly derives from the previously mentioned dialectic, in the form of a contemporary interpretation of the BHK semantics for typed languages. It represents the most recent and advanced legacy of the constructive and epistemic understanding of logic based on the correctness of judgemental acts as endorsed by the constructivist tradition. Such new paradigm endorses the most recent version of the Curry–Howard isomorphism: under this novel interpretation, proofs are equivalent to programs and types to specifications that programs validate. Logical types are thus used to characterize algorithms and to represent the operational semantics of programs. In the following, we shall refer to this version of the Curry–Howard isomorphism as the *computational paradigm*. The recent literature in the *Philosophy of Computer Science* has focused on the role of logical types to explore the nature of programs,[6] but still little has been said about the related notions of validity and correctness as compared to the standard logical traditions. We hope to offer in this chapter a first step in fulfilling this task.

A first understanding of the notion of validity for the computational paradigm is given in terms of syntactic correctness, i.e. correct output production when the program is executed.[7] In [37], the problem of output correctness was introduced for the first time in terms of a semantic explanation at the level of intentions:

> can computer systems satisfy *correctly* their designers aims?

In this formulation, a novel layer of analysis is offered that appears redundant when purely logical satisfaction is considered. Here correctness of the satisfaction relation refers to correct correspondence of a formally valid program with respect to the *intention* of the program's designer: does the program do what we want it to do? When considered under this computational perspective, syntactic and semantic output correctness receive a new formulation as discussed next.

Definition 1.3 (Output Correctness Problem) Can a logical system satisfy *correctly* the typing relation for which it has been formulated? Is it possible to formulate *correct* typing relations so that a certain proof system also satisfies a given validity relation?

Notice that here the qualification of correctness for the satisfaction and validity relations is not redundant: it does not refer to the formal correctness of the relation itself (in terms of holding under specified conditions); it rather points at its ability to satisfy the aim for which it is required, implemented and executed. This reading of correctness, given under the proofs-as-programs interpretation of the Curry–Howard isomorphism, significantly differs from the more traditional meaning of logical validity: it reflects typing relations that can crucially depend on the ability to formulate and access the resources needed for a given program to execute. Our main thesis is that in this new and extended sense, correctness basically differs from both the realistic and anti-realistic viewpoints mentioned above. A valid computation is satisfied by a practical aspect of term resolution, which in

[6] See, e.g., [3, 4] and [29].
[7] See [12, 17, 42].

functionally typed (programming) languages typically reduces to solving the type reconstruction problem.

Definition 1.4 (Type Reconstruction – Computational Version) Given a program p, there exists an output S and network $\mathcal{N}$ such that one can derive a formula saying that p produces an output of type S when executed in $\mathcal{N}$?

This way of determining the validity of a logical expression hinges on some essentially new problems and issues for the notion of (computational) correctness. We will present them by addressing the following questions:

1. Are conditions for execution and termination of p admissible?
2. Are resources in $\mathcal{N}$ reachable?
3. Where in $\mathcal{N}$ are processes of p valid as to satisfy S?
4. Finally, how to resolve non-valid executions of p?

In the following, we shall mainly use functional programming as the reference paradigm and typed systems as the underlying formal languages. Functional programming and typed languages are the most appropriate to analyse logical relations within the computational paradigm.

We shall start by presenting the difference in type theory between dependent types and subtypes, in order to illustrate the distinction between implicit and explicit treatment of computational information (Sect. 2).[8] At this first stage we maintain that, in view of a principle of typological non-neutrality for the data of computational acts, not all processes can be defined correct under the same conditions. We then focus on typed languages as information processing methods,[9] explaining how the standard correctness methods of proof-checking and type-inhabitation (Sect. 3) require in fact additional extensions for data-tracking (Sect. 4). Here we will suggest that the semantics of programs relies on such principles of data-tracking and localisation, which means to address the issue of resources accessibility as part of the correctness problem. The two issues combined, allow a further conclusion: given not all data are of the same logical kind, nor reachable by the same processes, *local validity* results from distinguishing between contextual and categorical correctness. Finally, we conclude by suggesting that this extension is helpful in representing levels of failure and, in turn, an appropriate notion of error resolution becomes crucial for a general understanding of logical correctness. A full taxonomy of failures can be provided, enhanced with resolution algorithms for fault-tolerant processes ([33]). As a result, correctness and so validity are restated as partially decidable properties, as it is formally stated computationally for distributed systems (Sect. 5).

[8] For the formulation of dependently typed programming languages, see, e.g., [2]; for a general introduction to the format of Martin-Löf's-type theory for programming see [28]; see [31] for a more logico-philosophical analysis.

[9] These represent a recently introduced theme in the literature in computer science, see, e.g., [11]. See also the 'Forgetting-Restore Principle' in [36] for the notion of abstraction and instantiation; its connection to types and information has been explored from a mainly historical point of view in [32].

2 Implicit and Explicit Information: Data Polymorphism

The proofs-as-programs identity underlying the semantics of typed programming languages allows us to treat the meaning of a program analogously to the constructive meaning of proofs. A basic syntax of a typed language and the corresponding intuitive semantics for programs in a network are given as follows:

- `type` $:= \{A, B, \dots\}$ are specifications that an object (program) can compute;
- `terms` $:= \{a, b, \dots\}$ are instances of programs;
- a formula of the system is an expression of the form $a : A$, meaning that a given program a holds satisfying specification A;
- an assumption $x : A$ is obtained by the process of forgetting the relevant computational information a of that expression, which can be restored at will: this means to express the ability of executing the given program for specification A, when needed;
- expressions of the language in their contextual format are read accordingly: $[x_1 : A_1, \dots, x_n : A_n] \vdash B\ type$ says that B is a valid specification under effective execution of subroutines $a_1, a_2, \dots, a_n$ for specifications $A_1, A_2, \dots, A_n$; each such type is dependent on the substitution of the free variable for the term in the preceding type.

In such a language, validity is given by establishing that every routine b of type B terminates, requiring β-reduction on the appropriate variables $\Gamma = \{x_1, \dots, x_n\}$ of types $A_1, \dots, A_n$ from which B depends. The formulation of such a context of assumptions Γ presupposes further presupposition judgements of the form $\langle A_1\ \mathit{type}, \dots, A_n\ \mathit{type}\rangle$, where each expression axiomatically declares the legality of the related specification (in absence of their values), with all A_i distinct. A specification B depending on a (set of) subroutine(s) $[a_1, \dots, a_n]$ is a valid program by stipulating that if A is a type and B is a family of types over $x : A$, then $B(x)[x : A]$ is the type of functions from A to B. Along with function formation, standard logical connectives for conjunction ($\wedge$) and disjunction ($\vee$) are interpreted for composition of programs or subroutines: the corresponding introduction rules explain the conditions under which a modular program is formulated for a complex specification.[10] From a programming perspective, the dependency relation of terms (and types) on types (and terms) allows thus to describe the behaviour of complex programs, providing a more precise typing procedure and allowing to exclude more of the bad-behaved terms. The related notion of correctness and thus of validity of programs relies therefore precisely on the ability to express and access the modules that a program requires to terminate. To do so, it is appropriate to distinguish between implicit and explicit formulation of data in the computational process.

Dependent typing as an implementation of modular programming relies on the role of witnessing information.[11] The design of modular program derivations, where the specification of B is obtained by the decomposition of sub-problem A, can either be expressed by making the entire informational load explicit, or by hiding it in terms of appropriate assumption formation rules. As an example, consider the distinction between the jus-

[10] For the standard set of computational rules for Intuitionistic Type Theory, see, e.g., [31, ch. 1].

[11] In [41], this problem is explicitly faced, to state the general thesis that it is not always practically feasible to make witnessing information explicit.

tification of a typed function to sort lists $f : [A] \Rightarrow [A]$ by an assertion of the kind $\forall x : A, \exists y : A.S(x, y)$ or rather by asserting the pair (f, p), where along with the program f one asserts the proof p that f is of type $S(x)$; hence the instance of the existential type $(\exists x : [A] \Rightarrow [A])S(x)$ which translates into: $(S(x, f(x)))[x : A]$.[12] Subtyping operations were in the first instance introduced to represent dependency relations of first-order formulae that were difficult to implement in programs.[13] By reading types as pairs (A, B), with A a type of the basic theory and B a propositional function defined over A, propositions are types of the polymorphic version of the theory which may contain quantifications over the new types.[14] By subtyping one expresses terms typed in A with property B; from a purely logical perspective, the subtype $\{x : A \mid B(x)\}$ of objects with property A that also have property B is a way to reduce the proof in B in terms of a substitution procedure such that $B[a/x]$. This operation is based on the principle of eliminating irrelevant computational information to express explicitly all and only the data needed for the program to be efficient. Subtyping can be restricted to certain sets of data, but it can result too complex to formulate all the needed subroutines in terms of subtypes. For this reason, the functional expression obtained by dependency remains the more general paradigm to express implicit computational information.

Two different aspects of computational processes are highlighted by the dependent/subtyping distinction:

1. witnessing information in the construction of a program is formally expressed by subtype formation;
2. implicit recalling of computational values is given by contextual and dependent types.

The former expresses the output production of all modules needed to satisfy the validity of a program. The latter formulation expresses the requirements needed to claim that a certain specification has an appropriate program satisfying it, even when the entire computational load is not directly accessible. This distinction justifies a ground principle of data typology: when considering recursive computational processes, not all data are of the same type. In particular, some processes require for their validity explicit computational content; other processes instead can be formulated by implicit computational content. This leads, in turn, to the following observation: contextually formulated processes, or procedures depending from implicitly given computational content depend on the accessibility and execution of the appropriate resources. The related notion of validity is thus *local* and this is reflected in the process of correctness checking.

[12] Cf. [40], pp. 256–257.
[13] See, e.g., the applications mentioned in [1] concerning systems such as LF, LEGO, *NuPrl*, *Coq*.
[14] For such extension see [28] and [40]. Under the types-as-specifications interpretation of A and B, programs satisfying specification A also satisfy specification B; in object-oriented languages, if one expects an object with interface B, an object with interface A can be safely used. See [1].

3 Correctness: From Proof-Checking to Type-Reconstruction

The satisfaction of the output-correctness problem for a dependently typed language requires well-typing of each contextual value in terms of β- and η-conversion rules for component elements of the program. Correspondingly, proof-checking proceeds first on the well-formedness of types and then performs computation on values.[15] Provided the implicit nature of the computational content of dependent types, to reconstruct values in view of proving correctness one needs to reconstruct both the *origin* and the *accessibility* of the data involved. This is equivalent to requiring runtime evaluation (or rather looping) of the programming language, referring to addresses of subroutines and their satisfaction of mobility rules, to establish where code is safely evaluated. This triggers our next observation: correctness as it is understood in the computational paradigm makes crucial use of metadata.

To explore this aspect, a general formulation of the type-checking problem is required.

Definition 3.1 (Type-checking Problem) Given a context Γ, term a and type A, is $\Gamma \vdash a : A$ a derivable expression?

where Γ is intended as the *information set* needed to enter a term a such that the program computes its type A in the current context, returning output `failure` if no such A exists. The generalization of this problem is notoriously given in polymorphic systems by requiring the notion of *type-scheme* and allowing the conversion rule, or by formulating identity conditions on proofs for uniqueness on program specification in the monomorphic version. However, subtype introduction can be seen as a declaration of the subroutine from which the main program can derive the needed values.[16] The normalization of routines in systems with subtyping is also valid only in the context of well-formedness.[17] To ensure such well-formedness of contexts, it is crucial to express location and accessibility of data. As a consequence, the decidability problem can be reformulated as the *type-reconstruction* problem.

[15] The other standard method for proof-checking consists in introducing an abstract data type (ADT) of theorems in a language of pure types, then considering an algorithm on constructions. The proof-checking problem by abstract data types presupposes a good treatability of computational information by pure types, and therefore goes along well with the idea of extending the system by subtypes. This format allows us to dispose at the programming stage with the complete set of needed witnessing information; the use of dependency relations is reduced by the formulation of an appropriate (and separate) semantic interpretation of propositions.

[16] See, as an example, the NUPRL-derivation described in [35], and further examples in the semantic justification of expressions of the form $A\ prop$ and $A\ true$.

[17] A very clear example is given by the calculus introduced in [1] which adds subtyping to λP: it first breaks dependency by providing an algorithmic version of the subtyping relation such that relates only to normal forms; then a second subtyping is introduced which splits reductions on terms and on types. The presupposition of well-formedness allows us to reduce sensibly the number of typing constraints for decidability. Even so, the required constraints might result computationally harder than (algorithmic) subtyping. Presupposition of well-formedness is avoided by interrupting circularity from subtyping to kinding, but in turn conversion on terms bounds types and transitivity holds restricted to bounded variables.

Definition 3.2 (Type-reconstruction Problem) Given a term a, there exists a type A and context Γ such that $\Gamma \vdash a : A$ is a derivable expression?

This formal definition corresponds to the informal one presented in Sect. 1: the context Γ formally expresses the set of sentences that describe all the required services in the Network $\mathcal{N}$; the term a is the formal counterpart of the program p and the type A corresponds to the output type or specification S. In this formulation of the validity for an expression, one refers not to an already given Γ, rather to its formulation, requiring explication of its content. The type-reconstruction problem is reducible to the type-checking problem in the logarithmic space for the decidable fragment of the language.[18] Hence, type-reconstruction is P-complete for example in simple-typed λ-calculi and equivalent predicative structures.[19] However, this is not the case for the formulation of second-order languages and, in particular, for languages with dependent types, i.e. where A is possibly of the form $B\ type[x : A]$, i.e. it embeds dependencies. In such a format, the type-inhabitation problem is undecidable, and so is type-reconstruction.[20] Decidability is thus affected from the dependency relation since a type-checking algorithm computes a type A for a term a and it has to assume all of its dependencies being satisfied. If dependency of $B\ type$ from some $A\ type$ is not reduced to termination, the usual behaviour of the program b for B is that of incurring in looping processes. Normalization for dependent types can be obtained by explicit assumptions checking, matching conversion for distinct occurrences of the same meta-variable and finally reduction to normal form.[21] But a type-checking algorithm will only check the base of a context and not the type of each variable in the leaves; hence one can expect to obtain soundness only for well-formed contexts.[22] Type-reconstruction represents therefore a more careful formulation of the decidability issue for the proofs-as-programs model, as it requires *explicit* account of contextual data and of their accessibility. Besides the static solution of fixing values of the letter schemas in the context so that one can always check correctness, this means that a tracking method is required for subroutines execution and termination to treat the preconditions encoded by the relation input–output of the main algorithm. Developing such explicit method requires an appropriate extension of the language, which in turn also affects the related notion of validity. This offers a new way of looking at how correctness is defined, involving accessibility and execution of computational resources.

[18] See, e.g., [38].
[19] See, e.g., [15].
[20] See, e.g., [43] and [6].
[21] See [9].
[22] See [2]. Other variants of type-checking algorithms, compromising between implicit and explicit typing, might for example allow omission of type parameters not causing any ambiguity. This is the case, for example, of the system introduced in [9], where in a sequence of type parameters those in the final part can be omitted.

4 Localizing Validity

In the literature on type-theoretical languages, undecidability by complete recursion is admitted in view of impredicative extensions, declaring legality for the set of all routines. The solution typically consists in admitting only predicative types. The weaker problem of non-terminating subroutines was treated in terms of extensional operations in [41] for the case of types without their entire informational load.[23] The same problem can be treated from an *intensional* viewpoint, adding expressive power by an extension to treat localities and indices of terminating routines. Standard logical machineries that can be used to this aim include modalities and labelling. When the propositions-as-types identity is instantiated by the proofs-as-programs paradigm, the role of these language extensions is to express conditions for termination of subroutines on input for the program to meet its specification. Modal versions of operational semantics and their calculi have been extremely useful in recent research to reason about validity and termination in distributed and concurrent computing.[24] Indices on terms specify now the locations where execution is performed. Polymorphism of specifications valid by globally terminating terms along with the type of specifications valid by locally terminating terms is now expressed by modal formulas. Under this reading, formulas receive a new interpretation:

- An expression $a_i : A$ is the declaration of value for specification A, *originating at some address* i.
- An expression $x_i : A$ is the declaration of legal code for specification A, *bounded at that given address* i.
- A judgement $[x_1 : A_i, \dots, x_n : A_n] \vdash B\ type$ says that a program valid for B is executed provided each of the routines $A_i \dots, A_n$ terminates.
- The necessity formula $\Box_i A$ is used to express *termination of* A *in every context*, i.e. being valid by subroutine reduction at runtime (*globally closed values*).
- The possibility formula $\Diamond_i A$ is used to indicate that a procedure for A *depends from locally terminating subroutines at* i; hence it is bounded to i and it is not valid in every context (*locally closed codes*).
- A contextual formula $\circ[x_1 : A_1, \dots, x_n : A_n] \vdash \circ B$ (with $\circ \in \{\Box, \Diamond\}$) expresses now the executable value of B within the resources of a network. The use of indexes allows us to distinguish among networks where all code is called-by-value, and networks where only valid code called-by-name can be used.

[23] In [41] the notion of *completely presented* type and proposition is used: a type T is called completely presented relative to a context Γ iff for some term t it holds that t inhabits a set abstracting from the content in Γ and assuming an instance in T and where the set of free variables of t is a subset of those in Γ plus the fresh free variable used for the assumption on T; a proposition ϕ will be then said completely presented relative to a context Γ iff for some term t, t is a construction for a proposition abstracted from the content in Γ and assuming ϕ, where the set of free variables of t is a subset of those in Γ plus the fresh free variable used for the assumption of ϕ. This procedure restores full computational information for the problem of type-inhabitation, reducing the form of dependent construction to a reduced normal form.

[24] See, e.g., [8, 10, 13, 18, 23, 25–27].

Semantically, truth becomes a localised concept: a judgement expressing validity for a specification A by terminating program a_i, without any additional requirement on the network will reduce to a notion of *global truth*; a judgement saying that a specification A is valid bounded to code execution within a network where the relevant program terminates reduces to a notion of *local truth*. Evaluation defines strong typing (normalisation) by reduction of expressions to globally valid counterparts, while localised ones are admissible procedural steps but may fail to produce a safe value (when called upon at wrong addresses). This makes (only) the overall safely evaluated formulas valid.

In such a calculus, one can treat cases where routines that are not globally valid become admissibile, inducing failure on execution when called at wrong locations,[25] thus reformulating the problem of programs that 'can fail at many different levels'.[26] In the following, we skip aspects of hardware-failure, which do not effect mechanical correctness from a purely logical and algorithmic point of view. Instead, we consider the language of programs and specifications as an information-processing system;[27] accordingly, the issue of failure can be for the present purposes approximated to an analysis of *levels of information failure*[28] affecting our understanding of logical validity in view of type correctness and inhabitation. Each information level expresses one step performed by a formal type-checker to control that a program gives a correct output. Routine correctness and correctness of sub-calls recursion refer to decidability from the viewpoint of type-reconstruction or type-inhabitation given in Sect. 3. In this *internal* sense of correctness, the source of failure for a system is identified by answering the following question:

> *Internal information failure*: 'at which step of program execution (routines, calls for sub-routines) does the termination process fail?'.

The tracking of termination procedures by modalities or labelling allows us to recognize the different levels of internal correctness with respect to sub-calls and recursion processes, specifying which ones may incur in local failure.

Correctness of data dependency and correctness of their retrieval refer, respectively, to the *structure* of the computation and to its *content*: formally, content is treated by considering environments (databases) from which data is extracted; structure is given by the dependency relation in contexts. The problem is here presented in its *external* form:

> *External information failure*: 'which data relevant for the computational process have not been retrieved or miss appropriate dependency, so that the termination process fail?'.

These external levels of correctness address therefore a different aspect of the output-correctness issue, and recall the 'failure-with-world' problem mentioned in [37].

Program checking under this new formulation requires selection and control over database information and mistakes override. The notion of failure as a process or a part of a process that does not satisfy some of the required logical dependency relations in the sub-routines calls or that fails to access some required data(-locations) is thus clearly central to correctness. In this sense, data retrieval procedures and dependency control processes put forward the role of process interaction in the analysis of validity.

---

25 For a full presentation of such a type system and of its procedural semantics, see, e.g., [34].

26 [37], p. 813.

27 See [30] and [16].

28 See [33].

5 Reformulating Validity

Already in [22], modern features of computing were given in terms of the following crucial properties:

1. the entity of interest is no longer a program performing computations, rather an agent in interaction;
2. the activity focused on is no longer the execution of a function but the performing of a process;
3. the way to compose parts in the computational process is no longer sequential, but parallel;
4. the resulting action of a computation is no longer the argument of the function, but a sent/received message.

In analysing computational correctness by local processes, these properties are expressed as process execution and meta-data control in terms of data typology and localization. These properties can be explicitly addressed in a computing process by stressing the role of contexts and local validity:

- programs act on different data-types (data are not all the same);
- they perform processes of meta-data control (regulating the when/how/where of data);
  - by formulating local conditions on network (processes occur as events);
  - by introducing originating locations (processes are user originated events);
- some such processes have restricted well-formedness and termination properties (not every process is explicitly formulated in its reduced form).

In a polymorphic language, some processes carry full informational load and guarantee termination and execution of distributed commands; other processes are admissible to carry external information about (mutable) conditions for termination. Allocation of meta-data for accessibility is treated by extensions such as labelling or modal operators, expressing validity over distinct locations and their interaction (union, intersection); these refer therefore to external data, such as the 'when', the 'how', and the 'where' of processes involved by the validity relation on types. In turn, termination of processes is restricted by their dependency relations.

We can now provide a more general definition of the related notion of validity:

Definition 5.1 (Computational Validity) A process P is valid iff at its execution, P is capable for any required P' of controlling

1. access to location(s) of P';
2. commands (reading/writing/exec/broadcasting) of P';
3. validity of P' (global/local w.r.t. its locations).

Operations are, in general, seen as producing an updated state of the system instantiating the interaction between processes P, P'. The principle of validity extracted from the previous analysis and conceptually redefined above is also reflected in Theoretical Computer Science in the so-called *consistency, availability and partition tolerance* principle:[29]

Theorem 5.2 (CAP Principle) One can only practically build a distributed system exhibiting any two of its three desirable characteristics: consistency, availability and partition-tolerance.[30]

As a result, the notion of consistency typically attributed to such systems is that of *eventual consistency*, i.e. dependent on the possibility that some states of the updated system will eventually lead to inconsistency; otherwise, there will be some state that might never become available; or, finally, the system is fault-tolerant with respect to some of its partitions. In the notion of eventual consistency, all the characteristics analysed above return promptly: consistency, and hence validity, is locally understood and locality requires description of failure conditions and mobility rules. A reinforced version of the same principle requires that consistency be preserved in the presence of updates, reconfigurations or failures.

From a logical viewpoint, the formation of classes of type constructors depending on terms has no computational difference with respect to the class of non-dependent types in constructive systems.[31] This means that, in principle, the reduction of a system with implicit information to one using explicit information requires no additional computational power for its realization. This corresponds to the assumption that equivalent terminations be assumed for all dependent inputs: every sequence of reductions leads to a result and every equivalent sequence of reductions leads to the same result. These are known, respectively, as the strong normalization principle and the Church–Rosser property, essential results to prove decidability of a system. But, according to the principle of eventual consistency, one is able to offer only a much weaker notion of consistent and valid state for distributed systems. In turn, the whole principle of strong normalization by reduction of procedure to their explicit output is to be reconsidered. The design of a system whose requests can add up to trillions of data at the same time (as in the case of highly distributed web-systems) has to account for cases of data unavailability, resources inaccessibility, transfer failure. The principle of replicating both data and services becomes the usual trick to reduce the possibility of failing conditions on execution and thus termination. This means that a validity relation among computational processes (as introduced in Definition 5.1, Sect. 5) is no longer understood in terms of *one* consistent execution (i.e. one consistent model), where, given a set of expressions (a database) accessible by a process P', the process P produces a unique new state of the system updated with an execution of a command of P'. In the early days of distributed systems, the unique *logical* nature of such model was reflected in the principle of *distribution transparency*: to the end-user

[29] Also known as CAP or Brewer's Theorem, the theorem originated in 2000 as a conjecture by Eric Brewer at the *Symposium on Principle of Distributed Computing*, and formally proved in 2002 in [20].

[30] By partition-tolerance in the context of distributed systems one understands the ability of the system to work even when parts of its network are isolated and no longer accessible.

[31] As shown in [5], this dependence corresponds to the coding of high-order proofs in the first-order arithmetic: both formats have the same typable terms and representable functions, equivalent strong normalization property and the set of logical theorems of a given λ-calculus is characterized proving that a logic with or without type dependence are equivalent (even if the former can prove some new theorems).

only a unique system is visible, and this requirement has priority over system stability.[32] Afterwards, data availability became the most important system feature, generating the tension between consistency, availability and tolerance expressed by the CAP theorem. The notion of consistent system is thus constrained only in view of the resulting state of each operation of the main process P, generating besides a strong version, the weak and eventual counterparts.

Definition 5.3 (Process' Strong Consistency) Every execution of P on P' generates a state of the system that is preserved under any independent reading of P on a different process P''.

Definition 5.4 (Process' Weak Consistency) Every execution of P on P' generates a state of the system that has to meet a number of pre-specified conditions to be accepted.

Definition 5.5 (Process Eventual Consistency) Every execution of P on P' generates a state of the system that, given no further execution, generates a state of the system that is preserved under any independent reading of P on a different process P''.

The latter notion can be further constrained in view of causal relations, session parameters, monotonic operations. The related notion of weak consistency can be generalized as follows for a system of processes.

Definition 5.6 (System Weak Consistency) A system is said to be weakly consistent if the number of processes involved by an update operation and the number of processes read by that operation is less or equal than the number of replicas of the updating process. Then the system is vulnerable to reading processes that have not yet received the update.

In view of this notion of system's consistency, the validity of the interaction relation among processes and the related understanding of correctness for computational events is crucially different from the notions of valid consequence and correct inference known from the logical traditions.

6 Conclusions

In this chapter, we have shown how the anti-realistic reading of logical validity can be given a truly novel interpretation based on the proofs-as-programs isomorphism, for example, as exploited in the language of dependent-type systems. Its meaning relies on a notion of contextual derivability which interprets logical validity in terms of correctness. Focusing on the role of programs in this analysis, we reviewed our current approaches to logical consequence and validity in terms of output correctness by interacting processes and shown how under this reading validity and correctness are understood as weak notions, with the exemplar case of eventual consistency for distributed systems.

[32] See [7].

A standard meaning of correctness for programs is to require that they do 'what you want it to do' ([37, p.813]). When correctness is understood as a context-dependent notion, this becomes clearly an insufficient description of what 'correct' means for a program. Our analysis stresses how program design understands correctness in a much broader sense, accounting for completeness, relevance and accessibility of data. Type-checking for correctness control is today to be seen as a very different technique than the kind criticised in [14], where it was suggested that 'because no [...] social process can take place among program verifiers, program verification is bound to fail'. Correctness is today better understood from the perspective of reliability within the standard approach of logical frameworks, reflecting the interactive nature of programs and the syntactic possibility of verification. Extensions of such structures are meant to improve methods of control and to suggest new forms of system reliability. In this way, they also offer the possibility to design a more complete understanding of correctness, and thus of validity, than the old-fashioned understanding of logical consequence, both in its realistic and in its anti-realistic vein.

Acknowledgment A preliminary version of this research was presented at the Workshop 'The Scope of Logic through History', co-located with the 4th World Congress on Universal Logic, April 03–07 2013, Rio de Janeiro Brazil. The author wishes to thank the participants to the workshop for the interaction and the financial support of the Flemish Research Foundation (FWO-Flanders).

References

1. Aspinall, D., Compagnoni, A.: Subtyping dependent types, Theoretical Computer Science **266**(1–2), 273–309 (2001)
2. Aspinall, D., Hofmann, M.: Dependent types. In: Pierce, B.C. (ed.) Advanced Topics in Types and Programming Languages, pp. 45–86. MIT Press, Cambridge, Massachusettes (2005)
3. Barendregt, H.P.: The Lambda Calculus: Its Syntax and Semantics. Studies in Logic and the Foundations of Mathematics, vol. 103, revised edition. North Holland, Amsterdam (1984)
4. Barendregt, H.P.: Lambda calculi with types. In: Abramsky, S., Gabbay, D.M., Maibaum, T.S.E. (eds.) Handbook of Logic in Computer Science, vol. 2: Background: Computational Structures, pp. 117–309. Clarendon, Oxford (1992)
5. Berardi, S.: Type Dependence and Constructive Mathematics. PhD thesis, Dipartimento di Informatica, Torino, Italy (1990)
6. Bezem, M., Springintveld, J.: A simple proof of the undecidability of inhabitation in λP. J. Funct. Program. **6**(5), 757–761 (1996)
7. Galtieri, C., Gray, J.N., Lorie, R.A., Price, T.G., Putzolu, F., Traiger, I.L., Wade, B.W., Lindsay, B.G., Selinger, P.G.: Notes on distributed databases. Report RJ2571(33471), IBM Research Laboratory (1979)
8. Borghuis, T., Feijs, L.M.G.: A constructive logic for services and information flow in computer networks. Comput. J. **43**(4), 274–289 (2000)
9. Cardelli, L.. Typechecking dependent types and subtypes. In: Boscarol, M., Aiello, L.C., Levi, G. (eds.) Proceedings of the Workshop on Foundations of logic and Functional Programming. Lectures Notes in Computer Science, vol. 306. Springer, Berlin (1987)
10. Cardelli, L., Gordon, A.D.: Anytime, anywhere. Modal logics for mobile ambients. In: Proceedings of the 27th Symposium on Principles of Programming Languages (POPL), pp. 365–377. ACM Press, New York (2000)
11. Colburn, T., Shute, G.: Abstraction in computer science. Minds Mach. **17**, 169–184 (2007)

12. Coquand, T.: An algorithm for type-checking dependent types. Sci. Comput. Program. **26**(1–3), 167–177 (1996)
13. Davies, R., Pfenning, F.: A modal analysis of staged computation. J. ACM **48**(3), 555–604 (2001)
14. De Millo, R., Lipton, J.R., Perlis, A.J.: Social processes and proofs of theorems and programs. Commun. ACM **22**(5), 271–280 (1979)
15. Dekkers, W.: Inhabitation of types in the simply typed lambda calculus. Inf. Comput. **119**(1), 14–17 (1995)
16. Fresco, N.: Explaining computation without semantics: Keeping it simple. Minds Mach. **20**, 165–181 (2010)
17. Gunter, C.A.: The semantics of types in programming languages. In: Abramsky, S., Gabbay, D.M., Maibaum, T.S.E. (eds.) Handbook of Logic in Computer Science, vol. 3: Semantic Structures, pp. 395–475. Oxford University Press, Oxford (1994)
18. Jia, L., Walker, D.: Modal Proofs as Distributed Programs. In: Programming Languages and Systems, ESOP2004. Lectures Notes in Computer Science, vol. 2986. Springer, Berlin (2004)
19. Kolmogorov, A.: Zur Deutung der intuitionistischen Logik. Math. Z. **35**, 58–65 (1932)
20. Lynch, N., Gilbert, S.: Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services (2002)
21. MacBride, F.: "Truthmakers", The Stanford Encyclopedia of Philosophy (Spring 2014 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/spr2014/entries/truthmakers/
22. Milner, R.: Turing, computing and communication. In: Smolka, S.A., Wegner, P., Goldin, D. (eds.) Interactive Computation, vol. 18, chapter 1, pp. 1–8. Springer, Berlin (2006)
23. Moody, J.: Modal logic as a basis for distributed computation. Technical Report CMU-CS-03-194, School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA, USA (2003)
24. Mulligan, K., Simons, P., Smith, B.: Truth-makers. Philos. Phenomenol. Res. **44**, 287–321 (1984)
25. Murphy, M.: Modal types for mobile code. CMU-CS-08-126. PhD thesis, School of Computer Science, Carnegie Mellon University (2008)
26. Murphy, T., Crary, K., Harper, R.: Type-Safe Distributed Programming with ML5. Lectures Notes in Computer Science, vol. 4912, pp. 108–123. Springer, Berlin (2008)
27. Murphy, T., Crary, K., Harper, R., Pfenning, F.: A symmetric modal lambda calculus for distributed computing. In: Ganzinger, H. (ed.) Proceedings of the 19th Annual Symposium on Logic in Computer Science (LICS'04), pp. 286–295. IEEE Computer Society Press (2004)
28. Nordström, B., Petersson, K., Smith, J.: Programming in Martin-Löf Type Theory. An Introduction. The International Series of Monograph in Computer Science, vol. 7. Clarendon Press, Oxford University Press, Oxford, UK (1990)
29. Pierce, B.C.: Types and Programming Languages. MIT Press, Cambridge, MA (2002)
30. Primiero, G.: An epistemic constructive definition of information. Log. Anal. **50**(200), 391–416 (2007)
31. Primiero, G.: Information & Knowledge – A Constructive Type-Theoretical Apporach. Logic, Epistemology and the Unity of Sciences, vol. 10. Springer, Berlin (2008)
32. Primiero, G.: Proceeding in abstraction. From concepts to types and the recent perspective on information. Hist. Philos. Log. **30**, 257–282 (2009)
33. Primiero, G.: A taxonomy of errors for information systems. Minds Mach. **24**, 249–273 (2014)
34. Primiero, G.: A multi-modal type system and its procedural semantics for safe distributed programming. In: Intuitionistic Modal Logic and Applications Workshop (IMLA11), Nancy (2011)
35. Salvesen, A.: On specifications, subset types and interpretation of propositions in type theory. BIT Numer. Math. **32**(1), 84–101 (1992)
36. Sambin, G., Valentini, S.: Building up a toolbox for Martin-Löf type theory: subset theory. In: Sambin, G., Valentini, S. (eds.) Twenty-five years of Constructive Type Theory, pp. 221–244. Oxford University Press, Oxford, UK (1998)
37. Smith, B.C.: Limits of correctness in computers. In: Kling, R. (ed.) Computerization and Controversy, pp. 810–825. Morgan Kaufman, Boston (1996)
38. Sørensen, M.H., Urzyczyn, P.: Lectures on the Curry–Howard Isomorphism. Studies in Logic and the Foundations of Mathematics, vol. 149. Elsevier, Amsterdam (2006)

39. Sundholm, B.G.: A double diamond of judgement. A perspective on the development of logical theory 1800–2000. In: Marion, M., Primiero, G., Rahman, S. (eds.) (Anti-)Realism: The Realism vs. Anti-realism Debate in the Age of Alternative Logics. Logic Epistemology and the Unity of Science, vol. 23, chapter 16. Springer, Berlin (2012)
40. Thompson, S.: Type Theory and Functional Programming. International Computer Science Series. Addison-Wesley, Boston (1991)
41. Turner, R.: Lazy theories of operations and types. J. Log. Comput. **3**(1), 77–102 (1993)
42. Turner, R., Eden, A.: "The Philosophy of Computer Science", The Stanford Encyclopedia of Philosophy (Winter 2014 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/win2014/entries/computer-science/
43. Urzyczyn, P.: Inhabitation in typed lambda-calculi (a syntactic approach). In: de Groote, P. (ed.) Proc. of 3rd Int. Conf. on Typed Lambda Calculi and Applications, TLCA'97, Nancy, France, 2–4 Apr. 1997. Lecture Notes in Computer Science, vol. 1210, pp. 373–389. Springer, Berlin (1997)

G. Primiero (✉)
Department of Computer Science, Middlesex University, London, UK
e-mail: G.Primiero@mdx.ac.uk

Epistemological and Ontological Paraconsistency in Quantum Mechanics: For and Against Bohrian Philosophy

Christian de Ronde

Abstract We interpret the philosophy of Niels Bohr as related to the so called linguistic turn and consider paraconsistency in the light of the Bohrian notion of complementarity. Following [16], Jean-Yves Béziau has discussed the seemingly contradictory perspectives found in the quantum mechanical double slit experiment in terms of paraconsistent viewpoints [7, 8]. This interpretation goes in line with the well-known Bohrian Neo–Kantian epistemological account of quantum mechanics. In this chapter, we put forward the idea that one can also consider, within quantum mechanics and departing from the philosophy of the danish physicist, a more radical paraconsistency found within one of the main formal elements of the theory, namely, quantum superpositions. We will argue that, rather than epistemological, the contradictions found within quantum superpositions could be interpreted as ontological contradictions.

Keywords Paraconsistency · Epistemological · Ontological · Viewpoints · Superpositions

Mathematics Subject Classification (2000) Primary 03G12 · Secondary 03A10

1 Interpreting Quantum Mechanics

Regarding its formal structure we could say that quantum mechanics (QM) seems to be a "finished theory." In terms of empirical adequacy, it provides outstanding results, its mathematical structure – developed in the first three decades of the 20th century by Werner Heisenberg, Pascual Jordan, Max Born, Erwin Schrödinger and Paul Dirac – seems able to predict any experiment we can think of. However, apart from its fantastic accuracy, even today, more than one century after its creation, its physical interpretation remains an open problem. In the standard formulation, QM assigns a quantum mechanical state to a system, but "the state" has a meaning only in terms of the outcomes of the measurements performed and not in terms of "something" which one can coherently relate to physical reality [29]. It is not at all clear, apart from measurement outcomes, what should be the interpretation of a vector in Hilbert space, in particular, and of the formal structure, in

This work was partially supported by the following grants: FWO project G.0405.08 and FWO-research community W0.030.06. CONICET RES. 4541-12 (2013-2014).

A. Koslow, A. Buchsbaum (Eds.), *The Road to Universal Logic*, Studies in Universal Logic,
DOI 10.1007/978-3-319-15368-1_27

general. If we are to ask too many questions regarding the physical meaning of the theory, problems start to pop up, and simple answers seem doomed to incoherency.

From the very beginning of the voyage, the problem of the founding fathers was to find a picture (an *anschauliche* content), a physical representation which would allow us to explain *what* QM was talking about. This idea guided the early attempt of Luis de Broglie with his matter-wave theory in 1924, and of Schrödinger with the introduction of his wave equation in 1926. It is also well known that Albert Einstein was very uncomfortable with the unclear reference of QM to physical reality. His concerns were matter of debate in the 1927 Solvay conference in Brussels and also, later on, his criticisms were exposed in the famous EPR paper written in 1935 [33]. However, very soon, it became clear that the interpretation of QM faced deep problems when attempting to provide a coherent description of physical reality. Due to this impossibility, very soon more pragmatic stances were developed. For example, Paul Dirac [32, p. 10] writes in his famous book, *The Principles of Quantum Mechanics*: "[. . .] the main object of physical science is not the provision of physical pictures, but is the formulation of laws governing phenomena and the application of these laws to the discovery of new phenomena. If a picture exists, so much the better; but whether a picture exists or not is a matter of only secondary importance." Taking distance from the physical representation of the formalism, Niels Bohr, one of the key figures in the creation and development of QM, developed a scheme in which he restricted the reference of the theory to classical phenomena. His choice, in resonance with the philosophical movements of the period, determined the problems and questioning of the future generations of physicists and philosophers of physics working on the fundamental questions about quantum theory.

2 Niels Bohr and the Linguistic Turn in Physics

Niels Bohr might have been the most influential figure who tried – and succeeded to great extent – to expel metaphysical questions from the debate regarding QM. As remarked by Arthur Fine:

> These instrumentalist moves, away from a realist construal of the emerging quantum theory, were given particular force by Bohr's so-called philosophy of complementarity; and this nonrealist position was consolidated at the time of the famous Solvay conference, in October of 1927, and is firmly in place today. Such quantum nonrealism is part of what every graduate physicist learns and practices. It is the conceptual backdrop to all the brilliant success in atomic, nuclear, and particle physics over the past fifty years. Physicists have learned to think about their theory in a highly nonrealist way, and doing just that has brought about the most marvelous predictive success in the history of science. [13, p. 1195]

At a distance from antimetaphysical construals, Einstein was a strong defender of the physical representation. As recalled by Wolfgang Pauli:

> *Einstein*'s opposition to [QM] is again reflected in his papers which he published, at first in collaboration with *Rosen* and *Podolsky*, and later alone, as a critique of the concept of reality in quantum mechanics. We often discussed these questions together, and I invariably profited very greatly even when I could not agree with *Einstein*'s view. "Physics is after all the description of reality" he said to me, continuing, with a sarcastic glance in my direction "or should I perhaps say physics is the

> description of what one merely imagines?" This question clearly shows *Einstein*'s concern that the objective character of physics might be lost through a theory of the type of quantum mechanics, in that as a consequence of a wider conception of the objectivity of an explanation of nature the difference between physical reality and dream or hallucination might become blurred. [47, p. 122]

It is important to remark that the debate which took place between Einstein and Bohr regarding physical reality can only be understood as part of the neo-Kantian tradition and discussion which was taking place in German-speaking countries at the end of the 19th and beginning of the 20th century. This discussion is very well exposed by Michael Friedman in his beautiful book, *A Parting of the Ways* [36]. From this perspective, both Einstein and Bohr were discussing from within representation, considering specifically the conditions of possibility to access phenomena. Despite the account provided by many, Einstein and Bohr were not part of the – extensively addressed in the philosophy of science literature – realist antirealist debate (see for discussion [40, 41]).

According to our reading of Bohr,[1] the preeminence of language within his own philosophical scheme can be only understood in relation to the bigger philosophical movement which was taking place in Europe and has been called the "linguistic turn."[2] So even though QM arose in relation to the criticism of Kant's epistemology – escaping from the domains of the *a priori* categories and forms of intuition –, it was soon *re-covered* by the neo-Kantian scheme of thought. It was the Danish physicist who was capable of introducing physics into the philosophical movement of the linguistic turn, shifting quantum theory from ontological concerns into epistemological ones. Niels Bohr was the first to initiate the re-turn of physics back into the domain of philosophy after the revolution produced by relativity theory and QM. Just like Immanuel Kant did with Newtonian mechanics, turning upside down the relation of power between physics and philosophy (see [45]), Bohr was able to constrain the strength of physics within the limitations imposed by that which would now play the role of the *a priori*: classical language.[3]

Niels Bohr's ideas have played a central role in the development of physics in the 20th century, placing the discipline within the main philosophical line of discussion of the period, namely, the problem of language and its relation to ontology and epistemology. The linguistic turn is a technical term in the history of philosophy according to which all problems in philosophy are problems of language. We do not claim that Bohr knew that movement or was explicitly part of it. Rather, we point to the fact that, quite independently of this movement, Bohr took for himself many of the discussions and problems involved within such philosophical stance. A clear statement regarding this point is the famous quotation by Aage Petersen. According to the long time assistant of Bohr, when asked whether the quantum theory could be considered as somehow mirroring an underlying

---

[1] We acknowledge there are almost as many interpretations of Bohr as physicists and philosophers of science. Even though the orthodoxy has been to interpret Bohr from a neo-Kantian perspective, there are also ontological interpretations of Bohr such as those proposed by Folse [34] and Dieks [30]. In particular, Dieks interprets complementarity as an ontological notion which relates "experimental situations."

[2] We take the linguistic turn to be a moment with multiple lines of philosophical investigation which can be comprised by the importance of language as a fundament. See the interesting analysis of Scavino in [53].

[3] It is then not a surprise to notice that the philosophy of Niels Bohr has been directly engaged with philosophers like Ludwig Wittgentstein (see [46]), Jaques Derrida (see [49]), and of course, Immanuel Kant (see [42] for an extensive review of the relation between Bohr and Kant's philosophy).

quantum reality Bohr [59, p. 8] declared the following: “There is no quantum world. There is only an abstract quantum physical description. It is wrong to think that the task of physics is to find out how nature is. Physics concerns what we can say about nature” The problem was then how to secure *communication*. Instead of going directly back into Kant’s *a priori*, Bohr made a detour into the realm of language, finding his cornerstone, his “clear and distinct idea” in the the language used by classical physics:

> Even when the phenomena transcend the scope of classical physical theories, the account of the experimental arrangement and the recording of observations must be given in plain language, suitably supplemented by technical physical terminology. This is a clear logical demand, since the very word *experiment* refers to a situation where we can tell others what we have done and what we have learned. [59, p. 7]

Making his point even more explicit, Bohr [59, p. 7] claimed that: “[...] the unambiguous interpretation of any measurement must be essentially framed in terms of classical physical theories, and we may say that in this sense the language of Newton and Maxwell will remain the language of physicists for all time.” In this same sense, Bohr seemed to take distance from ontological questioning. As also noted by Aage Petersen:

> Traditional philosophy has accustomed us to regard language as something secondary and reality as something primary. Bohr considered this attitude toward the relation between language and reality inappropriate. When one said to him that it cannot be language which is fundamental, but that it must be reality which, so to speak, lies beneath language, and of which language is a picture, he would reply, “We are suspended in language in such a way that we cannot say what is up and what is down. The word “reality” is also a word, a word which we must learn to use correctly” Bohr was not puzzled by ontological problems or by questions as to how concepts are related to reality. Such questions seemed sterile to him. He saw the problem of knowledge in a different light. [48, p. 11]

Bohr’s characterization of physics goes then together with his linguistically based pragmatic account:

> Physics is to be regarded not so much as the study of something a priori given, but rather as *the development of methods of ordering and surveying human experience*. In this respect our task must be to account for such experience in a manner independent of individual subjective judgement and therefor *objective in the sense that it can be unambiguously communicated in ordinary human language*. [9] (emphasis added)

However, this line of thought – which considers a physical theory in pragmatic and linguistic terms – might be regarded as ending up in a path which seems difficult to maintain, at least in the case we are still willing to state that there is something like “physical reality” of which our theories talk about. In a spirit very similar to the ideas expressed by Dirac, and after endless discussions regarding the meaning of the quantum, in the year 2000, exactly one century after the beginning of the voyage, Christopher Fuchs and Asher Peres finally took this line to its unavoidable conclusion in a paper entitled: *Quantum Theory Needs no “Interpretation”*. There, they wrote:

> [...] quantum theory does not describe physical reality. What it does is provide an algorithm for computing probabilities for the macroscopic events (“detector clicks”) that are the consequences of experimental interventions. This strict definition of the scope of quantum theory is the only interpretation ever needed, whether by experimenters or theorists. [37, p. 1]

Of course, this emphasis on prediction to the detriment of description can be severely questioned. The main objection against this instrumentalistic point of view is that the success of a theory cannot be explained, that is to say, we do not know how and why quantum physics is, in general, able to carry out predictions (and in particular with such a fantastic accuracy). Or in other words, there is no physical representation of what is going on according to the theory. Undoubtedly a "hard" instrumentalist may simply refuse to look for such an explanation, since it is in fact the mere effectiveness of a theory which justifies it, so that he may not be interested in advancing toward a justification of that effectiveness. If one takes such a position, there is nothing left to say. Just like the Oracle of Delphi provided always the right answer to the ancient Greeks, QM provides us with the correct probability distribution for every experiment we can think of.

3 The Re-Turn of Classical Metaphysics

Since the World War II, the philosophical analysis of science, and that of quantum theory in particular, has been an almost exclusive field owned by analytic philosophy. Although the analytical tradition is an inheritor – via logical positivism and logical empiricism – of a deep criticism to metaphysics, strangely enough, within analytic philosophy of physics the attempt to return to a classical metaphysical scheme of thought seems to be a recursive element, especially in the philosophy of QM during the second half of the 20th century.

Indeed, the position of Bohr, which can be very well regarded in close continuation with analytic concerns against metaphysics, was replaced after the war by approaches to QM, such as, for example, Bohmian mechanics and DeWitt's many worlds interpretation, which recovered the classical metaphysical foundation of thought and understanding. While Bohr attempted to analyze the logical structure of the theory and concentrated on the analysis of phenomena, these new lines of research intended "to restore a classical way of thinking about *what there is*" [4, p. 74]. It seems in this case a bit ironic that the aversion professed by many philosophers of physics – which are part of the analytic tradition – to Bohr's ideas does not recognize the profound connection of his thought to analytic philosophy itself. These same philosophers choose – knowingly or not – for metaphysical schemes going against their own tradition. In the case of many worlds interpretation, the metaphysical step goes as far as to propose non-observable worlds in order to explain the formal aspects of QM. Also, from a metaphysical point of view, the many worlds attempt seems to end up in an extreme violation of Ockham's principle: "Entities are not to be multiplied beyond necessity."[4] In the case of Bohmian mechanics, the metaphysical presuppositions got as far as postulating that QM must talk about classical particles with definite trajectories. Bitbol notices in this respect [6, p. 8] that: "Bohm's original theory of 1952 is likely to be the most metaphysical (in the strongest, speculative, sense) of all

[4] Although Lev Vaidman [56] claims that: "in judging physical theories one could reasonably argue that one should not multiply physical laws beyond necessity either (such a verion of Ockham's Razor has been applied in the past), and in this respect the many worlds interpretations is the most economical theory. Indeed, it has all the laws of the standard quantum theory, but without the collapse postulate, the most problematic of physical laws." One could argue however, that due to the existence of modal interpretations, which are also no collapse interpretations and share the same formal structure as many worlds, there is no clear argument why one should be forced into this expensive metaphysical extension.

readings of QM. It posits free particle trajectories in space-time, that are unobservable in virtue of the theory itself." Furthermore, that which should play the role of space-time in the mathematical formalism varies its dimension with the addition or substraction of particles breaking the initial attempt to recover trajectories in space-time. It is not at all clear that these kinds of attempts bring more solutions than problems.

The Danish physicist remained not only agnostic regarding the metaphysical concerns raised especially by Einstein, but also those of Heisenberg and Pauli. He tried by all means to restrict his analysis to the empirical data as exposed by classical physical theories and language, and not go beyond the interpretation of the formalism in terms of a new conceptual scheme. According to Bohr [59, p. 7] "it would be a misconception to believe that the difficulties of the atomic theory may be evaded by eventually replacing the concepts of classical physics by new conceptual forms." Contrary to this approach, Bohmian mechanics and many worlds interpretations, two of the most important interpretational lines of research in the present philosophy of QM, compose their analysis with heavy metaphysical commitments based to a great extent on the actualist picture put forward by Newtonian physics (see for discussion [28]). Rather than starting from the analysis of the formal structure of the theory, the metaphysical presuppositions constitute the very foundation and center of gravity of such interpretations. Some of these interpretations even attempt in some cases (e.g., Bohmian mechanics and GRW theory) to change the formalism in order to recover – at least some of – our classical (metaphysical) conception of the world.

4 Complementarity and Paraconsistency

Paraconsistent logics (PL) are the logics of inconsistent but nontrivial theories. The origins of PL go back to the first systematic studies dealing with the possibility of rejecting the PNC. PL was elaborated, independently, by Stanislaw Jaskowski in Poland, and by Newton da Costa in Brazil, around the middle of the last century (on PL, see, for example, [17]). A theory T founded on the logic L, which contains a symbol for negation, is called inconsistent if it has among its theorems a sentence A and its negation $\neg A$; otherwise, it is said to be consistent. T is called trivial if any sentence of its language is also a theorem of T; otherwise, T is said to be nontrivial. In classical logics and in most usual logics, a theory is inconsistent if and only if it is trivial. L is paraconsistent when it can be the underlying logic of inconsistent but nontrivial theories. Clearly, no classical logic is paraconsistent.

In the context of QM, da Costa and Krause have put forward [16] a PL in order to provide a suitable formal scheme to consider the notion of complementarity introduced by Bohr in 1927 during his famous "Como Lecture."[5] The notion of complementarity

[5] The logical understanding of complementarity has an interesting history which goes back to Carl Friedrich von Weizsäcker who wrote an article named: "Komplementarität und Natuurwissenschaft" [58] for Bohr's 70th birthday. In this article he explained the concept of complementarity in terms of *parallel complementarity* and *circular complementarity*. The difficulties to understand complementarity is exposed by a rectification added at the end of the same paper in which von Weizsäcker explains that he received

was developed in order to consider the contradictory classical representations found in the double-slit experiment; i.e., the representation provided by the notions of "wave" and "particle." According to Bohr [16, p. 103]: "We must, in general, be prepared to accept the fact that a complete elucidation of one and the same object may require diverse points of view which defy a unique description." The starting point of analysis was for Bohr the classical description of experimental arrangements univocally represented by classical language (with the aid of physics). Bohr considered the wave-particle duality present in the double-slit experiment as expressing the most important character of quantum theory. The resolution of this duality was provided via his own notion of complementarity. Bohr's agenda was focused in fulfilling the consistency requirements of the quantum formalism to apply the well known classical scheme. In this respect, Heisenberg's principle was only considered as providing the limits of certainty and applicability of classical concepts as such. The classical scheme would then remain that which secured the knowledge provided by QM and analogously, Heisenberg's *uncertainty relations* that secured the knowledge provided by the more general principle of complementarity.[6] As Leon Rosenfeld makes the point:

> Bohr wanted to pursue the epistemological analysis one step further [than Heisenberg], and in particular to understand the logical nature of the mutual exclusion of the aspects opposed in the particle-wave dualism. From this point of view the indeterminacy relations appear in a new light. [...] The indeterminacy relations are therefore essential to ensure the consistency of the theory, by assigning the limits within which the use of classical concepts belonging to the two extreme pictures may be applied without contradiction. For this novel logical relationship, which called in Bohr's mind echoes of his philosophical meditations over the duality of our mental activity, he proposed the name "complementarity," conscious that he was here breaking new ground in epistemology. [59, p. 59]

In [16], the proposal was to go further into the notion of *complementary theories*.[7]

> [...] we shall say that a theory T admits complementarity interpretation, or that T is a C-theory, if T encompasses "true" formulas α and β (which may stand for Jammer's D_1 and D_2 respectively) which are "mutually exclusive" in the above sense, for instance, that their conjunction yields to a strict contradiction if classical logic is applied. In other words, if $\vdash$ is the symbol of deduction

a letter from Bohr expressing that complementarity can be only defined with respect to phenomena, and as the Schrödinger wave equation is just an abstract magnitude of calculus and it does not designate in itself any phenomena, such circular complementarity is by no means possible and only parallel complementarity should be taken into account.

[6] It is important to notice that Heisenberg's relations can be directly derived from the mathematical scheme of the theory, as a direct consequence of the quantum postulate. Today, we have more elements to make precise the relation between both principles, see for example, the analysis of Pekka Lahti in his thesis [44]. As remarked by J. Hilgevoord and J. Uffink [39]: "On the one hand, Bohr was quite enthusiastic about Heisenberg's ideas which seemed to fit wonderfully with his own thinking. Indeed, in his subsequent work, Bohr always presented the uncertainty relations as the symbolic expression of his complementarity viewpoint. On the other hand, he criticized Heisenberg severely for his suggestion that these relations were due to discontinuous changes occurring during a measurement process. Rather, Bohr argued, their proper derivation should start from the indispensability of both particle and wave concepts. He pointed out that the uncertainties in the experiment did not exclusively arise from the discontinuities but also from the fact that in the experiment we need to take into account both the particle theory and the wave theory."

[7] This idea of considering complementary theories goes in line with the complementary descriptions approach proposed in [18].

> of classical logic, then, α and β being complementary, we have $\alpha, \beta \vdash \gamma \wedge \neg\gamma$ for some γ of the language of T. [16, p. 111]

The proposal of da Costa and Krause was also considered by Jean-Yves Béziau [8] stressing in his case the notion of *viewpoint* and taking into account the square of opposition. According to Béziau:

> In modern physics we have a subcontrary opposition between wave and particle in the sense that the proposition "K is a particle" and "K is a wave" can both be true but cannot both be false. [...] [O]ne may say that with particle and wave there is an opposition because something cannot be at the same time a particle and a wave, due to the very nature of wave and particle, in the same sense that something cannot be a square and a circle. But why then can we say that "K is a particle" and "K is a wave" can both be true but not that "K is a circle" and "K is a square" can both be true? In fact it is also possible to say that these two geometrical propositions are both true, but from a different perspective, which is not the usual flat one. [8]

Béziau then continues to analyze the situation taking into account more explicitly the position of Bohr and his complementarity approach.

> [Bohr] argues that there are no direct contradiction: from a certain point of view "K is a particle," from another point of view "K is a wave," but these two contradictory properties appear in different circumstances, different experiments. Someone may ask: what is the absolute reality of K, is K a particle or is K a wave? One maybe has to give away the notion of objective reality. [8]

At this point, it becomes clear that the paraconsistency implied by Béziau relates to the perspective assumed by the observer. The contradiction appears through the choice of the subject between the different experimental set-ups. Following this idea, he develops a logical theory where the central concept is the concept of *viewpoint*. The price to pay, as clearly acknowledged by Béziau, is to abandon the notion of objective physical reality. In the following section, we will argue that paraconsistency can also help us to develop a new idea of physical reality that would allow us to consider QM as a physical theory providing an objective representation of a world outside there.

5 Quantum Superpositions and Paraconsistency

In classical physics, every physical system may be described exclusively by means of its *actual properties*, taking "actuality" as expressing the *preexistent* mode of being of the properties themselves, independently of observation – the "pre" referring to its existence previous to measurement. Each system has a determined state characterized mathematically in terms of a point in phase space. The change of the system may be described by the change of its actual properties. Potential or possible properties are considered as the points to which the system might arrive in a future instant of time. As Dieks [31, p. 124] makes the point: "In classical physics the most fundamental description of a physical system (a point in phase space) reflects only the actual, and nothing that is merely possible. It is true that sometimes states involving probabilities occur in classical physics: think of the probability distributions ρ in statistical mechanics. But the occurrence of possibilities in such cases merely reflects our ignorance about what is actual." It is then important to recognize that the main character which makes possible this physical description in terms of actual properties, and more in general, in terms of an actual state of affairs is the fact that the mathematical structure allows a global valuation of all properties [29].

In QM, the representation of the state of a system is given by a ray in Hilbert space H and physical magnitudes are represented by operators on H that, in general, do not commute. As a consequence, the Kochen–Specker theorem precludes the possibility of a global valuation of all properties independently of the context [43]. It then becomes difficult to affirm that all quantum magnitudes are *simultaneously preexistent*. In order to restrict the discourse to different sets of commuting magnitudes, different complete sets of commuting operators (CSCO) have to be chosen. The choice of a particular representation (given by a CSCO) determines the basis in which the observables diagonalize and in which the ray can be expressed. Thus, the ray can be written as different linear combinations of states:

$$\alpha_i|\varphi_i^{B1}\rangle + \alpha_j|\varphi_j^{B1}\rangle = |\varphi_q^{B2}\rangle = \beta_m|\varphi_m^{B3}\rangle + \beta_n|\varphi_n^{B3}\rangle + \beta_o|\varphi_o^{B3}\rangle \quad (5.1)$$

Each linear combination of states is also called a *quantum superposition*. The Born interpretation tells us that the numbers that accompany each state in square modulus compute the probability of finding that particular state. It is also well known that such probability cannot be interpreted in terms of ignorance [51, 52]. As remarked by Schrödinger in a letter to Einstein:

> It seems to me that the concept of probability [in QM] is terribly mishandled these days. Probability surely has as its substance a statement as to whether something is or is not the case – of an uncertain statement, to be sure. But nevertheless it has meaning only if one is indeed convinced that the something in question quite definitely is or is not the case. A probabilistic assertion presupposes the full reality of its subject. [12, p. 115]

If we consider a typical Stern–Gerlach experiment, given a spin $\frac{1}{2}$-system whose state is $\frac{1}{\sqrt{2}}(|\uparrow_x\rangle + |\downarrow_x\rangle)$, orthodox QM tells us that we shall obtain either a "click" in the upper part of the screen with probability 0.5 (which relates to the state $|\uparrow_x\rangle$), or, a "click" in the bottom part of the screen also with probability 0.5 (which relates to the state $|\downarrow_x\rangle$). However, it is not clear at all what *is* the state before the measurement since, on the one hand, one cannot claim due to the formalism of the theory that the measurement discovers a preexistent actual reality, and on the other hand, actuality itself precludes the existence of the seemingly exclusive possibilities. According to our interpretation, there are good reasons to claim that, in the just mentioned example, both states ("$|\uparrow_x\rangle$" and "$|\downarrow_x\rangle$") should be regarded as existent – rather than exclusive possibilities. As a matter of fact, quantum superpositions "evolve" according to the Schrödinger equation of motion, "interact" with other superposition states – creating the famous entanglement of multiple states – and can "be predicted" according to the rules of QM. But if, within a physical theory there is a mathematical expression which allows us to calculate the evolution of its terms, its interaction with other mathematical expressions of the type and predict its possible results, after or before the interaction, then it becomes reasonable to claim that there is something "physically real" about such a mathematical expression.

The problem at this point is that while these two states, $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$, seem to exist before measurement as somehow contradictory states, the realm of actuality denies – by definition – the existence of contradictions. This weird fact about quantum superpositions was cleverly explained by Schrödinger through his famous cat experiment which produced a zombie cat half-dead and half-alive. At this point, one might consider two different interpretative strategies to solve the paradox. The first, to find the way to interpret quantum superpositions escaping contradictions, forcing the actual realm of existence in order to interpret the formalism – e.g., as many worlds or the modal interpretation of Dieks attempts to do. The second, more radical path, is to develop a realm of existence which is not that of actuality. It is this latter possibility which we find more interesting.

In [14], Newton da Costa together with the author of this chapter argued in favor of the possibility of considering quantum superpositions in terms of a paraconsistent approach. We discussed the idea that, even though most interpretations of QM attempt to escape *contradictions*, there are many reasons that indicate it could be worth while to assume that the terms in a quantum superposition can be, in principle, contradictory – as in the just mentioned example. It should be also stressed that the paraconsistent approach to quantum superpositions (PAQS) does not present an interpretation of QM but just a way to account for quantum superpositions in terms of paraconsistent logics. This formal approach might lead to different interpretations. However, one may also recognize that intuitively, the approach might favor an interpretation in which superpositions are considered as contradictory physical existents.

In [1], Arenhart and Krause raised several arguments against the PAQS, one of the main concerns having to do with the consideration of quantum superpositions as contradictory existents. In [25], the author of this paper argued, firstly, that the obstacles presented by Arenhart and Krause are based on a specific metaphysical stance, which we characterized in terms of what we called the orthodox line of research (OLR). Secondly, that this is not necessarily the only possible line, and that a different one, namely, a constructive metaphysical line of research (CMLR) provides a different perspective in which the PAQS can be regarded as a valuable prospect that could be used by different interpretations of QM.

It is interesting to point out that the OLR implicitly embraced by Krause and Arenhart reflects the Bohrian perspective toward the problem of interpretation in QM. Indeed, the OLR has always debated within the limits imposed by the Danish physicist – this has been Bohr's true success. As we have discussed in detail in [25], this path can be condensed in two main metaphysical presuppositions which block the conceptual development of quantum superpositions. The first presupposition relates to what Bohr called the *correspondence principle*, an idea which has been later on reconsidered in the literature in terms of what is known to be the quantum to classical limit [10].

1. **Quantum to Classical Limit:** The principle that one can find a bridge between classical mechanics and QM, i.e., that the main notions of classical physics can be used in order to explain quantum theory.

The second metaphysical principle can also be traced back to Bohr's claim that physical experience needs to be expressed exclusively in terms of classical language [11]. As a matter of fact, if one considers the core of the classical physical and metaphysical representation of the world, one is then stuck with two main concepts: "entity" and "actuality" (as a mode of existence). In QM, one can also encounter these metaphysical notions as basic elements of any interpretation.

2. **Quantum Systems and Actuality:** The principle that one needs to presuppose the metaphysics of entities together with the mode of being of actuality in any interpretation of QM.

The idea proposed in [14] to introduce paraconsistency in order to account for quantum superpositions might be regarded as opening the possibility of supplementing such formal scheme with new nonclassical physical notions. It should also be clear, however, that this is not necessarily the only option, since the PAQS might be regarded as only a formal proposal. But in case we would be willing to step into the realm of metaphysics and advance with an interpretation that accepts *contradictory existents*, then we must also come up with a realm of existence different from actuality, breaking the equation that has ruled physics since Newton: Actuality = Reality. One such proposal has been put forward in [23].

As we have argued extensively in [25], if we now go into the problem of interpretation, the idea presented by the PAQS can be supported by the CMLR which imposes a different perspective towards QM, its problems, and the questions that need to be answered. Taking into account the need to provide a coherent physical interpretation of QM, our CMLR is based on three main presuppositions:[8]

1. **Closed Representational Stance:** Each physical theory is closed under its own formal and conceptual structure providing access to a specific set of phenomena. The theory provides the constraints to consider, explain and understand such physical phenomena.
2. **Formalism and Empirical Adequacy:** The formalism of QM is able to provide (outstanding) empirically adequate results. Empirical adequacy determines the success of a theory and not its (metaphysical) commitment to a certain presupposed conception of the world. The problem is not to find a new mathematical scheme, on the contrary, the "road signs" point in the direction that *we must stay close to the orthodox quantum formalism*.
3. **Constructive Stance:** To learn about what the formalism of QM is telling us about reality, we might be in need of *creating new physical concepts*.

What is needed according to the CMLR is a radical inversion of orthodoxy and its problems. According to this inversion, for example, the question of contextuality is not a problem which we need to escape but rather a central feature that any interpretation of QM should respect. Also, the nonseparable character of QM, its specificity with respect to identity and individuality, its indeterminate and indeterministic aspects, etc., should be all considered – rather than obstacles – as the main roadsigns that should guide is in the development of a coherent interpretation of the theory. Going back to the meaning of quantum superpositions, instead of considering *the measurement problem*, we should focus instead on the analysis of what we have called *the superposition problem* (see for discussion [25]). The constructive stance assumes the radical possibility of considering a different mode of existence. Elsewhere [20, 21, 23], we have put forward a notion of *ontological potentiality* which allows us to discuss an independent realm of existence which – contrary to the orthodox approach to potentiality – cannot be reduced to actuality. Indeed, the PAQS,

[8] This proposal was put forward and discussed in [20, pp. 56–57]. See also [24].

properly supplemented by the CMLR, allows us to interpret quantum superpositions in terms of potential contradictory existents.

6 Epistemological and Ontological Contradictions

From the above discussions, it is interesting to notice that the use of paraconsistency in QM can be considered from two very distinct perspectives, according to the philosophical stance that one assumes in order to approach interpretational issues. On the one hand, we have the paraconsistency implied by the epistemological approach to QM which deals with the "contradiction" between multiple classical representations. As remarked by da Costa and Krause:

> It should be emphasized that our way of characterizing complementarity does not mean that complementary propositions are always contradictory, for α and β above are not necessarily one the negation of the other. However, as complementary propositions, we may derive from them (in classical logic) a contradiction; to exemplify, we remark that "x is a particle" is not the direct negation of "x is a wave," but "x is a particle" *entails* that x is not a wave. This reading of complementarity as not indicating strict contradiction, as we have already made clear, is in accordance with Bohr himself [. . .] [16]

This could be understood as an *epistemological contradiction* since it deals with different perspectives or, in terms of Beziau, "complementary viewpoints." However, the paraconsistency discussed within the PAQS implies a more radical stance regarding the meaning of contradictions. Indeed, as we have mentioned above, quantum superpositions – according to the PAQS – seem to open the door to consider "contradictory existents." Thus, one might argue that the PAQS attempts to put forward the introduction of *ontological contradictions*. As remarked by Aristotle: "It is impossible for the same thing to belong and not to belong at the same time to the same thing and in the same respect." [Metaph. IV 3 1005b1920]. It seems that this is not the case of epistemological contradictions which deal with different, rather than contradictory, representations. It makes no sense to say that "a wave is contradictory to a particle." On the other hand, ontological contradictions do talk about *the same property in the same respect and at the same time*. If we consider our earlier Stern–Gerlach experiment the state of affairs is described by the following quantum superposition: $\frac{1}{\sqrt{2}}(| \uparrow_x\rangle + | \downarrow_x\rangle)$, which includes the propositions: "spin up in the x-direction" and "spin down in the x-direction." Both propositions make reference to the same property of spin in a contradictory manner, and at the same time.

The introduction of the realm of potentiality reconfigures in itself the meaning of contradiction which has always been considered in terms of the actual realm. Indeed, once we accept the idea that we can have existents in the potential realm, we can go further and understand contradictions as related to the path between the potential and the actual realms. If we now take into account the square of opposition we must consider the following set of definitions:

Contradiction Propositions: α and β are *contradictory* when both cannot be true and both cannot be false.

Contrariety Propositions: α and β are *contrary* when both cannot be true, but both can be false.

Subcontrariety Propositions: α and β are *subcontraries* when both can be true, but both cannot be false.

Subaltern Propositions: α is *subaltern* to proposition β if the truth of β implies the truth of α.

The idea that potentiality determines a contradictory realm goes back to Aristotle himself who claimed that contradictions find themselves in potentiality. Of course, as remarked by Arenhart and Krause, the square of opposition discusses about actual truth and falsehood. Thus, potentiality is not considered in terms of a mode of existence but rather as mere logical possibility. The introduction of an ontological realm of potentiality changes things drastically. The interesting question is if our representation of quantum superpositions in terms of potential contradictory propositions is compatible with the square. We believe that it is easy to see that such is the case provided special attention is given to the realms involved in the discussion.

Truth and falsehood have always been considered in relation to actuality, since in the orthodox view this is the exclusive realm which imposes the limits of what can be understood as real. Contrary to the actualist scheme, our notion of ontological potentiality is completely independent of actuality. Thus, it makes perfect sense to extend "truth" and "falsity" to this mode of being. We have investigated this possibility in [26]. Our redefinition of truth and falsehood with respect to potentiality escapes any subjective choice and regains an objective description of physical reality. The price to pay is to abandon the idea that everything needs to be defined in terms of the actual realm.

Consider we have a Stern-Gerlach apparatus placed in the x-direction. If we have the following quantum superposition: $\alpha \mid \uparrow_x\rangle + \beta \mid \downarrow_x\rangle$, this means we have "spin up in the x-direction, $\mid \uparrow_x\rangle\langle\uparrow_x \mid$, with probability $|\alpha|^2$" and "spin down in the x direction, $\mid\downarrow_x\rangle\langle\downarrow_x\mid$, with probability $|\beta|^2$" which can be actualized. Is it *contradiction* or *contrariety* the best notion suited to account for such existent possibilities[9] in this quantum experiment? Given this quantum superposition, it is clear that both actualizations (elementary processes) "$\mid \uparrow_x\rangle\langle\uparrow_x \mid$" and "$\mid \downarrow_x\rangle\langle\downarrow_x \mid$" *cannot be simultaneously "true" in actuality*, since only one of them will become actual; it is also the case that both actualizations (elementary processes) "$\mid \uparrow_x\rangle\langle\uparrow_x \mid$" and "$\mid \downarrow_x\rangle\langle\downarrow_x \mid$" *cannot be simultaneously "false" in actuality*, since when we measure the quantum superposition we know we will obtain either the elementary process "spin up in the x-direction," "$\mid \uparrow_x\rangle\langle\uparrow_x \mid$," or the elementary process "spin down in the x-direction," "$\mid \downarrow_x\rangle\langle\downarrow_x \mid$." As we know, given a measurement on the quantum superposition, $\alpha \mid \uparrow_x\rangle + \beta \mid \downarrow_x\rangle$, one of the two terms will become actual (true), while the other term will not be actual (false), which implies that *both cannot be false*. This experimental fact goes clearly against the proposal of Arenhart and Krause of considering contrary propositions instead of contradictory ones.

[9] We have developed an interpretation in which such existent possibilities are discussed in terms of the notion of *power*. See [21, 23, 26].

7 Final Remarks

In this chapter, we have discussed different approaches regarding the meaning and use of paraconsistency in QM. We have shown how this subject, from very different perspectives, both epistemological and ontological, can provide an interesting discussion and development of the theory. We hope to continue this analysis in future papers.

Acknowledgment The author would like to thank Newton da Costa, Jean-Yves Béziau, Graciela Domenech, Hector Freytes, Décio Krause and Jonas Arenhart for many discussions regarding the subject of paraconsistency in quantum mechanics.

References

1. Arenhart, J.R., Krause, D.: Oppositions in Quantum Mechanics. In: Bèziau, J.-Y., Gan-Krzywoszynska, K. (eds.) New dimensions of the square opposition, pp. 337–356. Philosophia Verlag, Munich (2014)
2. Arenhart, J.R., Krause, D.: Contradiction, Quantum Mechanics, and the Square of Opposition. Logique et Analyse, accepted.
3. Arenhart, J.R., Krause, D.: Potentiality and Contradiction in Quantum Mechanics. In: Koslow, A., Buchsbaum, A. (eds.) The Road to Universal Logic (volume II), Springer (2015, in press)
4. Bacciagaluppi, G.: Topics in the modal interpretation of quantum mechanics. Doctoral Dissertation, University of Cambridge, Cambridge (1996)
5. Bernien, H., Hensen, B., Pfaff, W., Koolstra, G., Blok, M.S., Robledo, L., Taminiau, T.H., Markham, M., Twitchen, D.J., Childress, L., Hanson, R.: Heralded entanglement between solid-state qubits separated by three metres. Nature **497**, 86–90 (2013)
6. Bitbol, M.: Reflective Metaphysics: Understanding quantum mechancis from a Kantian standpoint. Philosophica **83**, 53–83 (2010)
7. Béziau, J.-Y.: The Power of the Hexagon. Log. Univers. **6**, 1–43 (2012)
8. Béziau, J.-Y.: Paraconsistent logic and contradictory viewpoints. Rev. Bras. Filos. **241** (2014, to appear)
9. Bohr, N.: The unity of human knowledge. In: Philosophical Writings of Neils Bohr, vol. 3. Ox Bow Press, Woodbridge (1960)
10. Bokulich, A.: Bohr's Correspondence Principle. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Spring 2014 Edition). (2014). http://plato.stanford.edu/archives/spr2014/entries/bohr-correspondence/
11. Bokulich, P., Bokulich, A.: Niels Bohr's generalization of classical mechanics. Found. Phys. **35**, 347–371 (2005)
12. Bub, J.: Interpreting the Quantum World. Cambridge University Press, Cambridge (1997)
13. Curd, M., Cover, J.A.: Philosophy of science. In: Norton and Company (eds.) The Central Issues. Cambridge University Press, Cambridge (1998)
14. da Costa, N., de Ronde, C.: The Paraconsistent logic of quantum superpositions. Found. Phys. **43**, 845–858 (2013)
15. da Costa, N., de Ronde, C.: The Paraconsistent Approach to Quantum Superpositinos Reloaded: Formalizing Contradictory Powers in the Potential Realm, preprint (2014)
16. da Costa, N.C.A., Krause, D.: The logic of complementarity. In: van Benthem, J., Heinzmann, G., Rebuschi, M., Vesser, H. (eds.) The Age of Alternative Logics: Assessing Philosophy of Logic and Mathematics Today, pp. 103–120. Springer, Berlin (2006)
17. da Costa, N.C.A., Krause, D., Bueno, O.: Paraconsistent logics and paraconsistency. In: Jacquette, D. (ed.) Handbook ofthe Philosophy of Science (Philosophy of Logic), pp. 791–911. Elsevier (2007)

18. de Ronde, C.: Complementary Descriptions (Part 1). Los Alamos Archive. arXiv: quant-ph/0507105v1 (2005)
19. de Ronde, C.: For and against metaphysics in the modal interpretation of quantum mechanics. Philosophica **83**, 85–117 (2010)
20. de Ronde, C.: The Contextual and Modal Character of Quantum Mechanics: A Formal and Philosophical Analysis in the Foundations of Physics. PhD dissertation, Utrecht University (2011)
21. de Ronde, C.: La noción de potencialidad en la interpretación modal de la mecánica cuántica. Sci. Stud. **1**, 137–164 (2012)
22. de Ronde, C.: Quantum Superpositions and Causality: On the Multiple Paths to the Measurement Result. Philosophy of Science Archive. http://philsci-archive.pitt.edu/10049/ (2013)
23. de Ronde, C.: Representing Quantum Superpositions: Powers, Potentia and Potential Effectuations. Philosophy of Science Archive. http://philsci-archive.pitt.edu/10155/ (2013)
24. de Ronde, C.: The problem of representation and experience in quantum mechanics. In: Aerts, D., Aerts, S., de Ronde C. (eds.) Probing the Meaning of Quantum Mechanics: Physical, Philosophical and Logical Perspectives, pp. 91–111. World Scientific, Singapore (2014)
25. de Ronde, C.: A Defense of the Paraconsistent Approach to Quantum Superpositions (Answer to Arenhart and Krause). Philosophy of Science Archive. http://philsci-archive.pitt.edu/10613/ (2014)
26. de Ronde, C.: Modality, Potentiality and Contradiction in Quantum Mechanics, WCP5. Springer, Berlin (2014)
27. de Ronde, C., Freytes, H., Domenech, G.: Interpreting the modal Kochen-Specker theorem: Possibility and many worlds in quantum mechanics. Stud. Hist. Philos. Mod. Phys. **45**, 11–18 (2014)
28. de Ronde, C., Freytes, H., Domenech, G.: Quantum mechanics and the interpretation of the orthomodular square of opposition. In: Béziau, J.-Y., Gan-Krzywoszynska, K. (eds.) New Dimensions of the Square of Opposition, pp. 223–242. Philosophia Verlag, Munich (2014)
29. de Ronde, C., Massri, C.: Revisiting the Orthodox Interpretation of 'Physical States' in Quantum Mechanics, preprint (2014)
30. Dieks, D.: Quantum Mechanics and Realism Conceptus XXII **57**, 31–47 (1988)
31. Dieks, D.: Quantum mechanics, chance and modality. Philosophica **83**, 117–137 (2010)
32. Dirac, P.A.M.: The Principles of Quantum Mechanics, 4th edn. Oxford University Press, London (1974)
33. Einstein, A., Podolsky, B., Rosen, N.: Can quantum-mechanical description be considered complete? Phys. Rev. **47**, 777–780 (1935)
34. Folse, H.J.: Niels Bohr's concept of reality. In: Lathi, P., Mittelslaedt, P. (eds.) Symposium on the foundations of Modern Physics 1987, pp. 161–179. World Scientific, Singapore (1987)
35. Freytes, H., de Ronde, C., Domenech, G.: The square of opposition in orthodmodular logic. In: Béziau, J.-Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition: Studies in Universal Logic, pp. 193–201. Springer, Basel (2012)
36. Friedman, M.: A Parting of the Ways: Carnap, Cassirer, and Heidegger. Open Court, Chicago (2000)
37. Fuchs, C., Peres, A.: Quantum theory needs no 'interpretation'. Phys. Today **53**, 70 (2000)
38. Heisenberg, W.: Physics and Philosophy. World Perspectives. George Allen and Unwin, London (1958)
39. Hilgevoord, J., Uffink, J.: The uncertainty principle. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Winter 2001 Edition). http://plato.stanford.edu/archives/win2001/entries/qt-uncertainty/ (2001)
40. Howard, D.: Was Einstein Really a Realist? Perspect. Sci. **1**, 204–251 (1993)
41. Howard, D.: Einstein, Kant, and the origins of logical empiricism. In: Salmon, W., Wolters, G. (eds.) Logic, Language, and the Structure of Scientific Theories: Proceedings of the Carnap-Reichenbach Centennial, University of Konstanz, 21–24 May 1991, pp. 45–105. University of Pittsburgh Press, Pittsburgh (1994)
42. Kauark-Leite, P.: The transcendental approach and the problem of language and reality in quantum mechanics. PhD thesis, Centre de Recherche en Epistémologie Appliquée – École Polytechnique (2004)

43. Kochen, S., Specker, E.: On the problem of hidden variables in quantum mechanics. J. Math. Mech. **17**, 59–87 (1967). Reprinted in Hooker 293–328 (1975)
44. Lahti, P.: Uncertainty and complementarity in axiomatic quantum mechanics. Int. J. Theor. Phys. **19**, 789–842 (1980)
45. Meillassoux, Q.: Après la finitude. Essai sur la nécessité de la contingence. Éditions du Seuil (2006)
46. Niniluoto, I.: Varities of Realism. In: Lathi, P., Mittelslaedt, P. (eds.) Symposium on the Foundations of Modern Physics 1987, pp. 459–483. World Scientific, Singapore (1987)
47. Pauli, W.: Writings on physics and philosophy. In: Enz, C., von Meyenn, K. (eds.). Springer, Berlin (1994)
48. Petersen, A.: The philosophy of Niels Bohr. Bull. At. Sci. Sep 1963, 8–14 (1963)
49. Plotnyski, A.: Complementarity: Anti-Epistemology After Bohr and Derrida. Duke University Press, Durham (1994)
50. Priest, G.: In Contradiction. Nijhoff, Dordrecht (1987)
51. Rédei, M.: Von Neumann's concept of quantum logic and quantum probability. In: Rédei, M., Stötzner, M. (eds.) John von Neumann and the Foundations of Quantum Physics, pp. 153–172. Kluwer Academic, Dordrecht (2001)
52. Rédei, M., Summers, S.J.: Quantum probability theory. Stud. Hist. Philos. Mod. Phys. **38**, 390–417 (2007)
53. Scavino, D.: La filosofía actual. Paidos, Buenos Aires (2001)
54. Schrödinger, E.: The present situation in quantum mechanics. Naturwissenschaften **23**, 807 (1935). Translated to english in: Wheeler, J.A., Zurek, W.H. (eds.) Quantum Theory and Measurement. Princeton University Press, Princeton (1983)
55. Smets, S.: The modes of physical properties in the logical foundations of physics. Log. Log. Philos. **14**, 37–53 (2005)
56. Vaidman, L.: Many-worlds interpretation of quantum mechanics. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Fall 2008 Edition). (2008). http://plato.stanford.edu/archives/fall2008/entries/qm-manyworlds/
57. Van Fraassen, B.C.: Quantum Mechanics: An Empiricist View. Clarendon, Oxford (1991)
58. von Weizsäcker, C.F.: Komplementarität und Naturwissenschaft. Naturwissenschaften **42**, 19–20 (1955). Translated as Complementariedad y Lógica in: La Imagen Física del Mundo. Biblioteca de Autores Cristianos, Madrid (1974)
59. Wheeler, J.A., Zurek, W.H. (eds.): Theory and Measurement. Princeton University Press, Princeton (1983)

C. de Ronde (✉)
Philosophy Institute "Dr. A. Korn", University of Buenos Aires, CONICET, Buenos Aires, Argentina
Center Leo Apostel and Foundations of the Exact Sciences, Brussels Free University, Brussels, Belgium
e-mail: cderonde@gmail.com