Studies in Universal Logic

Amirouche Moktefi
Sun-Joo Shin
Editors

Visual Reasoning with Diagrams

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.

Amirouche Moktefi • Sun-Joo Shin
Editors

Visual Reasoning with Diagrams

Birkhäuser

Editors
Amirouche Moktefi
IRIST
Université de Strasbourg
Strasbourg, France

Sun-Joo Shin
Department of Philosophy
Yale University
New Haven, CT, USA

ISBN 978-3-0348-0599-5 ISBN 978-3-0348-0600-8 (eBook)
DOI 10.1007/978-3-0348-0600-8
Springer Basel Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013944050

Mathematics Subject Classification (2010): 03-XX, 03Bxx, 03Axx

Printed on acid-free paper

Springer Basel is part of Springer Science+Business Media (www.birkhauser-science.com)

Preface

1. Valid Reasoning and Formalization

Deductive logic, all of us would agree, is the study of valid reasoning. Valid reasoning is the process of extracting certain information from given information. Logical systems are invented to make this process almost mechanical so that we may adopt them to save time and effort as well as to carry out valid reasoning in an accurate way. Realizing that mechanical processes go hand in hand with formalization, we should not be surprised to encounter much formalism in the literature on logic. At the same time, one could be intrigued by the following observation: The formalism studied in the logic literature is quite homogeneous—limited to almost only symbolic formal systems. *Why does symbolization almost exclusively dominate the enterprise of formalism in logic?* This is our opening question. We do not even pretend to answer the question decisively, but only aspire to reveal some important aspects about valid reasoning in terms of different forms of representation and to highlight some of the main motivations behind the project of the current volume.

One might be puzzled by the opening question itself, if one assumes that formalization is identical with symbolization. According to this outlook, symbolic systems are the only kind of medium to formalize our valid reasoning processes, and hence, there is no mystery about the homogeneity of logical systems. Let us examine this view by breaking it into two steps: To explore (i) the relation between valid reasoning and symbols/diagrams and (ii) the relation between formalization and symbols/diagrams. After outlining theoretical issues involved in each step, we will illustrate our positions about (i) and (ii) in the second and the third sections, respectively.

First, is valid reasoning (which is the primary mission of logic) itself tied up with a certain type of medium, that is, symbolic representation? Hence, is it the case that information-extraction is carried out only through symbol-manipulation? If so, it would not be surprising that formalization, which aims to mechanize valid reasoning process, should be limited to symbolic systems only. However, if valid reasoning itself does not dictate a certain form of medium, there would be no prima facie reason to equate formalization and symbolization. For now, we would simply like to point to examples in our daily life reasoning process: Maps, pictures, charts, and diagrams, as well as sentences and symbols are all used to carry out reasoning. Of course, pictures may be misleading. But, so may sentences. In the next section, we present extensive and historical examples to illustrate visual reasoning, that is, how valid reasoning is carried out by the use of diagrams. Thus, we would like to conclude that symbols are one of the main methods adopted in carrying out valid reasoning, but are far from being the only method. Then, why have symbols been the almost exclusive medium of formal systems?

What is the essence of formal systems? Not being a case-by-case approach, formalization aims to mechanize processes so that errors may be eliminated and at the same time effectiveness may be achieved. Indeed, accuracy and efficiency are the goals of a formal system. A system that does not assure accuracy would be useless. On the other hand, if we cared only about accuracy, that is we could take as much time as we want and we could rely on processes as elaborate as we want, then there would be no point of having a system, either. We explore further how these two desiderata are obtained so that we may get to the essence of formalization. Then, our opening question about the dominance of symbolic formal systems might find some answer.

Let us start with the accuracy desideratum. An error, we say, takes place in the case of deductive reasoning when we (wrongly) infer a false piece of information from given true information. In order to secure accuracy, we want to have a mechanism to prevent a move from true to false information. An obvious obstacle to this enterprise is that there is an infinite number of cases to get to falsity from truth. How do we come up with a way to predict and prevent these non-denumerably infinite cases? This is a dilemma almost every system has to face, and at the same time it is precisely the reason why we desire to have a system, instead of case-by-case approaches. Another perplexing element is how to deal with semantic values, that is, truth and falsity, in a system. To sum up: How can we manage an infinite number of semantic relation cases at a mechanical level?

A clever solution to this challenge is to stipulate a finite number of permissible syntactic manipulations. Only those permissible inferences are allowed in a system, and since permissible steps are finite, a system can block an error by checking each move against a finite set of rules. It should be noted that these are syntactic transformations, not semantic ones. How do we know that a finite number of syntactic manipulations guarantee the accuracy that a formal system strives for? First, we would like to determine how a finite number of rules conquer an infinite number of valid/non-valid reasoning cases. Second, we need to theorize two-way traffic between syntax and semantics. Valid reasoning is understood in terms of semantic concepts, that is, true or false. At the same time, permissible inference is defined in terms of permissible syntactic alteration.

These questions demand a meta-level of justification: The soundness and completeness proofs of a system (semi-)resolve the tension between infinity and finiteness and at the same time uphold a legitimate exchange between syntax and semantics. If a system is sound, any inference obtained by its rules is valid, and if a system is complete, any valid inference is obtainable in terms of syntactic inference. In spite of a conceptual discrepancy, as far as the extension goes, syntactic and semantic inferences coincide. Here is a triumph of formal systems: A finite number of mechanical syntactic manipulations conquered the infinite cases of semantic territory.

Let us stop here to relate the above discussions about accuracy to our question—the relation between formalization and symbols. In order to achieve the accuracy desideratum all we need is to prove that each permissible syntactic inference is a semantically valid step. There is no constraint on a medium of a formal system, as long as we can establish syntax and semantics in a non-ambiguous way. Are symbols the only kind of medium for which we can set up syntax? Many might think so, since we are very much used to thinking about syntax only in terms of sentences. On the other hand, since we have been using many different forms of representation, e.g. sentences, pictures, and sounds, to carry out reasoning, obviously semantics is not limited to particular kinds of media. However, some might argue that while ordinary piecemeal reasoning can be carried out in multi-representational forms, only symbolic systems allow systematic formal semantics.

We claim syntax and semantics are not bound to a certain form of representation at all. Pictures, just like sentences, can have their own syntax, and as long as they represent something, semantics can be defined as well. Syntax should tell us what are the vocabulary, the well-formed units, and the transformation rules. It is often the case that we judge whether a given representation is symbolic or diagrammatic depending on what its vocabulary is. In the case of Euler representation, we see circles as its vocabulary, and we therefore call it diagrammatic. Likewise, almost every logical system in logic textbooks adopts symbols, e.g., A, B, x, y, z, ∀, ∃, as vocabulary, and such systems are therefore considered symbolic systems. Well-formedness of a unit is defined so that certain kinds of arrangement, either spatial or linear, are acceptable in a system. Transformation rules can be stipulated as to which manipulations between two well-formed units are allowed in a system. Semantics, again, does not have to be tied to one form or another to carry out the job. Hence, there is nothing intrinsic about symbols or diagrams in terms of the feasibility of setting up syntax and semantics. Now, our opening question—the exclusiveness of symbolic formalization—has become even more mysterious. The third section revisits this mystery and at the same time presents a relatively recent movement for visual formalization. But first, more historical visual reasoning considerations are in order.

2. Valid Reasoning and Diagrams

A look at historical literature shows that diagrams seem to have always been used by logicians. However, uses vary. Of course, visual devices have long and often been used in educational contexts as heuristic and mnemonic tools. For instance, logic students have for centuries been familiar with squares of opposition and logic trees. These and similar structures offer in a glance a survey of the relations between propositions (Fig. 1) or illustrate the working of a logic process such as dichotomy division (Fig. 2). As such, they do usually accompany logical arguments developed in words or with the appeal to symbolic notation. However, the widespread use of such devices should not make us forget that other schemes have also been designed to carry out logical reasoning independently. Such diagrams, known to John Venn as analytical diagrams, were particularly appreciated in the 18th and 19th centuries, a period which could fairly be considered as the golden age of logic diagrams. Interestingly, this period also witnessed a growing interest in symbolic notation.

Both diagrammatic and symbolic methods were widely used during that period to solve logical problems. Broadly speaking, a syllogistic problem was understood as checking the validity of an inference where premises and conclusions were given. In the mid-nineteenth century, symbolic logicians working under the influence of mathematical practices, tended rather to offer a set of premises and look for what conclusion(s) is/are to be drawn. In both situations, logicians invented tools and methods to solve those problems with the appeal of symbolic, diagrammatic or even sometimes mechanical devices. Leonhard Euler used diagrams alone, while George Boole made use of symbols merely. Charles S. Peirce devised

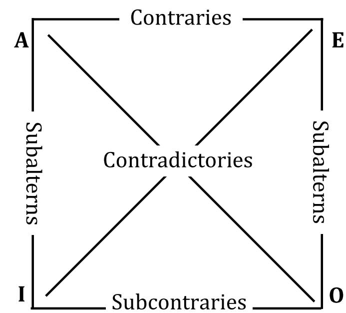


Fig. 1 Square of opposition

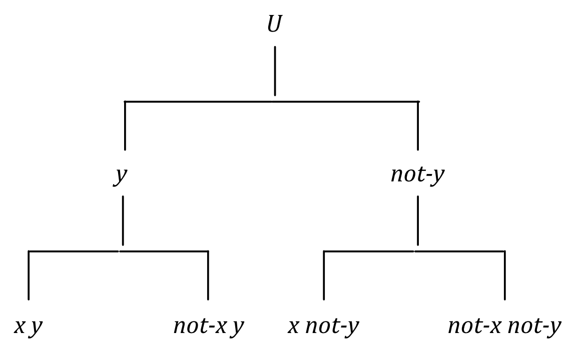


Fig. 2 Logic tree

both symbolic and diagrammatic methods. For the same purpose, Lewis Carroll designed a board and colored counters that were sold with his books. Some other logicians like William Stanley Jevons and Allan Marquand invented logic machines too. One appeals to either method depending on the problem being faced and what sounds convenient. Sometimes, the combination of several methods is also highly appreciated when dealing with complex problems. Finally, an appeal to more than one method, say both diagrammatic and symbolic, has proven useful in ascertaining individual results, each method being carried out independently in order to compare those results with each other.

This account should not be understood as the story of a happy and continuous development. Not only were different methods also in rivalry, but within diagrammatic reasoning itself various schemes were in competition. Of course, diagrams needed to be accurate in order to enter the contest. However, other criteria were in order as to the efficiency of the diagrams, the type of information they could represent, their naturalness and the visual aid they would provide. If we consider the method of representation, one can broadly distinguish two main types of diagrams. The first method, usually attributed to Euler though it was known prior to him, aims at representing given information in strictness. For instance, if we were asked to represent the proposition "All x are y", we need only to draw a circle x within a circle y (Fig. 3). This way, we do represent what is actually known: class x is included in class y. However, it might be observed that our diagram represents x as being strictly included in y, while our proposition holds also for the case where x and y are identical. In order to represent this potential information, we would do well to appeal to another method of representation attributed to Venn. The idea is to first represent all possible relations between terms, then to mark the cells to indicate their state. In the above proposition, two terms were involved (x and y), which means that there are four

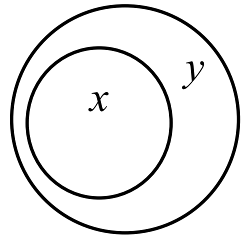


Fig. 3 Euler diagram

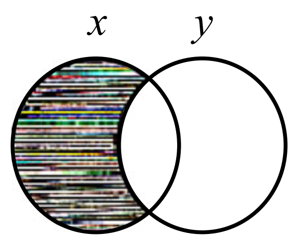


Fig. 4 Venn diagram

possible combinations: xy, x *not*-y, *not*-x y, *not*-x *not*-y, as shown in Fig. 2. These can be represented with two intersecting circles that divide the space in the desired manner. Then, the representation of proposition "All x are y" is obtained simply by shading the compartment x *not*-y to indicate its emptiness (Fig. 4).

Both diagrams do represent the same proposition "All x are y". Figure 3 looks more intuitive, but Fig. 4 is more accurate as it leaves room for the possibility of having classes x and y be identical. So far we have only represented information with these diagrams. Solving logic problems requires manipulating information in order to extract a conclusion from a set of premises. We will not give a detailed account here as such examples will be found in the chapters of this volume. However, we will explain the general idea for working such problems. First, we have to represent diagrammatically the propositions given as premises. Then, we remove the figures (here circles) that correspond to the terms that we do not want to have in the conclusion. In the case of a syllogism, that would be the middle term. Consequently, we get a diagram that represents new information as to the relation between the saved terms. Now, all we have to do is to express that relation in concrete form in order to tell what the conclusion of the argument is.

In the above example, we appealed to circles in order to represent classes. It must be noted that other shapes could have been used as well. Actually, linear diagrams have also been used by logicians and it can be easily shown that their method of representation might also be recognized as being Euler or Venn types. The same can be said about tabular diagrams that were praised in the late decades of the 19th century, thanks to their representation of a closed universe and to their advantageous use when the number of terms increases. Of course, when it comes to solving logic problems involving a high number of terms, say above six terms, diagrams become complex and lose substantially the visual aid one would expect from such devices. However, this is a practical inconvenience that should not dispute the theoretical feasibility of solving such problems diagrammatically. It must also be said here that the problem arises similarly with symbolic methods.

The variety and rivalry of diagrammatic methods that were used throughout the 19th century should not be considered as a symptom of their ambiguity. Actually, the very same situation can be observed with symbolic notations, many being aimed by their inventors to supersede rival symbolisms. Early notations used by Boole, and his followers Venn and Jevons, were equational. Later, several inclusional systems were developed by Peirce, Hugh MacColl and Ernst Schröder. Other logicians such as Carroll and Oscar H. Mitchell rather favored notations with subscripts. It is interesting to note that even for symbolic notation, visual properties certainly played a crucial role in their conception, acceptation or rejection. For instance, Peirce and Schröder's symbols for inclusion ("$-\!\!\!<$" and "$=\!\!\!<$" respectively) look like a combination of mathematical symbols "$=$" and "$<$". As such, they suggest that a class is strictly included or is identical to another class the same way a number is strictly inferior or is equal to another number. In addition, those symbols were asymmetrical the same way modern notations for implication are. As such, they were more convenient to represent inclusion than MacColl's symbol (":") which was symmetrical. Christine Ladd-Franklin adopted Peirce's symbol for inclusion, but unlike him, she preferred to invent new symbols that would be symmetrical for intersection ("$\vee$") and exclusion ("$\overline{\vee}$").

So far we have argued that both diagrammatic and symbolic methods were known to logicians and were widely used, together or separately, to solve logic problems. We also pointed out similarities in their conception, development, use and status in the 19th century. Consequently, one legitimately wonders why modern logic turned out to be almost exclusively symbolic in the 20th century. In the next section, we will briefly provide some explanations connected to the logical path explored by Gottlob Frege, and subsequently investigated by Giuseppe Peano and Bertrand Russell. However we would like to end the present section by recalling that Frege himself appealed to a system of graphs in a way that does not differ that much from Peirce's use of existential graphs. These two graph systems show that diagrammatic representations can be used effectively for advanced logic systems, beyond the class logics that we discussed throughout this section.

3. Logical Systems and Diagrams

At the end of the first section, we concluded that a formal system itself does not require any specific kind of representation as long as we can set up its syntax and semantics. In principle, symbols and diagrams have the same status, it seems. Do we, then, actually have visual formal systems in a strict sense? The answer is "yes." Peirce, who was a founder of modern logic, invented both symbolic and graphic logical systems. John Sowa's semantic network is a prime example of how Peirce's Existential Graphs could be adopted as a formal knowledge representation system. Late in the 20th century, Barwise and Etchemendy initiated the project of diagrammatic formalization, and much work has been produced under their leadership. Shin presented a modified version of Venn diagrams as a formal system equipped with its own syntax and semantics. Just as with symbolic systems, the soundness and completeness of a diagrammatic system were proved. Subsequently more diagrams have been formalized and utilized in computer science. Hence, our theoretical discussions of formalization in the first section have materialized.

This is, however, far from being a satisfactory response to our opening question—*Why does symbolization almost exclusively dominate the enterprise of formalism in logic?* On the contrary, one could be more puzzled. We have shown a persistent practice of visual reasoning throughout history in the second section, have argued for a theoretical ground for non-symbolic formal systems in the first section, and have presented diagrammatic logical systems in the above. Then, how could we explain the dominance of symbolic systems over visual systems throughout the 20th century, since modern formal logic was born? Many have pointed out that there has been a prejudice against diagrams. However, the existence of a bias does not explain the dominance of symbolic systems but reiterates it. It would be almost like answering "It is just because I like them" to the question "Why do you like apples?"

One classic complaint against diagrams is that they are misleading. We can easily find examples of misleading diagrams in geometric proofs. For example, in the case of proving a property of a triangle in general the user happens to draw an isosceles triangle and mistakenly uses in the proof the property that the triangle has equal sides. Hence, many adopt diagrams only as a heuristic tool, but not as part of a formal proof. This line of thought, however, misses entirely the main spirit of formalization. As seen in the first section, a formal system is needed precisely to prevent ambiguity and misleading of our reasoning steps, by stipulating permissible inference rules. In a sound formal diagrammatic system there would be no room for misusing diagrams and, hence, diagrams would not be able to mislead us. Then, why have we seen much more symbolic formalization than diagrammatic formalization? Is there any obstacle to formalizing diagrams? We have paved the way to clearing up any theoretical obstacles to formalization of any medium: Formal syntax and semantics do not have to belong to symbols only. Then, is it a practical choice of symbolic formalization over diagrammatic formalization? Acknowledging that this is an extremely important inquiry, we would like to encourage researchers to work on specific diagrammatic systems so that we may find more fine-grained differences among various kinds of representation systems. Fortunately, much work has been under way, and the papers in the current volume are excellent illustrations of the work in this direction. We applaud each author's valuable and creative project presented here in the volume.

Witnessing the surge of interest in visual formalization, one cannot help raising the following question: *What is the source of the new movement for visual formalization?* We suspect that this question is the flip side of our opening puzzle about the long-standing preferred practice of symbolic systems. The dominance of symbols and the new revival of diagrams are directly related to the two main goals of a formal system mentioned in the first section—accuracy and efficiency. In the world of mathematics and logic, the turn of the 20th century brought many surprises, some of which were alarming and quite devastating. Discovering inconsistencies, contradictions and limits of mathematical systems, scholars made accuracy the top priority of formalization and did everything possible to avoid any error. In this context, diagrams with their reputation of leading inference astray could not attract any serious attention, and diagrams were not considered to be something we could formalize. That is, hyper attention to accuracy along with a slightly misunderstood concept of formalization produced a wrong equation between formalization and symbolization.

If we can explain why accuracy became almost the only desideratum of our formal disciplines when facing Russell's paradox, the inconsistency of set theory, different geometries, incompleteness theorems, etc., we can also explain why the other goal of a formal system, that is, efficiency, has recently attracted new attention in our efforts at formalization, in terms of what has been recently pursued in formal systems. Not surprisingly, the age of the computer can easily tell us why efficiency has been highlighted in the research on formalization. Given more than one logically equivalent system, it is natural for us to compare their efficiencies and to choose the most efficient system. Unlike with accuracy, it is not easy to define what efficiency is and how it might vary depending on context. We would like to leave this exciting task for further work and discussions.

4. Overview of the Volume

This volume contains 10 essays on visual reasoning with diagrams. These original essays come from three main sources. First, some essays were presented at the "Logic diagrams" workshop that took place in Lisbon (Portugal), during the third Universal Logic Congress, from 22 to 25 April 2010. Second, it happened that the event coincided with the famous Eyjafjallajökull eruptions in Iceland, which substantially disturbed transportation systems in Western Europe. This prevented other contributors from attending the workshop. Happily, some of those contributions are now included in this volume. Finally, further essays have been collected thanks to a call for papers that was spread in summer 2010. Papers from these three categories have been reviewed, selected and revised between 2011 and 2012, before being offered here to the reader.

The opening essay by Catherine Legg discusses the complex question of defining what a logic diagram is. This is a classical problem that all diagram users and scholars have faced, whether in the domain of logic or not. It might look intuitive in most cases to agree on whether a given "inscription" on a page is a diagram or not. However, consensus is not always reached. And even when agreement is obtained, it is not necessarily clear what makes us consider such a representation to be diagrammatic or not. One thing is for sure: the sign itself doesn't suffice to tell what its nature is. If you consider for instance the sign "=" on a page, it could legitimately be seen as a diagram with two parallel segments. However, if you add two letters on both sides to get: $x = x$, then the sign "=" is likely to be understood as a symbol for equality. In this essay, Legg develops an expressivist view of logic diagrams and draws on Peirce's concept on iconicity. Diagrams, as signs, are thus inspected in their relationship to the objects that they represent, and in the use we make of them, not just as "pictures on the page".

The second and third essays in this volume introduce new diagrammatic systems for syllogistic calculus. Traditionally, solving syllogisms has been a necessary (but not sufficient) test for new schemes that were offered for consideration. Indeed, syllogisms are often seen as the simplest pieces of formal reasoning. As such, it is expected from any diagrammatic system to be able to solve such problems, before making further claims. New diagrammatic systems regularly appear in logical literature. Two such systems are presented here and are of special interest. The first system, known as Sophie diagrams, is developed by Richard Bosley who argues that a demonstrative syllogism depends upon its figure. As such, his diagrams represent syllogistic figures rather than specific syllogisms proper. Consequently, handling diagrammatically the second and third figures of syllogisms prevents their conversion into the first figure, as is usually done. The second system, devised by Ruggero Pagnan, seems to combine symbolic and diagrammatic elements, and is said by its author to incorporate both "a graphical appearance and an algebraic nature". Pagnan uses here linear diagrams to support a systematic treatment of syllogistic calculus and extends them to handle n-term syllogisms and to syllogisms with completed terms.

The fourth and fifth essays explore similar paths as they discuss the use of logic diagrams for problems that goes beyond traditional syllogistic. In the fourth essay, Amirouche Moktefi discusses a diagrammatic solution to a 4-term problem provided by

Lewis Carroll. This example illustrates the difficulties that are raised from working problems more complex than syllogisms, and also throws some light on what was considered to be a logic problem for early symbolic logicians. Finally, the fifth essay by Ferdinando Cavaliere introduces a new diagrammatic scheme of his invention, known as the numerical segment. As it might be guessed from its name, the purpose is to take into account quantitative considerations as one finds in non-classical logics.

The next essays, sixth and seventh, provide two instances of a formal diagrammatic system, in the style that has been pursued for about two decades. Indeed, diagram studies witnessed a revival in recent years, as has been explained in the previous section. Several diagrammatic systems have been elaborated, with syntax, semantic and manipulation rules formally defined, and have been proved to be sound and complete. The sixth essay by Jørgen Fischer Nilsson provides an example of a diagrammatic system which carefully pays attention to computing needs, notably for computer assisted reasoning. Nilsson develops a diagrammatic visualization and reasoning language, considering various logical relationships between classes. For this purpose, he uses diagrams that are transformations of Euler diagrams, augmented with higraphs. The diagrammatic system presented in the next essay, the seventh, also makes use of extended diagrams, known as spider diagrams. This essay, co-authored by Gem Stapleton, John Howse, Simon Thompson, John Taylor and Peter Chapman, provides an illustration of the type of work carried out within the Visual Modelling Group (University of Brighton, UK). In this essay, the previously published spider diagrams are augmented with constants to mark individuals. The new system is then proved to be sound, complete and decidable.

The last three essays in this volume might be connected to what has been recently known as the philosophy of mathematical practice. This new trend aims at paying more attention to the real practices of mathematicians and logicians, rather than standing with the classical problems as to the foundations of mathematics and logic alone. In the eighth essay, Valeria Giardino argues that ambiguity is an inherent feature of diagrams. This ambiguity should not be tamed because it makes diagrammatic reasoning productive by opening the way to interpretation and imagination. As such, it is the manipulation practices shared by the community that fix the meaning of the diagrams on each occasion.

The next two essays continue to consider the role of diagrams in mathematics from the viewpoint of practices, providing concrete examples from different mathematical disciplines. In the ninth paper, Zach Weber relies on category theory to look for a mathematical answer to the philosophical question that he is investigating. There, he uses the idea of natural transformation to describe what is to be considered as a good mathematical representation, be it a formula or a figure. Finally, in the last paper of this volume, Mitsuko Wate-Mizuno discusses the development of the diagrams used in graph theory. She examines Dénes König's representations in *Theorie der endlichen und unendlichen Graphen* (1936) and compares them with his predecessors to provide an account of how those diagrams evolved and took shape.

We hope that the essays in this volume will convince the reader of the richness and energy of current research in diagram studies. Contributions to this volume are made by scholars from various disciplines: philosophy, mathematics, logic, history, etc. As such, they demonstrate the need and interest in fostering interdisciplinary work. This could be carried out only by further exchanges and collaborations between scholars from different disciplines and countries, as is now the case in several international meetings such as the *Diagrams* and the *Universal Logic* congresses. We offer this volume as one further step in that direction.

Acknowledgements We would like here to express our thanks to all those who helped us during the preparation of this volume, notably the authors who contributed to this volume and the referees who accepted to review the papers. We would like to thank particularly Jean-Yves Béziau for his support both for the organization of the "Logic diagram" workshop in 2010 and for the publication of this volume in the series of *Studies in Universal Logic* which he is editing. Finally, we would like to thank the Birkhäuser team, notably Barbara Hellriegel and Sonja Gasser, for their help and patience during the preparation of this volume.

Strasbourg, France — Amirouche Moktefi
New Haven, CT, USA — Sun-Joo Shin

Contents

What is a Logical Diagram? 1
Catherine Legg

The Geometry of Diagrams and the Logic of Syllogisms 19
Richard Bosley

A Diagrammatic Calculus of Syllogisms 33
Ruggero Pagnan

Beyond Syllogisms: Carroll's (Marked) Quadriliteral Diagram 55
Amirouche Moktefi

A Diagrammatic Bridge Between Standard and Non-standard Logics: The Numerical Segment 73
Ferdinando Cavaliere

Diagrammatic Reasoning with Classes and Relationships 83
Jørgen Fischer Nilsson

On the Completeness of Spider Diagrams Augmented with Constants 101
Gem Stapleton, John Howse, Simon Thompson, John Taylor, and Peter Chapman

A Practice-Based Approach to Diagrams 135
Valeria Giardino

Figures, Formulae, and Functors 153
Zach Weber

Representation of Graphs in Diagrams of Graph Theory 171
Mitsuko Wate-Mizuno

What is a Logical Diagram?

Catherine Legg

Abstract Robert Brandom's expressivism argues that not all semantic content may be made fully explicit. This view connects in interesting ways with recent movements in philosophy of mathematics and logic (e.g. Brown, Shin, Giaquinto) to take diagrams seriously—as more than a mere 'heuristic aid' to proof, but either proofs themselves, or irreducible components of such. However what exactly is a diagram in logic? Does this constitute a cleanly definable semiotic kind? The paper will argue that such a kind does exist in Charles Peirce's conception of iconic signs, but that fully understood, logical diagrams involve a structured array of normative reasoning practices, as well as just a 'picture on a page'.

Keywords Logic · Mathematics · Diagram · Proof · Icon · Existential graphs · Expressivism · Pragmatism · Peirce · Brandom · Ayer

Mathematics Subject Classification Primary 00A66; Secondary 03A05

1 Introduction: 19th Century "Picture Shock"

20th century mainstream analytic philosophy was almost entirely neglectful of diagrams in its theorizing about *semantic content*, and *proof*. It is worth understanding the historical background to this arguably contingent state of philosophical affairs.

The trend began in mathematics. In the 19th century this field was revolutionized by an arithmetization movement, and some of the key developments foregrounded ways in which our "visual expectations in mathematics"[1] might deliver the wrong answer about mathematical fact. A famous example is the claim that a function which is everywhere continuous must be differentiable, which is in fact false. Attempting to evaluate this using visual imagination, one may imagine that if a function is continuous then it contains no 'gaps' or 'breaks', and then one seems to 'see' that at some sufficiently fine-grained level it must present a smooth surface, which would have a gradient, and thus a derivative. However, to the surprise of many, Weierstrass and Bolzano proved that certain functions are infinitely finely jagged, yet still gap-free in a way that fits the formal definition of continuity.[2] Another example is whether a 1-dimensional line might fill a 2-dimensional

[1] This phrase is taken from Marcus Giaquinto [17, p. 3].

[2] This example is nicely discussed in [17, pp. 3–4], and [22, pp. 3–4].

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic, DOI 10.1007/978-3-0348-0600-8_1, © Springer Basel 2013

region. Any attempt to mentally picture something resembling an infinitely thin thread unspooling into a finite area and thereby 'filling it in' seems to show that the claim is false, but Peano proved it true.[3] Such examples prompted some of the most influential mathematicians of the 19th century to draw strong morals about the potential for error in diagrammatic reasoning. As Marcus Giaquinto writes:

> Such cases seemed to show not merely that we are prone to make mistakes when thinking visually…but also that visual understanding actually conflicts with the truths of analysis [17, pp. 4–5].

Hilbert famously wrote, "a theorem is only proved when the proof is completely independent of the diagram" [17, p. 8], drawing on an almost identical remark by Moritz Pasch in his influential *Lectures in Modern Geometry* (1882). So, remarkably, even the field of *geometry*, it came to be seen, needed to be purged of diagrams.[4] The end result was a "prevailing conception of mathematical proof" which John Mumma describes as "purely sentential", as follows:

> A proof…is a sequence of sentences. Each sentence is either an assumption of the proof, or is derived via sound inference rules from sentences preceding it. The sentence appearing at the end of the sequence is what has been proven [22, p. 1].

This suspicion of 'visual expectations' then flowed into Frege's work on the foundations of mathematics. Cognizant of the errors which his fellow mathematicians had learned to skirt, Frege attempted to entirely remove 'intuition' from the logic with which he set to put mathematics on an entirely new and more rigorous foundation. Famously, he remarked of his own concept-script:

> So that nothing intuitive could intrude here unnoticed, everything had to depend on the chain of inference being free of gaps [16, p. 48].

Frege argued against the empiricism of John Stuart Mill that numbers were not properties abstracted from the physical world, but definable purely analytically.

Frege in turn was an enormous influence on *logical positivism* (Carnap studied under him, for instance), which in turn set the scene for mainstream analytic philosophy's aims and methodologies in many ways that are still being worked out today. The movement's early strict focus on clarifying *meaning* owed much to Frege's vision of an ideal language all of whose inferential steps are explicitly stated, and use a set of rules specified in advance.[5] Thus A.J. Ayer laid down a strict definition of "literal significance" as confined to claims which have "factual content" by virtue of offering "empirical hypotheses" [1, p. 2]. Thus, to illustrate by way of a simple example, "The cat is on the mat" is literally significant because there is a cat-being-on-the-mat type of *experience* which might be had—or not—in the relevant situations.

---

[3] Discussed in [17, pp. 4–5].

[4] "A body of work emerged in the late 19th century which grounded elementary geometry in abstract axiomatic theories…This development is now universally regarded as a methodological breakthrough. Geometric relations which previously were logically free-floating, because they were understood via diagrams, were given a firm footing with precisely defined primitives and axioms" [22, p. 6]. Non-Euclidean geometries are another key example, and I am grateful to an anonymous referee for pointing this out.

[5] Although transposed into a rigidly empiricist setting which truth be told sits oddly with Frege's thinking—and arguably has caused significant problems in the philosophy of mathematics.

Claims which lack "literal significance" fall into two camps. Either they can be "literally false" but somehow "the creation of a work of Art" which is gestured towards as valuable, though Ayer is somewhat vague about how. Or, worse, claims might be "pseudo-propositions"—disguised nonsense. Any claim lacking literal significance is not the purview of philosophy [1, p. 2]. It is hard to see how a diagram could offer an empirical hypothesis, and thus have literal significance in Ayer's sense. And he briskly dismisses the idea that a philosopher might be "endowed with a faculty of intellectual intuition which enabled him to know facts that could not be known through sense-experience" [1, p. 1]. Likewise, the early Carnap [10] claimed that statements were meaningful if syntactically well-formed and their non-logical terms reducible to observational terms in the natural sciences.

It is well-known that crisp criteria for what constitutes a genuine empirical hypothesis were much more difficult to find than Ayer imagined they would be. Carnap dropped back from demanding verifiability to requiring "partial testability" [11], and confirmation became a more and more holistic affair, until finally Quine acknowledged that what meets the tribunal of experience is in an important sense the whole of science. By way of consolation for thus sounding verificationism's death-knell, Quine offered a new criterion of what might be called 'factuality': if we could imagine our science collated and regularized into a single theory expressed in first-order logic, its bound variables would have values. In a pseudo-science such as witchcraft they would not [33]. Now we can say that "The cat is on the mat" is factual because in the logical formula $\exists \mathtt{x}\ (\mathtt{Cx}\ \&\ \mathtt{Oxm})$ suitably interpreted, the variable $\mathtt{x}$ binds to George.[6]

Philosophers' banishment of diagrams from semantics and theories of inference arguably reached a high-water mark in the 1970s with the publication of Quine's colleague Nelson Goodman's *Languages of Art*. Here Goodman made an influential argument that resemblance plays no interesting or important role in signification. Rather, he claimed that denotation, "is the core of representation and is independent of resemblance" [18, p. 5]. His reasoning was that while the resemblance relation is symmetric (if $\mathtt{X}$ resembles $\mathtt{Y}$ then $\mathtt{Y}$ resembles $\mathtt{X}$), the representation relation is not.[7]

However, a profound challenge to this more than century-long neglect of diagrams is 'in the air'. It seeks to reconceive diagrams as more than a mere 'heuristic aid' to proof in mathematics and logic. Rather diagrams may be understood as capable of serving either as proofs themselves, or irreducible components of such. Thus James R. Brown writes:

> ...the prevailing attitude is that pictures are really no more than heuristic devices...I want to oppose this view and to make a case for pictures having a legitimate role to play as evidence and justification—a role well beyond the heuristic. In short, pictures can prove theorems [9, p. 96].

John Mumma writes:

> In the past 15 years, a sizable literature consciously opposed to [the attitude that pictures do not prove anything in mathematics] has emerged. The work ranges from technical presentations of formal diagrammatic systems of proof...to philosophical arguments for the mathematical legitimacy of pictures... [22, p. 8].

---

6 (A cat.)

7 Randall Dipert has argued against this that it no more follows that resemblance is 'entirely independent of' representation because the former relation is symmetric and the latter is not, than that the brother relation is 'entirely independent of' the uncle relation as the former is symmetric and the latter is not [15].

Meanwhile Marcus Giaquinto writes:

> . . .a time-honoured view, still prevalent, is that the utility of visual thinking in mathematics is only psychological, not epistemological. . .The chief aim of this work is to put that view to the test [17, p. 1].

Other authors have returned to ancient Greek mathematical texts to argue that one cannot understand them fully without taking their diagrams more seriously [13].[8]

Meanwhile, in logic, Sun-Joo Shin argues that although, "[f]or more than a century, symbolic representation systems have been the exclusive subject for formal logic" [36, p. 1], this should be widened to also consider "heterogeneous systems", which "employ both symbolic and diagrammatic elements" [36, p. 1]. "Heterogeneous systems" is an influential term which derives from Jon Barwise [2]. Shin argues that symbolic and heterogeneous reasoning systems have different strengths and weaknesses, and we should do a thorough study to get the best out of both, bearing in mind that different disciplines which might draw on such systems (such as logic, artificial intelligence and philosophy of mind) might have different needs.

This paper seeks to join these authors while at the same time to put this goal in a broader context, namely a movement which is also aimed at unbuilding the simple picture of "literal significance" that has been so influential in the 20th century—*expressivism*.

2 Expressivism: Saying, Doing and Picturing

Expressivism has a metaethical incarnation, as a view that, ". . .claims some interesting disanalogy between. . .evaluations and descriptions of the world" [12, p. 1]. By contrast, Robert Brandom has put forward a semantic expressivism whose main point is that not all semantic content may be made fully *explicit*. This view contrasts with a widespread view often thought to be intuitively obvious, and arguably a downstream specter of Ayer's notion of literal significance. I will call it a *metaphysical realist semantics*. The juxtaposition here is deliberately somewhat controversial, given that many metaphysical realists take great pains to make a clear separation between metaphysical and semantic questions, and to claim that their view lies firmly on the metaphysical side. An argument will be put forward later in the paper that this self-assessment is problematic.

A metaphysical realist semantics holds that the purpose of language is to state "facts" which, if the propositions stating them are true, form part of language-independent reality. Thus, to return to our earlier example, "The cat is on the mat" (suitably disambiguated as to cats and mats) is thought to present a 'content' which it is sufficient to know the meaning of the statement's words to fully understand. Brandom calls the view *representationalism*. By contrast, he argues that the primary purpose of language is to transform what we *do* into something that we can *say*:

> By expressivism I mean the idea that discursive practice makes us special in enabling us to make *explicit*, in the form of something we can *say* or *think*, what otherwise remains *implicit* in what we *do* [30, p. 7].

[8] See also, from a more philological perspective, the work of Reviel Netz, e.g. [23].

Crucially, this renders the explicit statement semantically parasitic on the implicit practice, in that one cannot fully understand the statement without antecedently understanding the practice which it "expresses". Thus Brandom writes:

> ...we need not yield to the temptation...to think of what is expressed and the expression of it as individually intelligible independently of consideration of the relations between them...And the explicit may not be specifiable apart from consideration of what is made explicit [8, pp. 8–9].

Consider for example, the invention of *musical notation*. This freed musicians from having to learn music by directly copying a live musician's *actions*. Instead a musical score substitutes dots on a page for string-pluckings, key tappings, and all other actions which might produce a note. In this way a musical score can *say what musicians do* (with added bonuses such as that the score can be indefinitely copied, survive longer than any living musician, and be readily compared and contrasted with other scores). However, it is not possible to fully understand a musical score without having some antecedent understanding of the practices of music which it is expressing. For instance, if aliens were to stumble upon the score for Beethoven's 5th symphony, it is highly unlikely they could perform it without some observation of human musical performance.

This commitment to a parasitism of the explicit statement on the implicit practice renders expressivism a form of *pragmatism*. It claims that certain practices are not fully explicated in language, but presupposed by it [7]. Pragmatism is frequently seen as a form of antirealism, merely internal realism,[9] non-cognitivism,[10] non-factualism,[11] or as some would put it "quasi-realism".[12] But the conclusion of this paper will consider other views on this.

Such an expressivism may make sense for musical notation, but might it be generalized? For Brandom wishes it to be a global view, concerning all language. In particular, might expressivism be applied to talk about *logic*? Surely the matters of truth-preservingness and validity are a paradigm of practice-independent fact? Not so, according to Brandom. He claims that logic also should be seen as a way of *saying* what we are *doing* when we actually make inferences, in ways that can guide our reasoning in systematic and useful ways. In fact he self-consciously highlights the practice of philosophy itself as a particularly sophisticated pulling of unselfconscious implicit *practices* into explicit *statements* that might be critically appraised [8, pp. 56–57].

Brandom's expressivism may be linked in interesting ways with Wittgenstein's Picture Theory of Meaning.[13] In the *Tractatus* Wittgenstein drew a famous distinction between what is *said* (namely atomic facts, and truth-functional combinations of them) and what is *shown* (the laws of logic, the limits of the world and, interestingly in the expressivist context, ethics). In the spirit of Brandom we might describe the former as "explicit" and the latter as "implicit". However, having drawn this distinction between saying and showing, Wittgenstein made the further claim that what is shown *cannot be said*.[14] Early Wittgenstein and Brandom stand out amongst mainstream semantics in their bold claim that not

---

[9] [5, 32].

[10] Price suggests Rorty approaches a global non-cognitivism in [29].

[11] This term derives from [6].

[12] The term was coined by Simon Blackburn, see in particular [4]. Its links with pragmatism are explored in [28], though [21] argues that the two views share important similarities *and* differences.

[13] I have argued this previously elsewhere: [19, 20].

[14] Thus for instance *Tractatus* 6.42 states, "...there can be no ethical propositions..." [37].

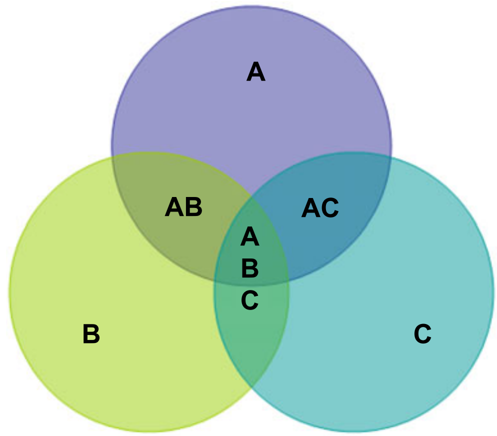


Fig. 1 A paradigmatic diagram

everything true can be made explicit, or stated. However they are also different in two ways. Firstly, where Wittgenstein suggested that "the said" and "the shown" consist in two irrevocably sundered 'camps' of content, Brandom allows any implicit practice to be made explicit—an example would be noticing a pattern in one's reasoning and naming it *Modus Ponens*. He merely notes that this can only happen against a background of further implicit practices (in the example just cited—argument categorization). Secondly, although a 'semantic parasitism' exists in both views, it is apparently traveling in opposite directions. For where we saw that for Brandom the *explicit* is parasitic on the *implicit*, for the early Wittgenstein it appears that what is *shown* is parasitic on what is *said*.

I will now develop an expressivist view of logical diagrams.

3 Defining the Diagrammatic

What exactly *is* a diagram in logic? Can we give a definition which would cover all cases which we would want to call logical diagrams, and not cover any cases which we would not? Let us begin by trying to define a diagram more generally. Figure 1 would appear to be a paradigm case—so what is 'diagrammatic' about it?

First of all, it seems capable of conveying some kind of meaning, as it is so structured. However whatever meaning it has certainly seems different to that which would be conveyed by a piece of prose. But how exactly? I will now work through a series of possible definitions of 'diagram' which attempt to capture this. Although some of these definitions might appear naïve, or held by few—the order in which they are arranged is intended to embody a useful learning process in the reader.

We might begin by attempting the following definition:

(i) "*X is a diagram iff there are no words on the page*".

However Fig. 2 fails this criterion. And yet it is arguably a diagram. One might protest that the 'part on the left' of Fig. 2 is the truly diagrammatic part, while the 'part on the right' is just a label. But one would not wish to say that the part on the right is not part of Fig. 2. It is pointing directly towards it and we have conventions for interpreting such arrows. Thus

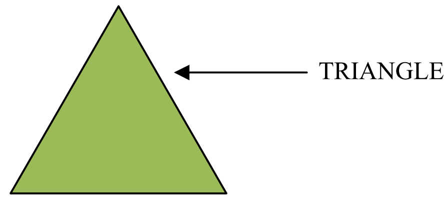


Fig. 2 A challenge to definition (i)

having a 'truly diagrammatic part' seems to be sufficient for being a diagram, at least in this case. So perhaps we can capture this in a new definition:

(ii) "*X is a diagram iff there are pictures on the page*".

But what is a picture? Can we give a definition which would cover all cases which we would want to call pictures, and not cover any cases which we would not? Perhaps we could say that a picture, unlike a piece of prose, is made of joined-up lines? This produces another possible definition:

(iii) "*X is a diagram iff there are joined-up lines on the page*".

However Fig. 3 has no joined up lines, only letters. Yet it too is arguably a diagram. Although it is composed solely of words, they are arranged in a *structure*, and this seems to render it diagrammatic. So maybe we could capture this with a definition something like:

(iii) *X is a diagram iff there is some 'non-word component' on the page*.

This however seems awfully vague, and even so it is probably not a sufficient criterion. (Punctuation? Page numbers?)

One might wonder at this point if we are attempting to define something too basic and fundamental to be put into words. Or perhaps we are attempting to define something too heterogeneous—maybe the concept of a diagram is not a cleanly demarcatable semiotic kind? However this would be to give up too easily. I will now offer a definition which draws on Charles Peirce's concept of an *iconic sign*. Our key problem so far has in fact been trying to craft our definition merely by inspecting *the sign itself*. Peirce believed the key criterion is the sign's *relationship to its object*.

4 Peirce's Icon: The Sign Which Resembles

Peirce's 'philosophy of language' falls within a much wider theory of *signs*, or *semiotics* [25, 26]. Including pictures and diagrams is part of the point of this broader disciplinary

Fig. 3 A challenge to definition (iii)

purview. Peirce defined a sign very broadly as any irreducibly triadic relation between a *representation*, an *object*, and an *interpretation*. As well as attributing triadic structure to the sign itself, he taxonomized signs using a series of three-way distinctions. The icon is part of a triad comprising *icon*, *index* and *symbol*, corresponding to the three different ways in which Peirce believed a sign could be associated with an object.

Symbols symbolize what they do via some arbitrary habit or *convention which must be learned*. Thus we must learn that the word "banana", in English, means bananas. All words are symbolic to some degree, as they belong to shared public language. However as Perry [27], and others showed, language also includes signs which pick out an object not by arbitrary habit or convention but *via* some direct *'indicating'* or *'pointing'* relationship (e.g. "here", "now"). Peirce called these signs *indices*. Finally *icons* are signs which *resemble what they signify* [24, 2.304]. This category is distinct from symbols, as resemblances need not be established by convention but can be perceived anew (e.g. "That cloud looks like a frog"). Examples include maps, paintings and but also, crucially, mathematical diagrams which function by *mimicking the structures they signify*.

This definition of the icon immediately raises skeptical concerns in the minds of many. "Resemblance is cheap", it is thought. Anything can be argued to resemble any other thing in *some* respect. For instance, a photograph of Richard Nixon might be thought to resemble other objects *qua* male (e.g. Brad Pitt), *qua* brunette (e.g. Elizabeth Taylor), *qua* oval-headed (e.g. an egg), or in many other more recondite ways (e.g. "Something in his eyes reminds me of Mt Everest..."). To top it off, the Löwenheim-Skolem theorem is often vaguely invoked at this point, following Putnam [31], to suggest that all the points on one object can be mapped onto all the points on any other object to produce an 'isomorphism', so in some suitably impressive mathematico-logical way, a triangle could be 'like' a square, a cow could be 'like' a flock of birds and at that point the whole business of likeness dissolves into arbitrariness.

How is a serious theory of language, not to mention reasoning, to lean on such an apparently subjective pillar? There seems to be an unarticulated yet profound intuition in contemporary analytic philosophy that this is why semantics should be grounded squarely on *reference*, as to make room for the contingencies of the mind's ability to creatively notice likenesses would introduce theoretical chaos.

The worry is very understandable, and properly addressing it would require excavating and settling a number of very deep issues. I will mention three. (i) One will need to argue for a *realism about structures* which is arguably underappreciated since the Quinean equation of ontological commitment with the bound variable noted above.[15] For Peirce claims that the parts of an icon bear the same relationship to one another as do the parts of the object the icon represents [24, 3.363]. This definition seems to be a good way of explicating *structure*, and it is worth highlighting that icons are the structural signs par excellence, in terms of *their means of signification*. For although indices may be words and to that degree possess internal complexity (in the individual letters), their *signifying* function is to serve as a pure pointer. The word 'here' indicates a location in space—it does not 'say' anything else. And although the convention by means of which the symbol symbolizes what it does may have structure, that structure is not *internal to the sign itself*. For instance the word 'dollar' represents what it does in NZ society by a complex set of

[15] Pace the recent *structuralist* movement in philosophy of science.

conventions involving different colored notes, numbers on a screen in Internet banking, and so on, but none of these 'convention-parts' are related in the same way as are the parts of the word 'dollar' itself (once again: its individual letters). However in a map of the North Island of New Zealand we discover that Hamilton is West of Napier by observing the relevant spatial relations *in the map itself*.

(ii) The second issue concerns properly understanding the role of the icon, which is not to generate ontological commitments in the simple denoting way that many in the Quinean tradition envisage. Denotation is the role of the index—but it can only perform this function when appropriately supported by the other two sign-types. In fact Peirce believed that icons, indices and symbols play three different functional roles whose co-presence and coordination is vital for language to function as it does. It is worth explaining these roles. Firstly, the symbol's conventional nature means that it signifies *general properties*, because conventions are "general rules" [24, 3.360] which can be applied any number of times in situations which display the appropriate (general) features. Secondly, due to their pure pointing function, indices designate *particular existence*—Peirce writes, "[a]n indexical word. . .has force to draw the attention of the listener to some hecceity common to the experience of speaker and listener" [24, 3.460]. Finally, icons designate *neither general facts nor particular existences*. Rather they signify hypotheses, possible situations [24, 3.362]. Relatedly, it is important to note that these three sign-types are not mutually exclusive in that a single sign can serve as icon, index and symbol in different respects all at the same time.[16]

(iii) As for the re-introduction of the mind's creative idiosyncrasies to semantics and the theory of inference—the issue here is to embrace it, unnerving as it may be—this is long overdue. There is a long tradition of debate in philosophy over whether the thinking mind is essentially *active* or *passive*, with rationalists generally preferring the former camp, empiricists the latter. It is a major difference in the thought of Kant and Hume. (Although it must be conceded that Hume did highlight the imagination, seemingly an active faculty, to a degree unmatched by other British empiricists, nevertheless he explicated this faculty within a naturalistic perspective with strong determinist implications, and when he discussed the problem of free will arguably offered a compatibilist "cop-out".) At any rate, achieving this re-introduction will require mounting some deep challenges concerning what a semantics or theory of inference is *for*. The fields have arguably strayed from giving an *account* of the phenomena in question, too far towards a felt need to provide an *algorithm to predict* them.[17]

Lacking time to properly pursue these three large inquiries, I will mount a preliminary phenomenological argument for the worth of the icon in the robustly mind-independent field of logic by demonstrating Peirce's iconic logic *in use* (arguably not a terrible strategy for a pragmatist to use).

[16] An example is a shadow-clock, which is iconic insofar as it represents the 24 hour structure of our day, indexical insofar as it relies on the sun physically casting a shadow to tell the time, and symbolic insofar as the numerals on the clock-face have meanings which must be learned.

[17] This might have something to do with the fact that key researchers in semantics and logic in the 1960s and 70s also worked in artificial intelligence.

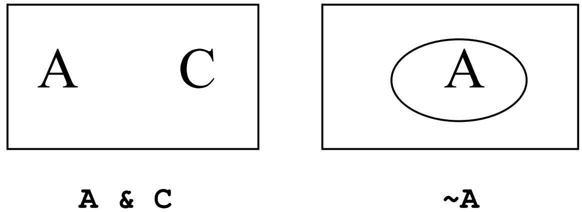


Fig. 4 Connectives in Peirce's Alpha Graphs

5 Peirce's Existential Graphs

Later in his career Peirce developed a diagrammatic logic which he called the Existential Graphs (henceforth: EG), claiming that "all necessary reasoning without exception is diagrammatic". His next remark about how these diagrams are to be used is interesting—he writes, "…we construct an icon of our hypothetical state of things and proceed to observe it" [24, 5.162]. I will demonstrate this system in its simplest, Alpha Graph form, which is provably equivalent to modern propositional logic [34]. But first, here is a traditional (natural deduction) proof for purposes of comparison:

Proof I, "Symbolic Logic": $\vdash (\mathtt{P} \vee \sim\mathtt{P})$

1	$\neg(P \vee \neg P)$	H
2	P	H
3	$P \vee \neg P$	I$\vee$ 2
4	$\neg(P \vee \neg P)$	IT 1
5	$\neg P$	I$\neg$2, 3, 4
6	$\neg P$	H
7	$P \vee \neg P$	I$\vee$ 6
8	$\neg(P \vee \neg P)$	IT 1
9	$\neg\neg P$	I$\neg$6, 7, 8
10	P	E$\neg$9
11	$\neg\neg(P \vee \neg P)$	I$\neg$1, 5, 10
12	$P \vee \neg P$	E$\neg$11

This proof *seems* to be purely sentential in Mumma's terms. Observe the length required to establish the simplest of tautologies!

By contrast, the EG uses no connectives between sentence letters. Rather it represents *conjunction* by placing two sentence letters on the same "sheet of assertion", and *negation* by drawing a line or 'cut' around the proposition ('graph') in question (see Fig. 4). A benefit that has been claimed for logical diagrams is so-called *free rides* [3]. One can immediately 'see' certain logical equivalences.[18] Thus, Fig. 5 may be observed to elegantly represent: $\sim(\mathtt{A}\ \&\ \sim\mathtt{C})$, $(\mathtt{A} \vee \sim\mathtt{C})$ and $(\mathtt{A} \supset \mathtt{C})$ simultaneously. This is an immediate perception. How about a proof? First we need some rules.

The following 5 Alpha Graph rules each consist merely in a permission to *write* or *erase* a graph on the sheet of assertion.

[18] Shin also calls this the "multiple carving principle" [35, p. 77].

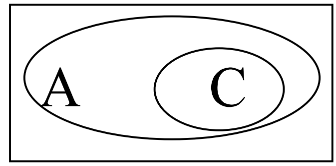
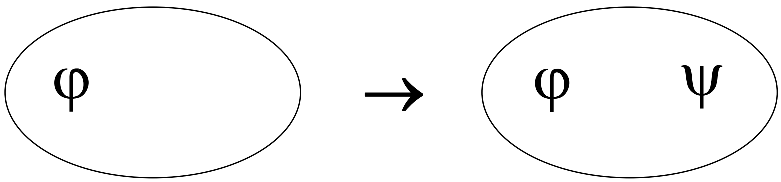


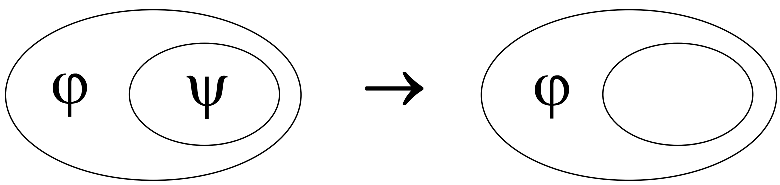
Fig. 5 Perceivable logical equivalence

Proof Rules for Alpha Graphs Diagrams are from Zeman [38, 39].

1. Insertion in odd: In an oddly enclosed area, any formula may be written.

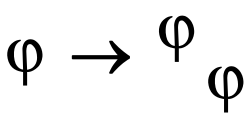


2. Erasure in even: In an evenly enclosed area, any formula may be erased.

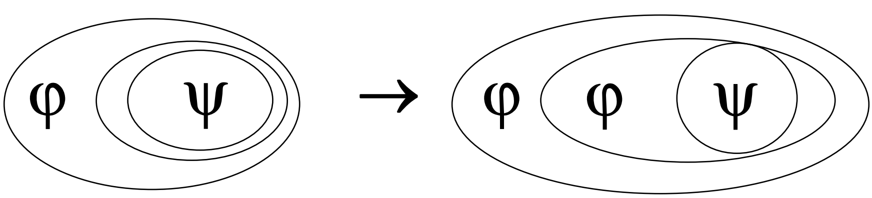


3. Iteration: Any formula may be written again:

(**a**) in the same area:

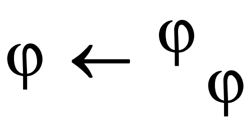


(**b**) by ‘crossing cuts’ in an inward direction:

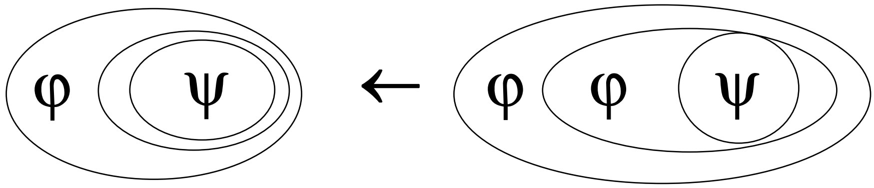


4. Deiteration: Any formula derivable through iteration may be removed:

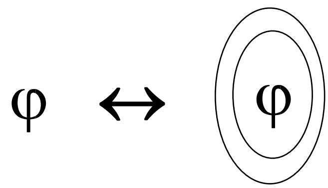
(**a**) in the same area:



(**b**) ‘crossing cuts’:

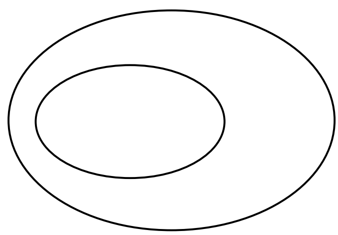


5. Biclosure: A double negation may always be added or removed.

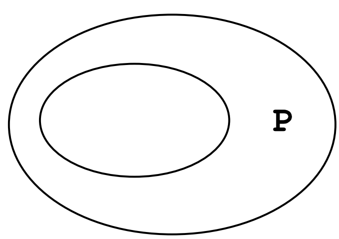


I will now perform two proofs, beginning with the simple tautology proven above:

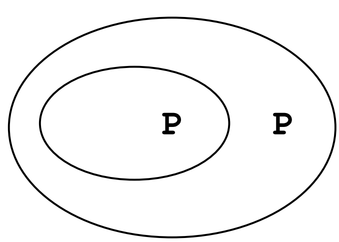
Proof II, EG: ⊢ (P ∨ ∼P) First, rule **5** allows the adding of a double negation (frequently a first step in these proofs):



Rule **1** allows any graph whatsoever to be added in an oddly enclosed area. Not surprisingly, we choose P:

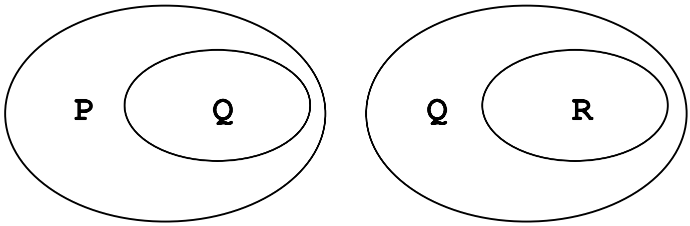


Rule (**3b**) (Iteration "Crossing Cuts") allows P to be re-scribed into the evenly enclosed space within:

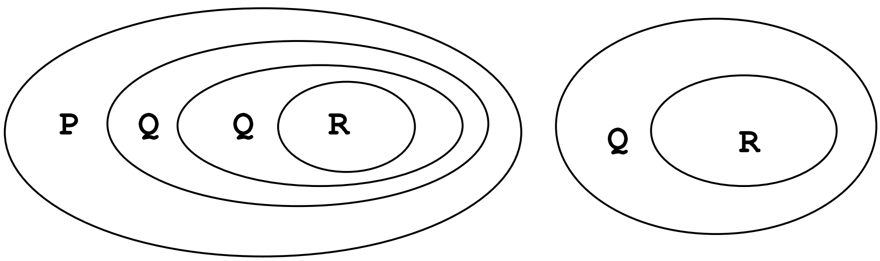


Our tautology is now most easily proven.

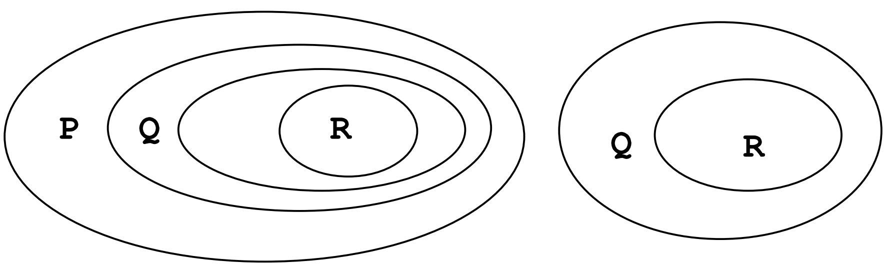
Proof III, EG: (P ⊃ Q), (Q ⊃ R) ⊢ (P ⊃ R) (*Hypothetical Syllogism*) We begin by scribing the premises on the sheet of assertion:



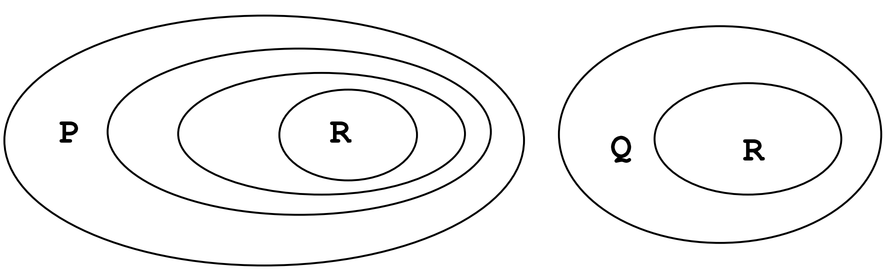
Next we iterate the right-hand portion of the diagram inside the left-hand portion, using Iteration Crossing Cuts (**3b**):



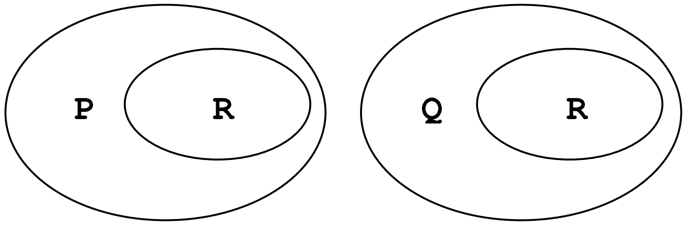
We now deiterate the innermost Q, by (**4b**):



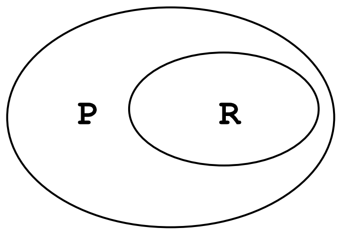
We now deiterate the other Q on the left-hand side, using Erasure in Even (**2**):



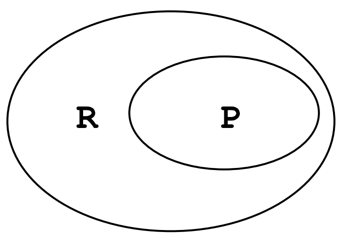
We may now remove the double negation ("biclosure") around R on the left-hand side (**5**):



And the final result is achieved by using (**2**) to erase the right-hand side (functionally equivalent to conjunction elimination):



Now think about how this procedure might work if we attempted to prove the invalid: $(\mathtt{P} \supset \mathtt{Q}), (\mathtt{Q} \supset \mathtt{R}) \vdash (\mathtt{R} \supset \mathtt{P})$. The proof is not possible as we would need to produce the following:



However our proof puts P in the outermost (odd) enclosure, and there is no way consistent with the rules that it can be removed. Graphs in odd enclosures may be iterated inwards, but never *removed*. Yet it would need to be removed in order to realize the conclusion as pictured above.

6 Logical Diagrammatic Forcing

We may now see that a "logical diagram" includes its rules of use, not just pictures on the page. It is no accident that the EG rules were presented before the diagram was constructed to furnish the proof. These rules are of course explicit representations of implicit reasoning practices,[19] to use Brandom's terminology [7]. We do not understand the rules without understanding the *actions* (literally adding and removing graphs from the diagram) which they represent. Once we understand the rules and most importantly—use

[19]This is not to say that ordinary reasoners would necessarily recognize them as such. This is theoretical not applied logic (what Peirce called *logica docens*, opposing it to *logica utens*, a distinction that medieval logicians drew).

them, we *directly experience* the impossibility of rendering a contradictory proposition or invalid argument on an EG.[20] We are *forced* to recognize how some part of what we are trying to realize 'has to give'. We thereby 'see' (in some arguably metaphorical, but powerful sense, which means something like 'structurally perceive') logical necessity. This is the "faculty of intellectual intuition" so facilely dismissed by Ayer. It is crucial to note that this intuition occurs not by having epistemic contact with any further 'necessary object' (whatever that might be), but merely by fully *grasping* the relationships amongst the diagram's different parts, already present on the page.

At the same time it is important to note that not all aspects of Peirce's logical diagrams are forced by their structure, and thus iconic. Some aspects are *symbolic*—for instance one must learn the convention that letters correspond to propositions, and not to, say, predicate letters applied to an object represented by the larger circle. Some aspects of the graphs are also *indexical*. For instance the sentence letters serve to indicate particular propositions (which are indicated in somewhat vestigial form, as propositional logic is characterized by abstracting away from atomic propositional content, yet remain as crucial place-holders). This is how Peirce's semiotics works—all three kinds of sign need to be present and to work together to create significance.

We have just seen that so-called 'iconic logic' is not purely iconic. At the same time, so-called 'symbolic logic' is not purely symbolic. Natural deduction, used in Proof I above, also has rules, forces certain results and forbids others, and is iconic *to that degree*. We therefore do not have symbolic logic and iconic logic, strictly speaking. We have logical systems whose iconicity is more or less *perspicuous*. If we bear in mind our initial definition of the icon, that its parts are related in the same way that the objects represented by those parts are themselves related, perspicuity consists in as many of the relationships on the diagram as possible representing logical, as opposed to arbitrary relationships. Peirce believed his system was more perspicuous than the 'algebraic' logic he had worked in prior to developing the EG,[21] and that it would be useful, not for proving results the other couldn't, but for studying logical form more clearly and minutely.

7 Conclusion

This paper has presented an expressivist view of logical diagrams. Thus Brown is vindicated in his claim that, "pictures hav[e] a legitimate role to play as evidence and justification". In fact we now see that *all* formal logic is essentially diagrammatic, although we have seen that the diagrams may be more or less perspicuous. We can also see that Mumma's definition of the 'purely sentential proof' he wishes to argue against ("[a] proof…is a sequence of sentences. Each sentence is either an assumption of the proof, or is derived via sound inference rules from sentences preceding it") wears incoherence on its sleeve. For once one adds 'sound inference rules' to the mix, one has *more* than just

[20] Although this is counterintuitive to some, the fact that the experience is only possible once one has grasped the proper interpretation of the EG rules does not undermine its *directness* once those rules have been grasped.

[21] It is worth noting that this logic articulated the first version of the Russell-Peano notation for quantifiers which is standard today. Thanks are due to an anonymous referee for suggesting this.

a sequence of sentences. Sound inference rules convey logical form and the actions which respect and mirror it.

One might ask: But what about the visual-expectation-derived mistakes which 19th century mathematicians learned to avoid? If we embrace diagrammatic reasoning, how do we know we won't be led into error in this field? This is a good question. The short answer is: Get better diagrams ('better' here meaning more perspicuous). The long answer, which would involve formulating a principled account of which structural features of any given diagram represent necessary truths, and which do not, will take much further work to determine.

Through diagrams such as Peirce's EG, logical necessity is presented to the human mind in such a way that it can be understood and learned. And what more could we ask in order to say that a system of signs represents something, or has genuine content? But at the same time the diagrams do not *state* logical necessity in anything like Ayer's sense of literal significance. The graphs do not put forward an empirical hypothesis.[22] They provide the means for us to exercise our rational intuition. Moreover, if we turn once again to Quine's criterion for ontological commitment, the basis for the metaphysical realist semantics pervasive today, we can see the graphs do not fit this model either. They do not gain their content by denoting further objects in the way that "The cat is on the mat" denotes a cat, and a mat. Rather, as icons, everything needed for logical insight is already internal to them as signs—one merely has to attend to their structure.

We might pause in closing and ask: What are the implications of this expressivism for *realism about logic*? For instance, does it show that the discourse of logic does not 'talk about real mind-independent things'? My answer is: No, but that our notion of 'real mind-independent things' requires some surgery.

It was noted that a great deal of recent metaphysics is semantics-driven,[23] although metaphysical realists sometimes rightly express some discomfort about this, as it would seem their very realism should lead them to try to keep semantics and metaphysics separate. I would say, actually it is fine to derive one's metaphysics from one's semantics—just please, please get a less simplistic semantics! We may understand Quine's criterion of ontological commitment in Peircean semiotic terms as an attempt to place the full burden of representing reality onto *indexical signs*. This leads philosophers with realist sympathies to feel they need to ask a raft of questions of the form: "Does term X [e.g. ethical or aesthetic predicates, number-terms…] denote a real object?" If we recall that indexical signs pick out sign-independent particulars, it often seems hard to answer "yes" to this question for key terms in manifestly important human discourses (such as ethical or aesthetic predicates, number terms…). On the other hand, those who are unsatisfied with metaphysical realism's problematization of such areas, and those who wish to recognize a manifest social input to human language-games, often oppose metaphysical realism with some form of *conventionalism* which argues that term X does not denote a real object

[22]One might object that this is incorrect, since a diagram such as a map may be understood to posit the shape of a real-world country. However *considered purely qua diagram,* a map does not yet have that semiotic function. To interpret it as saying something about a country is: (i) to peg it to a real-world object (thereby rendering it also an *index*), (ii) to claim something general about that object's shape (rendering it also a *symbol*). These are *further* signs.

[23]Thus for instance, Chrisman sums up much recent metaethics by writing, "The realism debate has been pursued (mostly) by investigating the appropriate semantic account of ethical statements" [14, p. 334].

but has some other socially sanctioned and taught *function*. We may understand such a conventionalism in Peircean terms as trying to understand all signification as performed with *symbolic signs*.

Metaphysical realism and conventionalism are widely assumed to be polar opposites. So many dialectics in so many papers in so many areas of philosophy revolve around this, so that an argument against metaphysical realism is more often than not assumed without question to be an argument for conventionalism, and an argument against conventionalism to be an argument for metaphysical realism. But this is a false dichotomy. A third kind of signification exists which does not consist in brute denotation *or* arbitrary convention, but which presents structure directly to the mind's eye. It is barely glimpsed in formal semantics today. And yet it is this kind of sign that represents *logical form*—hardly a trivial part of our conceptual scheme. If we could only recognize that the symbol, the index and the icon all have a unique and irreducible semantic role to play, and that reality correspondingly is comprised of *real habits*, *real particulars* and *real intrinsic structures*, we would take an unanticipated leap towards understanding that most contested concept.

References

1. Ayer, A.J.: Language, Truth and Logic. Dover, London (1953)
2. Barwise, J.: Heterogeneous reasoning. In: Conceptual Graphs for Knowledge Representation. Lecture Notes in Computer Science, vol. 699, pp. 64–74 (1993)
3. Barwise, J., Shimojima, A.: Surrogate reasoning. Cognit. Stud. **4**(2), 7–27 (1995)
4. Blackburn, S.: Spreading the Word. Clarendon, Oxford (1984)
5. Blackburn, S.: Wittgenstein, Wright, Rorty and minimalism. Mind **107**, 157–181 (1998)
6. Boghossian, P.: The status of content. Philos. Rev. **99**(2), 157–184 (1990)
7. Brandom, R.: Making It Explicit: Reasoning, Representing, and Discursive Commitment. Harvard University Press, Cambridge (1998)
8. Brandom, R.: Articulating Reasons: An Introduction to Inferentialism. Harvard University Press, Cambridge (2000)
9. Brown, J.R.: Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures. Routledge, London (1999)
10. Carnap, R.: The elimination of metaphysics through logical analysis of language. Erkenntnis **2**, 219–241 (1932) (trans.: Pap, A. In: Ayer, A.J. (ed.) Logical Positivism, pp. 60–81. Free Press, New York (1959))
11. Carnap, R.: Testability and meaning. Philos. Sci. **3**(4), 419–471 (1936)
12. Carter, A., Chrisman, M.: Is epistemic expressivism incompatible with inquiry? Philos. Stud. **159**(3), 323–339 (2012)
13. Catton, P., Montelle, C.: Traces upon the Rhodian shore: the Greek mathematical diagram as metonym for an epiphany (2010, unpublished manuscript)
14. Chrisman, M.: Expressivism, inferentialism, and saving the debate. Philos. Phenomenol. Res. **77**(2), 334–358 (2008)
15. Dipert, R.R.: Reflections on iconicity, representation, and resemblance: Piece's theory of signs, Goodman on resemblance, and modern philosophies of language and mind. Synthese **106**, 373–397 (1996)
16. Frege, G.: Begriffsschrift (trans.: Beaney, M.). In: Beaney, M. (ed.) The Frege Reader, pp. 47–78. Blackwell, Oxford (1997)
17. Giaquinto, M.: Visual Thinking in Mathematics. Oxford University Press, Oxford (2007)
18. Goodman, N.: Languages of Art: An Approach to a Theory of Symbols. Hackett, London (1976)
19. Legg, C.: The problem of the essential icon. Am. Philos. Q. **45**(3), 207–232 (2008)

20. Legg, C.: 'The hardness of the iconic must': can Peirce's existential graphs assist modal epistemology? Philos. Math. **20**(1), 1–24 (2012)
21. Macarthur, D., Price, H.: Pragmatism, quasi-realism and the global challenge. In: Misak, C. (ed.) The New Pragmatists, pp. 91–121. Oxford University Press, Oxford (2007)
22. Mumma, J.: Proofs, pictures and Euclid. Synthese **175**(2), 255–287 (2010)
23. Netz, R.: The limits of text in Greek mathematics. Boston Stud. Philos. Sci. **238**(4), 161–176 (2005)
24. Peirce, C.S.: In: Hartshorne, C., Weiss, P. (eds.) Collected Papers. Harvard University Press, Cambridge (1931–1958)
25. Peirce, C.S.: In: Houser, N., Kloesel, C. (eds.) Essential Peirce, vol. 1: Selected Philosophical Writings (1867–1893). Indiana University Press, Indianapolis (1992)
26. Peirce, C.S.: In: Houser, N., Kloesel, C. (eds.) Essential Peirce, vol. 1: Selected Philosophical Writings (1893–1913). Indiana University Press, Indianapolis (1998)
27. Perry, J.: The problem of the essential indexical. Noûs **13**, 3–21 (1979)
28. Price, H.: Two paths to pragmatism II. Eur. Rev. Philos. **3**, 109–148 (1998)
29. Price, H.: One cheer for representationalism (2008, unpublished manuscript)
30. Price, H.: Expressivism for two voices (2011, unpublished manuscript)
31. Putnam, H.: Reason, Truth and History. Cambridge University Press, Cambridge (1981)
32. Putnam, H.: Realism with a Human Face. Harvard University Press, Cambridge (1992)
33. Quine, W.V.O.: On what there is. In: From a Logical Point of View, pp. 1–19. Harvard University Press, Cambridge (1953)
34. Roberts, D.: The Existential Graphs of Charles S. Peirce. Mouton, The Hague (1973)
35. Shin, S.-J.: Kant's syntheticity revisited by Peirce. Synthese **113**, 1–41 (1997)
36. Shin, S.-J.: The Iconic Logic of Peirce's Graphs. MIT Press, Cambridge (2002)
37. Wittgenstein, L.: Tractatus Logico-Philosophicus (trans.: Anscombe, G.E.M.). Routledge, London (1961)
38. Zeman, J.: Peirce's philosophy of logic. Trans. Charles S. Peirce Soc. **22**, 1–22 (1986)
39. Zeman, J.: Introduction. The graphical logic of C.S. Peirce. http://www.clas.ufl.edu/users/jzeman/ (2002). Accessed 21 Feb 2011

C. Legg (✉)
Philosophy Programme, University of Waikato, Private Bag 3105, Hamilton, 3240, New Zealand
e-mail: clegg@waikato.ac.nz

The Geometry of Diagrams and the Logic of Syllogisms

Richard Bosley

Abstract Aristotle accounts for three figures on which syllogisms are formed. On the first figure it is possible to prove the completeness of all of the possible syllogisms. But on the second and on the third figures completion is not possible; therefore, premises based on the second or on the third figure are converted in such a way as to count as premises on the first figure. Aristotle's procedure leads to difficulties discussed and corrected in the course of this paper.

Keywords Short and long lines · Figure · Intervals · Terms · Syllogisms · Proof · Reduction

Mathematics Subject Classification Primary 03A10; Secondary 03B10

The aim of this paper is to argue the dependence of a demonstrative syllogism upon its figure. What a syllogism is can be briefly explained: a syllogism is formed from three linguistic parts: a major premise which affirms or denies a sentence composed of a subject and a predicate and a mark of affirmation or negation along lines suggested by the commentator John Philoponus; see his commentary on Aristotle's *Prior Analytics* [3, 66.27–67.14].

Aristotle's text differs from Alexander's and Philoponus's commentary in this way: having laid out the structure of a figure and its intervals Aristotle remarks on a basis which either is or is not adequate for the formation of a syllogism. But the commentators turn immediately to writing out a sketch of a combination of premises which either is or is not adequate for a conclusion. The commentators accordingly leave out the work which should be done by means of diagrams. Guenter Patzig [2], on the other hand, expresses high respect for the commentary of John Philoponus. W.D. Ross [4], in his commentary on the *Prior Analytics*, promotes the idea of diagrams drawn by Aristotle but not preserved.

Since it is argued that a demonstrative syllogism depends upon its figure, let us turn first to the shape of the figure by defining the three kinds of diagrams which guide the forming of the syllogisms as depending upon the figures presented by the diagrams. I call them Sophie diagrams, distinguishing them from Venn diagrams; the distinction between the two kinds will be made clear.

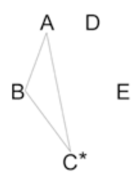
A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic,
DOI 10.1007/978-3-0348-0600-8_2,

1 Affirmation, Negation, and the Resolution of Negation

I shall first try to make clear how a diagram is to be drawn. It is formed by marking points and by drawing lines. For example, a diagram of the first figure is drawn as follows: first, mark a point and let the point indicate a term; Aristotle assigns to such a term—for example—the letter A. We can then draw a line from the point A to the point B. The line indicates an interval terminated by two terms or limits. Every syllogism showing either plain belonging or belonging with necessity has two intervals indicated by the premises of the syllogism or, if the syllogism proves contingency, there are three intervals.

I have so far assumed that the premises which depend upon terms and intervals are affirmative. But affirmation itself is not registered on the diagram. For a diagram is formed from points and lines; premises of a projected syllogism, of course, are said to be either affirmative or negative. Now since affirmation and negation are acts of intelligence but are not either points or lines marked or drawn on a diagram which itself exists not as a spatial but rather as a temporal particular.

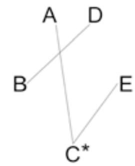
It is of course possible to resolve the negation embedded in a premise: denial is resolved not on a diagram but rather within the mind of the reasoner—resolved in favor either of correcting for deficiency or in favor of correcting for excess. Correction for deficiency is shown by drawing another line on the diagram (we will shortly have an example of this); correction of excess, on the other hand, is shown by adding another point opposite to the point A already registered (assuming that the major premise is negative). The force of correcting for excess is this: there are alternatives and so a choice is made, thinking "Let D belong to the whole of the limit B." But the choice is not joining terms but rather drawing a line from a point to a point. The faculty of reason will, for example, draw a line from the point D to the point B. This drawing helps complete a diagram.



The diagram above is drawn on the first figure which shows the concluding interval A-C composed of letters which stand for terms or limits one of which belongs to the upper end of an interval running to a middle limit; the interval is indicated on the diagram by a straight line running from an upper point, i.e. 'A', to a lower point, i.e. 'C'; the second interval runs from the middle point, i.e. 'B' to the extreme bottom point, i.e. 'C'. The concluding interval A-C is marked with an asterisk indicating that its long interval completes this instance of the first figure. The drawing of the long line is not meant to suggest that the long interval does not follow the two short intervals A-B and B-C.

In the course of the ancient tradition of teaching the syllogistic the commentator Alexander of Aphrodisias [1] tends to conflate an interval and a premise, even though Aristotle himself makes it clear that they are distinct and separate. Had Aristotle's diagrams survived, doubtless the tendency to merge a syllogism and its figure would not have developed.

Aristotle's text differs from Alexander's and Philoponus's commentary in this way: having laid out the structure of a figure and its intervals Aristotle remarks on a basis which either is or is not adequate for the formation of a syllogism. But the commentators turn immediately to writing out a sketch of a combination of premises which either is or is not adequate for a conclusion. The commentators accordingly leave out the work which should be done by means of diagrams.



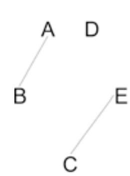
Aristotle does not himself make a distinction between the conclusion of a syllogism and the long interval drawn from the letter ‘A’ to the letter ‘C’. This interval is necessary for completing a second figure.

It is characteristic of every complete figure that there is a middle term whereby the long interval from the top to the bottom of the diagram can be shown. Showing the long interval on the first figure is the objective of a proof on the first figure. It is of course Aristotle’s view that the other two figures do not admit of a proof of their long intervals. For the long intervals are laid down by edict regarding the second two figures—i.e. on the second figure “Let the term A belong to the whole of the extreme limit C and let the term A belong to the whole of the middle term B.” (Discussion of the second two figures will add a crucial qualification to the initial description of the second two figures.)

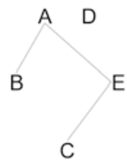
It was Aristotle’s view (as comes clear in early sections of the *Posterior Analytics* [4]) that it is not possible to prove an immediate interval, i.e. “The upper limit A belongs to the whole of the middle limit B.” In fleeing the difficulty of proving the existence of a short interval he confronts another: How is it possible to reshape a diagram in such a way that a long interval uniting two terms becomes a short interval joining two terms?

There is an important respect in which a syllogism differs from a diagram: a diagram is drawn and is not an immediate object of affirmation or negation. It is not a part of a diagram to have a part indicating negation. Negation, after all, is an act and not an object: affirming and denying, on the other hand, are contrary actions and not contrasting objects. For this reason no mark on a diagram indicates either affirming or denying; accordingly, an interval drawn between two points is neither an affirmation nor a negation. Therefore, in drawing a complete diagram it is necessary to resolve the negation embedded in the premises intended for forming parts of a syllogism.

Suppose we have the two premises on the first figure: “the limit A belongs to the whole of the limit B” and “the limit B does not belong to any part of the limit C.” Since a negative premise does not correspond to an interval running between two limits, it is useful to resolve the negation. The principle of the resolution of negation anticipates one of two possible ways in which negation can be resolved: in a first way we correct for the wrong term by drawing a contrary term on the diagram. Suppose we draw a line from the upper letter ‘A’ to the middle letter ‘B’; in resolving the negation of the second premise “the limit B belongs to none of the extreme bottom limit C” we should draw the following diagram on the first figure:



It is apparent that the diagram does not show a continuous long line from A to C. The diagram is therefore not suitable for showing a long interval necessary for a figure and sufficient for the formation of a syllogism. The point is not being made that we turn to showing by counterexamples that there is no conclusion; the point is rather that it is possible to correct the diagram for deficiency by adding an interval between A and E; for it is entirely possible that two species fall under the same genus A. The diagram can now be expanded, yielding:



There is accordingly a long interval A-E-C whereby the middle term E mediates the long interval. The syllogism formed upon the figure need not write a premise which describes the A-B interval.

The middle figure is particularly interesting in that it exemplifies the binary theory of the syllogistic regarding its longitudinal structure; the figure is of course also triadic with respect to a middle term whereby a long interval is mediated. A consideration of the middle figure also provides an opportunity to discuss a difficulty which Aristotle must face in his own dealing with the second figure—a difficulty to which I shall shortly turn.

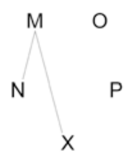
At the center of my study of diagrams and in what ways they provide plans for the formation of syllogisms is its theory of the resolution of negation, briefly discussed above. This theory makes it possible to give a coherent account of diagrams, on the one hand, and, on the other, to show how Aristotle's need of a reduction to the first figure can be eliminated; in place of such a reduction is a procedure for proving three classes of syllogisms: one class depends upon the first, another on the second, and the third class is established on the third figure. The demonstration of the equality of the three figures in relevant ways also provides a rationale for holding that there are only three distinct and separate figures.

The incoherence in Aristotle's execution of his syllogistic program is due to several demands which cannot be coherently satisfied together. To make the conflict clear let us begin with the view of proof which the philosopher evidently endorses: all proof is executed by means of the posit of a middle term. Let us accordingly imagine three terms so arranged that the upper extreme term A belongs to the whole of the middle term B and that the middle term belongs to the whole of the lower extreme term C. If the interval between the limits A and B and the interval between the limits B and C are immediate, Aristotle supposes that the existence of such immediate intervals (each limited by a pair of terms) cannot be proved. The long interval between A and C can be proved by means of a middle term which unites the two intervals A-B and B-C. For this reason Aristotle holds that syllogisms proved on the first figure are complete; but the premises based upon the second and the third figures cannot be the source of a complete syllogism a pair of premises of which has been formed upon the basis of either the second or upon the basis of the third figure. Like wood, prepared for becoming the parts of a chair formed by one mill, but which must be taken to a different mill for becoming a real and complete chair, the premises which originate upon the second or upon the third figure must be transformed into sentences which reflect the first figure thereby becoming the premises of a syllogism which itself depends immediately upon the first figure.

But what do you suppose happens to the sentence underway from the second back to the first figure? When it was first formed on the second figure, let us say, it had a job, namely to describe either a short or a long interval between two terms: perhaps it described a short interval from the limit A running to the limit B; perhaps, on the other hand, it described a long interval from A to C. And yet it is just this responsibility—of bearing a specified use sufficient for representation—that is dropped. For by means of its being converted it falsifies its own prior statement of the interval which was to be represented. The shame of the loss and the deception of Aristotle's pretensions are ameliorated by supposing that there never was a figure or a diagram presenting terms, their intervals, and their modalities—never was a figure, I was about to write, the representation of which was entrusted into the hands of the departing premise or sentence.

The job of a figure or a diagram is to present the terms and their intervals as they are—as we would expect from a floor-plan or from a map of a neighborhood. (A premise, on the other hand, is not a drawing of terms and intervals.) The stability of a diagram is particularly important for its presentation of the modalities of necessity and contingency; for if A is necessary for B, it does not follow that B is necessary for A. If C depends upon B, it does not follow that B depends upon C.

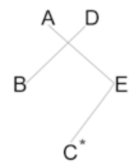
Consider the following two diagrams, defining the second figure ([4], I.5; 26b34–38), and reduction to the first figure ([4], I.5; 26b36–39):



Suppose that a diagram of the second figure shows both a long line connecting the letters 'M' and 'X', as the diagram just above shows, and also shows a short line connecting 'M' and 'N'. But the information provided by the diagram is destroyed by converting the long line 'M-X' into a short line 'M-X' shown above on the second diagram. In trying to show that there is a short, concluding line from 'M' to 'X', conversion turns the initial long line 'M-X', into a short line 'M-X'; this move destroys the long line which was meant to represent a long interval between M and X. So Aristotle's solution to the problem of forming syllogisms on the second and the third figures is in effect to undercut the point of the second and of the third figures and to undercut the uses of sentences which originally have uses whereby a grid of terms and relations can be indicated. We should accordingly expect that Aristotle's mishandling of the second and the third figures would have the effect of making the syllogistic a largely linguistic matter which would be interpreted pretty much as Alexander and John Philoponus did interpret it. Further, such an interpretation blurs a picture and a diagram of the general structure of the syllogistic: first comes the formation of demonstrative syllogisms ([4], I.4ff) upon the basis of a figure diagrammed to show a grid of terms and relations and then, having completed the project of the forming of syllogisms, Aristotle turns ([4], I.32ff) to the dissolution of a syllogism back down to its proper figure.

But let us return to my story of the black-market which has flourished in the obscurity of what happens when sentences formed as premises on either the second or the third figure are transported to the first figure now pregnant with a syllogism, illegitimate on the first figure but no longer at home on the second or the third figures.

Here is an example of the second figure and its diagram:



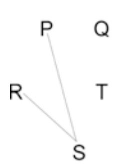
The two lines D-B and A-E-C indicate two intervals one long and one short. (This instance of the binary theory is the result of the resolution of the negation embedded in what Alexander would call a major, universal, and privative premise.) It is obvious that the short and concluding interval does not run between the terms B and C; for in that case the bottom extreme would fall under contrary genera—which is impossible. Therefore, there is a short and concluding interval between the terms E and C forming the concluding interval necessary for completing both the figure and the syllogism formed upon the basis of it.

The long line A-C represents the long interval A-C which in truth is mediated by the term E. The long interval being mediated by means of the middle limit E is not necessarily shown on the diagram drawn just above; nor is the fact mentioned in the syllogism formed on the basis of the second figure. If we were to draw the long interval first from A to E and then from E to C, we would have a different version of the proof of mediation (namely in the style of the second figure and not in that of the first figure). The question arises what difference there is between the proof which depends upon a proper figure and the proof which depends upon contraposition and return to the first figure for the basis of the syllogism. A syllogism on the second figure can be formed as follows:

1. All B are D (or, restoring negation): "No B is A"
2. All C are A
3. So, no C are B

It is evident upon the basis of the diagram that if some C were B, some part of B would fall under a contrary upper term, which is impossible. It is therefore not necessary either to convert or to contraposit the names of the terms in order to form the conclusion.

But let us consider Aristotle's text further—a text which makes it clear in what way the philosopher plays unfairly with the formation of the third figure and therefore with the syllogism to be formed upon its basis. We again require two diagrams: the first is the initial sketch of the third figure; the second shows the result of conversion in order to forge a copy of the first figure. According to Aristotle's suggestions for drawing the two figures, here again is one possibility:



Defining the third figure ([4], I.6; 28a10–22)

Reduction to the first figure ([4], I.6; 28a10–17)

In defining the third figure Aristotle identifies the possibilities whereby the same limit S is the common term to which both the upper term and the middle term belong, as indicated on the third figure. Aristotle remarks at the beginning of I.6 that he calls middle that of which both the limits predicated are predicated. The extreme and major limit P is further from the middle limit S; the limit closer to P is minor. The middle limit is outside the extreme limits and is last by position. Therefore, a complete syllogism does not come to be nor a complete syllogism which depends upon this figure. But a syllogism will be possible both when the terms in relation to the middle limit are whole and not whole.

In order to correct for the deficiency of the position of the limit S we convert the minor interval and thereby make our return to the first figure holding now that P belongs to the whole of S and that S belongs to some R, as the second diagram just above shows. Accordingly, it is proved on the first figure that the limit P belongs to part of the lower limit R.

When a figure has been identified and an interval drawn on a diagram according to which the limit B does not belong to any part of the lower limit C, an argument can be made similar to that argument laid out above concerning the second figure. The denial refers to the middle limit B, to be sure, but does not indicate that middle term by means of which a long interval can be identified by means of which the long interval is mediated. We are therefore ignorant of how the force of the upper term is delivered to the lower term.

The two mistakes attributed to Aristotle's handling of the second figure have counterparts in his handling of the third figure: rather than converting the upper interval he now converts the lower interval R-S forming the short interval S-R. The two unlawful conversions somewhat differ: the first disturbs the interval from the top down; the second disturbs the interval from the middle down.

What is disturbing in particular is the distinct roles of the two forms of modality: the upper term may be necessary for the middle term but the upper term does not determine the middle; the middle is rather determined by the bottom term. So Aristotle's solution to the problem of forming syllogisms on the second and the third figures is in effect to undercut the point of the second and of the third figures and to undercut the uses of sentences which originally have uses whereby a grid of terms and relations can be indicated.

A solution of the difficulty is possible in two steps: the first is an adequate diagram whereby the binary theory which represents two separate terms at the top of the diagram is given adequate representation; the binary theory also requires a noting of two middle terms; the lower level requires only one indication of a term usually taken by the commentators to be a class of particulars (something which Aristotle himself suggests in the *Prior Analytics* [4, I.27]).

Let us turn now to the logic of syllogisms which employ the modalities of necessity and contingency. The syllogisms which posit necessity have aroused less puzzlement than those which posit contingency. But even so, a certain disagreement broke out presumably during Aristotle's life concerning the logic of necessary syllogisms and, in particular, the conclusion which follows from premises one of which indicates pure belonging and the other necessity.

Diagrams can help eliminate irrelevant confusion. Consider the following example.

As the diagram stands, we can interpret the figure which it presents in the following way: the upper limit A is necessary for the whole of the limit B; the limit B, in turn, simply belongs to the whole of the extreme limit C. The part of the diagram drawn as ‘n-’ indicates not the necessity of the term A for the whole of the extreme lower term C; it rather indicates the modality of the concluding and long interval A-C, namely as necessary for the completion of the figure. By contrast the mark ‘n’ attached to the short line A-B (which presents the short upper interval A-B) indicates the modality of the upper term A for the whole of the middle term B. It follows that the diagram does not present the upper term as being necessary for the extreme term C. But commentators who assume that the middle term B is contained within the upper term A and that the lower term C is contained within the middle term B, are bound to confuse the two uses of the modalities.

A point made about the diagramming of the second figure can be made again: the long line indicates the long interval A-C but without showing the course of the long interval and thereby making clear what the mediating term is (nor would this information be shown by a syllogism which depends upon the figure); we are also not told how the middle term is able to receive an interval from a necessary upper term and mediate the long interval in such a way that the interval B-C is of plain belonging. Nevertheless, the interval is necessary for the coherence of the figure and for the syllogism formed upon the figure.

A syllogism formed on the third figure can be written as follows:

1. All C are B.
2. All C are A.
3. So, all B are A.

It is again evident upon the basis of the diagram that if some B are not A, then some part of C would be both A and not A. It is equally obvious that the long interval is mediated by the middle limit B. This feature of the figure is not made evident by a syllogism based upon it.

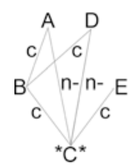
In Sect. I.14 of the *Prior Analytics* [4] Aristotle continues his study of modal syllogisms and their figures by taking up contingency. Just as some terms are necessary for others, so some terms are contingent for others. Unlike the posit of necessity the posit of contingency requires a correlated posit of two terms from which two intervals connect with a single term for which there is contingency. For example, both sitting and walking are contingent for humans and waking and sleeping are contingent for animals. A diagram which presents such terms and the modality of contingency requires two upper terms and a single lower term upon which the two intervals converge. It is not possible to draw a diagram of this sort of modality without a coordinated pair of intervals converging upon a lower term.

It was a mistake of Aristotle’s to try to expand the syllogistic into both the modality of necessity and the modality of contingency while maintaining a single column of terms. It was in fact a strange mistake for the philosopher to make since he was aware of and worked within a binary theory in several respects. Aristotle follows a modal path which is particularly rocky when some intervals are necessary, some contingent and others mixed both with the modalities and also with plain belonging. Aristotle has succeeded in puzzling his commentators; he leaves some modern commentators both perplexed and dismayed. Application of a binary theory is the least of what help lies at hand along with the resolution of negation. It is then possible to sketch diagrams which make consistent sense of the philosopher's project.

Let's begin with contingency defined in such a way as to accommodate the following diagrams. The text in question begins with Sect. I.14 of the first part of the *Prior Analytics* [4]. Aristotle writes:

> So when the limit A is contingent for the whole of the limit B and the limit B for the whole of the limit C, there will be a complete syllogism [showing] that the limit A is contingent for belonging to the whole of the limit C. This is plain from the definition; for we spoke in this way of the fact that being contingent for belonging to the whole is so. [4, I.14; 32b38–33a1]

The following diagram presents the figure described above:



Four premises need to be written in order to form a syllogism which depends upon the whole of the figure represented by the diagram above. The first premise represents the short contingent interval A-B; the counterpart premise represents the second short contingent interval D-B; the lower premise represents the short contingent interval B-C, and the second lower premise represents the short contingent interval E-C. There are two concluding long intervals both necessary for the completion both of the figure and also for the completion of the syllogism formed upon the basis of the figure. The course of the first long interval is evident, namely mediated by the middle term B. But how is the second long interval mediated? The binary theory applied to this instance of a figure suggests mediation by the middle term E.

Before turning to counterexamples employed by Aristotle to prove that there are no further syllogisms in addition to those already proved there is a question as to whether a distinction between the ontic and the modal should be part of the logic of syllogisms. If we begin with actuality—as Aristotle seems to do in laying down paradigms of figures—we may assume that Aristotle does not suppose that a posit of actuality implies the presence of necessity; on the other hand, he does suppose that a posit of necessity implies a posit of actuality; he makes an analogous assumption regarding contingency and actuality. And yet it is clear that a posit of actuality does not cohere with a posit of contingency. It is possible to hold that both graying and keeping the color of one's hair are contingent for belonging to Fred. The simultaneous actuality of both contingencies is not possible. Therefore, it is possible for there to be the contingency of A for B and also possible for there to be the contingency of D for B.

A parallel question arises regarding necessity, namely whether it is possible to hold that the limit A is necessary for B even though it is not implied that A actually belongs to B. Plausible examples support the suggestion. My neighbor's new house depends upon a foundation necessary for it even though a house has not yet been built.

There remain two matters to be discussed in bringing this paper to a conclusion. The first is about counterexamples and the second and last is about reduction to the impossible.

Regarding the first John Philoponus complains that Aristotle sometimes doesn't do a good job of laying out terms and their relations in showing that a syllogism is not possible. The question appropriate to this paper is whether Aristotle would have used a diagram in giving his argument or would rather have formed premises which are to make it clear that no syllogism can be formed by means of such premises. The first examples appear in I.4 of [4] after Aristotle has shown how terms and intervals in the first figure make syllogisms possible; at I.4; 26a2ff he lays out terms and intervals to show when there is not a figure sufficient for the formation of a syllogism. He writes:

> But if the first interval follows upon the middle term but if the middle limit belongs to none of the extreme limit, there will not be a syllogism of the extreme terms; for nothing necessary results by virtue of these limits and intervals; for it is possible for the first limit to belong both to the whole and to none; but when nothing necessary arises by means of these, there will not be a syllogism. [4, 26a2–8]

It is evident that at the same time as Aristotle lays out his argument, he is not forming a syllogism; he must rather be sketching a diagram in order to show why there cannot be a syllogism. To make his point obvious he then likely mentions terms on a diagram sketched out with indications of limits and intervals. How then does such a diagram look when exemplary terms are indicated on the diagram? It is likely that he would draw two intervals; the first would be drawn from A to B; perhaps we can guess how the second would look upon the basis of his example.

He writes, "terms of belonging to the whole: animal-human-horse; of belonging to none: animal-human-stone." Since Aristotle stands by the principle that syllogisms are formed upon the basis of a figure consisting of three terms and two intervals, he has inadequate means for producing a diagram whereby it is clear that no figure is formed which is sufficient for a syllogism.

Although it is clear to Philoponus (see his commentary [3, 76.1–76.24]) that Aristotle is dealing with terms and not with premises, the commentator does not diagnose Aristotle's failure to lay out a figure by means of a diagram clear enough to persuade us that there is a diagram either deficient or excessive for diagramming a figure—a figure which would itself be adequate for the formation of a syllogism.

The last topic to be taken up in this paper is Aristotle's account of reduction to the impossible. Book II of the *Prior Analytics* takes up topics which seem to imply that there is no perfecting correspondence between a syllogism and its underlying structure of terms and their relations. Aristotle's account of a syllogism with false premises and a true conclusion implies the absence of perfecting correspondence; such a syllogism is therefore not a demonstrative syllogism but should rather be called hypothetical.

There are several questions and difficulties which arise in taking up an examination of Aristotle's account of reduction to the impossible. The first question (1) is just what a syllogism *by the impossible* is; a second question (2) is what the role of reduction to the impossible is in completing the syllogistic project; a third question (3) is how it is possible to present the impossible within the general project of the syllogistic. The fourth question (4), finally, is whether there are two ways of resolving the impossibility shown on a reduction diagram.

Aristotle opens his discussion of a syllogism by or through the impossible in the following way:

> So then what converting is and how [it is effected], depending upon each figure, and what syllogism comes to be, is clear. [4, II.11; 61a17–18]

(Comment. A pair of premises depends upon, for example, the second figure; the application of conversion is required in order to transform a pair of premises on the second figure into the premises of a complete syllogism formed on the first figure. In earlier parts of this paper it has been shown how there can be syllogisms some of which fully depend upon the second and others upon the third figure. It is still unclear how we're to understand reduction to the impossible if the procedure implies that there are syllogisms which fully depend upon the second and the third figures. Aristotle writes:)

> So what then converting is and how on each figure and what syllogism comes to be, is clear. [4, II. 61a17–25]

(Comment. Aristotle's summary remark leaves the question unanswered just when and how a syllogism is formed on either the second or the third figure. When we left the first book, we had no reason provided as to how complete syllogisms have been formed on the second and the third figures. But concerning a syllogism by the impossible Aristotle writes:)

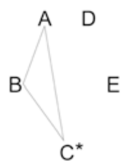
> Now the syllogism by the impossible is demonstrated when the contradictory of its conclusion has been set out and another premise is taken in addition; [such syllogisms come to be depending, respectively, upon] all the figures: for [the procedure] is like conversion, although it differs to the extent that [the conclusion] of the syllogism, which has come to be and of both of the premises which have been assumed, is reduced to the impossible not by virtue of the fact that the contrary has been agreed to before, but rather because it is apparent that it is true. And the terms are disposed in a similar way in both [figures]. For example, if the limit A belongs to the whole of the limit B, and the limit C is middle and if the limit A is set down either belonging not to the whole or to none of the limit B but to the whole of the limit C. [4, II.11; 61a17–25]

The first question stated above is answered in this way: a syllogism by the impossible is a syllogism which is demonstrated with the help of a second syllogism, a premise of which is the contrary or the contradictory of the conclusion of the first syllogism; a premise of the first syllogism is copied into the framework of the second syllogism. A conclusion emerges which is in apparent conflict with one of the premises of the original syllogism. Some debater in the discussion, who holds that premise dear, may find the conclusion of a reduction syllogism intolerable and therefore endorses the conclusion of the original syllogism.

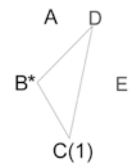
Aristotle contrasts two procedures for generating a second syllogism: just as conversion of a premise (part of a pair of premises on either the second or the third figure) results in a syllogism constructed upon the first figure, so reduction to the impossible results in a second syllogism which is based on the first figure (as Aristotle remarks in I.7) one premise of which is inconsistent with the conclusion of the second syllogism.

But the two procedures differ in this way: conversion begins with a syllogism already agreed upon; reduction to the impossible is different. For the conclusion of the first syllogism is not agreed upon; the second syllogism is constructed beginning with the denial of the conclusion of the first syllogism. For, in constructing the second syllogism, one assumes that the conclusion of the first syllogism has not been proved (or is not evident to those who are arguing with one another). Further, an additional premise for the second syllogism is copied from a premise in the first syllogism. The conclusion of the second syllogism conflicts with a premise helping to compose the first syllogism—a premise held to be true by those arguing, as just suggested.

In Sect. 12 Aristotle remarks that all the projections in the first figure are proved by means of the impossible; the whole affirmative projection based on the first figure is not proved by the impossible [4, II.12; 62a20–22]. If we draw a proper diagram, it is possible to suggest a reason why Aristotle thinks that reduction to the impossible is not possible regarding the first figure. Consider the following pair of diagrams:

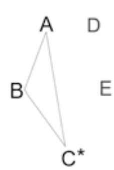


A diagram of a first figure wanting proof

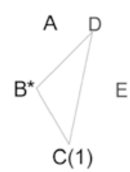


A diagram of a reduction to the impossible

It is evident from the diagrams above that the reduction diagram posits the denial of the long concluding interval charted on the first diagram: that denial is read from the long interval from D to part of the limit C. This first step is not problematical for Aristotle. We then copy the interval indicated by the minor premise into the reduction diagram. The interval and their terms yield a conclusion according to which the limit D belongs to some part of the middle term B. It is commonly supposed that if there is denial of the major interval A-B, and if we assume that negation is embedded not in the interval but in the premise "A-B", it follows that the statement of the denial yields: "So D belongs to some B". But this cannot be since no highest term belongs to part of the middle term when it belongs at all. On the other hand, if we were denying a major premise on account of universal quantification, we would take the resulting major premise to be "So D belongs to some B". But the middle term is neither quantified nor universalized. Therefore, the diagrams should be paired as follows:

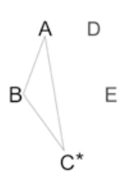


A diagram of a first figure wanting proof

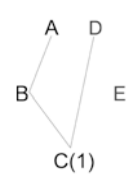


A diagram of the third figure

The second diagram is not part of the process of reduction to the impossible—a process in the service of proving the validity of the diagram on the left and therefore proving the validity of a syllogism formed upon the basis of the figure. Aristotle's mistake is perhaps evident: if we're to show the necessity of the long interval A-C for completing both the figure and also the syllogism formed upon the basis of the figure, we must posit the contrary of the long concluding interval A-C. When we do so, we form the following pairs of diagrams:



A diagram of a first figure wanting proof



A diagram showing the impossible

The point is not to continue syllogistic formation as if to complete a demonstration of impossibility. A syllogism itself is not defined for that task; the task rather falls to the drawing of diagrams which, at the same time, are not intended to present the impossible buried in its figure. What we need is what the second diagram just above gives us, stripped bare of any presumption to reveal the actual situation within a syllogistic figure. No part of the bottom term can fall under both of the two upper indications of the upper terms.

The final question before us is whether there are two ways of resolving the impossibility. With the resolution of the negative prefix 'im' I answer in the affirmative and in the following way. When impossibility becomes apparent upon the basis of a diagram, the faculty of intellect can return to itself with an apparent need to resolve the impossibility either for deficiency or for excess. Resolution of the impossible opens upon two paths: correcting for deficiency takes one path and posits possibility; correcting for excess takes the other and posits the necessity of the syllogism. The necessity is hypothetical necessity with its subject as the figure under discussion. But what is proved with hypothetical necessity is not the relative necessity of a final interval relative to the other intervals described by the diagram; nor is the relative necessity of the conclusion proved relative to its premises.

It is tempting to think that a choice between two courses of resolution depends upon the structure of the two sorts of diagrams: the first plotting a long interval as conclusion and the second two plotting short intervals as conclusions. If a reduction diagram shows impossibility regarding a sketch of a syllogism, resolution of the impossible opens upon two paths, as suggested just above: correcting for deficiency takes place on one path and writes 'possibility' below that projected syllogism; correcting for excess takes the other and posits the necessity of the syllogism. The modality is necessity, and its subject is the syllogism in question. If a reduction diagram rather shows possibility regarding a sketch of a syllogism, the diagram in question is itself possible but not sufficient for the construction of a syllogism.

At the beginning of the paper I argued that the formation of a demonstrative syllogism should not depend upon the conversion or the contraposition of premises of a syllogism to be. The paper has shown how syllogisms can be completed without invoking either conversion or contraposition. Reinforcement is provided by means of diagrams. I hope it has come clear how Sophie diagrams and Venn diagrams differ: by virtue of overlapping a Venn diagram denies the role of intervals which keep the representation of terms apart. A Sophie diagram, on the other hand, give space for all the parts.

References

1. Alexander of Aphrodisias: Commentary on Aristotle's Prior Analytics. Wallies, M. (ed.). Berlin (1883) (CAG II.1)
2. Patzig, G.: Die Aristotelische Syllogistik. Vandenhoeck & Ruprecht, Goettingen (1963)
3. Philoponus, J.: Commentary on Aristotle's Prior Analytics. Wallies, M. (ed.). Berlin (1905) (CAG XIII.2)
4. Ross, W.D.: Aristotle's Prior and Posterior Analytics, A Revised Text with Introduction and Commentary. Clarendon, Oxford (1949)

R. Bosley (✉)
Professor Emeritus, Department of Philosophy, University of Alberta, 2-40 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada
e-mail: rbosley@ualberta.ca

A Diagrammatic Calculus of Syllogisms

Ruggero Pagnan

Abstract A diagrammatic calculus of syllogisms is introduced and discussed, so that a syllogism is valid if and only if its conclusion follows from its premises by calculation. The calculus at issue allows the easy retrieving of the traditional rules of the syllogism and of the laws of the square of opposition. Moreover, it extends to n-term syllogisms and to syllogisms with complemented terms. In this respect, a comparison with De Morgan's *spicular notation* is treated.

Keywords Syllogism · Syllogistic diagram · Syllogistic inference

Mathematics Subject Classification 03B65

1 Introduction

In this paper we aim at the extended description of a diagrammatic formalism supporting a calculus for the syllogistic reasoning. More precisely, we introduce suitable diagrammatic linear representations of the fundamental Aristotelian categorical propositions and show that they are closed under the syllogistic canon of inference which is the deletion of the middle term, so peculiarly implemented to let the formalism simultaneously incorporate a graphical appearance and an algebraic nature. We prove that a syllogism is valid if and only if its conclusion follows from its premises by calculation. A feature of the calculus at issue is that it is naively algorithmic, meaning that no specific knowledge or particular ability is needed in order to understand it and to use it. It also supports a criterion for the rejection of the invalid syllogisms on the base of which the easy retrieving of the traditional rules of the syllogism is possible. Moreover, it allows to retrieve the laws of the square of opposition too. All these facts are shown in Sect. 3. In Sects. 4 and 5 we show that the calculus extends to n-term syllogisms and to syllogisms with complemented terms, respectively. In Sect. 4 the well known result that the number of valid n-term syllogisms is $3n^2 - n$ is reobtained, see [9]. In Sect. 5 a comparison with De Morgan's spicular notation, see [2], is treated.

We point out that other linear diagrammatic formalisms for the syllogistic reasoning exist, notably [12] and [3] but see [7].

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic, DOI 10.1007/978-3-0348-0600-8_3,

2 Preliminaries on Syllogistic

We here recall some well known facts on syllogistic in its traditional form. In doing this we will introduce some of the terminology and notations that will turn out to be useful in the sequel of the paper. For further details on the subject, the reader may consult [8], for example.

Since Aristotle, syllogistic is based on the following *categorical propositions*:

$$\mathbf{A}_{XY}: \text{ All } X \text{ is } Y \text{ (universal affirmative)}$$
$$\mathbf{E}_{XY}: \text{ No } X \text{ is } Y \text{ (universal negative)}$$
$$\mathbf{I}_{XY}: \text{ Some } X \text{ is } Y \text{ (particular affirmative)}$$
$$\mathbf{O}_{XY}: \text{ Some } X \text{ is not } Y \text{ (particular negative)}$$

in which, the letters X and Y denote meaningful expressions of the natural language, that is *terms*, and will be referred to as *term-variables*. Specifically, the term-variable X is the *subject* whereas the term-variable Y is the *predicate* of the considered proposition. A *syllogism* is an argument form that involves three categorical propositions that are distinguished by referring to them as *first premise*, *second premise* and *conclusion*. More precisely, a syllogism involves exactly three term-variables S, P and M as follows: M occurs in both the premises and does not occur in the conclusion whereas P occurs in the first premise and S occurs in the second premise. The term-variables S and P occur as the subject and predicate of the conclusion, respectively, and are also referred to as *minor term* and *major term* of the syllogism, whereas M is also referred to as *middle term*.

Remark 2.1 What we are referring to as syllogisms are *traditional syllogisms* in the terminology of [8], where a detailed discussion of the difference between this notion and that of *Aristotelian syllogism* can be found. Such a difference will not affect the present treatment. We only mention that according to [8], an Aristotelian syllogism is a proposition of the type "If A and B, then C", whereas a traditional syllogism is an argument form with two premises and one conclusion like "A, B therefore C", which in its entirety does not form a compound proposition. Thus, whereas an Aristotelian syllogism can either be true or false, a traditional syllogism can either be valid or not. However, there is to say that considering a syllogism as a proposition is contested by some authors, see [1] for example, and that in general the notion of syllogism still deserves to be debated. The interested reader is invited to consult [11].

The *mood* of a syllogism is the sequence of the kinds of categorical propositions by which it is formed. The *figure* of a syllogism is the position of the term-variables S, P and M in it. There are four possible figures as shown in Table 1 and a syllogism is completely determined by its mood and by its figure together.

Remark 2.2 Undoubtedly the first three figures in Table 1 were introduced by Aristotle. Concerning the controversy about who introduced the fourth one, we mention that according to [8] Aristotle was completely aware of the existence of it and of all the valid syllogistic moods for it, see Tables 2 and 3 below.

Table 1 Figures of the syllogisms

	Fig. 1	Fig. 2	Fig. 3	Fig. 4
first premise	MP	PM	MP	PM
second premise	SM	SM	MS	MS
conclusion	SP	SP	SP	SP

We write syllogisms so that their mood and figure can be promptly retrieved, by also letting the symbol $\models$ separate the premises from the conclusion, as in

$$\mathsf{A}_{MP}, \mathsf{A}_{SM} \models \mathsf{A}_{SP}$$

for example, where it is possible to recognize from left to right the first premise, the second premise and the conclusion, the mood which is **AAA**, and the figure which is the first one. Tables 2 and 3 list the syllogisms which are known to be valid since Aristotle. They are 24 in total, divided into 15 plus 9, the latter being those syllogisms that are also said to be *strengthened*, that is valid under *existential import*, which is the explicit assumption of existence of some S, M or P, as indicated.

3 The Calculus

In this section we introduce a diagrammatic calculus on the base of which, the valid syllogisms listed in Tables 2 and 3 will turn out to be exactly those whose conclusion follows by calculation.

To each categorical proposition we associate the diagrammatic representations

$$X \xrightarrow{\mathsf{A}_{XY}} Y \qquad X \longrightarrow \overset{\mathsf{E}_{XY}}{\bullet} \longleftarrow Y$$

$$X \longleftarrow \overset{\mathsf{I}_{XY}}{\bullet} \longrightarrow Y \qquad X \longleftarrow \bullet \xrightarrow{\mathsf{O}_{XY}} \bullet \longleftarrow Y$$

to be correspondingly read. We will henceforth collectively refer to them as *syllogistic diagrams*. We think of each syllogistic diagram as to an *abstract copula*, that is

> [...] a formal mode of joining two terms which carries no meaning, and obeys no law except such as is barely necessary to make the forms of inference follow.

quoting from [2].

Two or more syllogistic diagrams can be concatenated and reduced, if possible, by formally composing two or more consecutive and accordingly oriented arrow symbols

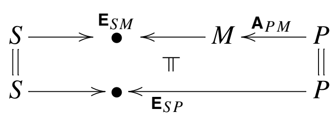
Table 2 The valid syllogisms

Fig. 1	Fig. 2	Fig. 3	Fig. 4
$\mathsf{A}_{MP}, \mathsf{A}_{SM} \models \mathsf{A}_{SP}$	$\mathsf{E}_{PM}, \mathsf{A}_{SM} \models \mathsf{E}_{SP}$	$\mathsf{I}_{MP}, \mathsf{A}_{MS} \models \mathsf{I}_{SP}$	$\mathsf{A}_{PM}, \mathsf{E}_{MS} \models \mathsf{E}_{SP}$
$\mathsf{E}_{MP}, \mathsf{A}_{SM} \models \mathsf{E}_{SP}$	$\mathsf{A}_{PM}, \mathsf{E}_{SM} \models \mathsf{E}_{SP}$	$\mathsf{A}_{MP}, \mathsf{I}_{MS} \models \mathsf{I}_{SP}$	$\mathsf{I}_{PM}, \mathsf{A}_{MS} \models \mathsf{I}_{SP}$
$\mathsf{A}_{MP}, \mathsf{I}_{SM} \models \mathsf{I}_{SP}$	$\mathsf{E}_{PM}, \mathsf{I}_{SM} \models \mathsf{O}_{SP}$	$\mathsf{O}_{MP}, \mathsf{A}_{MS} \models \mathsf{O}_{SP}$	$\mathsf{E}_{PM}, \mathsf{I}_{MS} \models \mathsf{O}_{SP}$
$\mathsf{E}_{MP}, \mathsf{I}_{SM} \models \mathsf{O}_{SP}$	$\mathsf{A}_{PM}, \mathsf{O}_{SM} \models \mathsf{O}_{SP}$	$\mathsf{E}_{MP}, \mathsf{I}_{MS} \models \mathsf{O}_{SP}$	

Table 3 The valid strengthened syllogisms

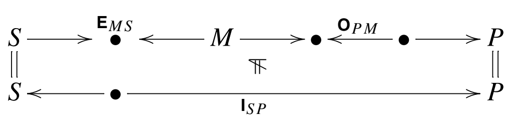
Fig. 1	Fig. 2	Fig. 3	Fig. 4	Assumption
$\mathbf{A}_{MP}, \mathbf{A}_{SM} \models \mathbf{I}_{SP}$	$\mathbf{A}_{PM}, \mathbf{E}_{SM} \models \mathbf{O}_{SP}$		$\mathbf{A}_{PM}, \mathbf{E}_{MS} \models \mathbf{O}_{SP}$	some S exists
$\mathbf{E}_{MP}, \mathbf{A}_{SM} \models \mathbf{O}_{SP}$	$\mathbf{E}_{PM}, \mathbf{A}_{SM} \models \mathbf{O}_{SP}$			some S exists
		$\mathbf{A}_{MP}, \mathbf{A}_{MS} \models \mathbf{I}_{SP}$	$\mathbf{E}_{PM}, \mathbf{A}_{MS} \models \mathbf{O}_{SP}$	some M exists
		$\mathbf{E}_{MP}, \mathbf{A}_{MS} \models \mathbf{O}_{SP}$		some M exists
			$\mathbf{A}_{PM}, \mathbf{A}_{MS} \models \mathbf{I}_{SP}$	some P exists

separated by a single term-variable, thus deleting it. Such a reduction will be henceforth referred to as *syllogistic inference*. By means of syllogistic inferences, the syllogistic diagrams can be used to verify the validity of the syllogisms. This is obtained by using three syllogistic diagrams, as the first premise, the second premise, and the conclusion of the syllogism, involving three distinguished term-variables, denoted S, M and P, in such a way that M occurs in both the syllogistic diagrams in the premises and does not in the conclusion, whereas S and P occur in the conclusion as well as in the premises. Following the tradition, P will occur in the first premise whereas S in the second. We will show that a syllogism is valid if and only if there exists a syllogistic inference from the concatenation of the syllogistic diagrams in the premises to the syllogistic diagram which is the conclusion. Syllogistic inferences will be represented by planar diagrams filled in with the symbol $\models$ upside down, so to explicitly underline the fact that the notion of syllogistic inference is a directed one. Thus, for example, the syllogistic inference associated with the valid syllogism $\mathbf{A}_{PM}, \mathbf{E}_{SM} \models \mathbf{E}_{SP}$ is



(1)

whereas, the syllogism $\mathbf{O}_{PM}, \mathbf{E}_{MS} \models \mathbf{I}_{SP}$ is not valid since its conclusion cannot be obtained by syllogistic inference, as shown by diagram



(2)

in which no formal composition can be computed to delete the middle term M.

Remark 3.1 Verifying the validity of the syllogisms by means of syllogistic inferences is a naively algorithmic procedure as we mentioned in Sect. 1. This is different from the informal employment of Venn's diagrams for the same purpose, because of the necessary ability which is required to the user in order to read off the diagram for the conclusion from the diagrams for the premises. The crucial step consists in understanding if the former is "contained" in the latter or not. Formal reasoning with Venn's diagrams is treated in [5] and [10].

The syllogistic diagram which is the conclusion of a syllogistic inference contains as many bullet symbols as the syllogistic diagrams for the premises. This fact turns out to be useful in showing that a syllogism is not valid. So, without drawing diagram (2), the syllogism in consideration can be immediately said to be not valid since a single bullet symbol occurs in the conclusion, whereas three of them occur in the premises. However, this criterion does not always apply. It suffices to consider the syllogistic inference

$$\begin{array}{ccccccccc} S & \longrightarrow & \overset{\mathsf{E}_{SM}}{\bullet} & \longleftarrow & M & \longleftarrow & \overset{\mathsf{I}_{MP}}{\bullet} & \longrightarrow & P \\ \| & & & & \top\!\!\top & & & & \| \\ S & \longrightarrow & \bullet & \longleftarrow & \underset{\mathsf{O}_{PS}}{} & & \bullet & \longrightarrow & P \end{array}$$

in which as many bullet symbols occur in the premises as in the conclusion and does not correspond to any valid syllogism, coherently with (3) in Remark 3.5 below, since in the diagram for the conclusion the rôles of S and P has been illicitly exchanged. Anyway, by the application of suitable rules of reduction, already known to Aristotle, specifically by exchanging the order of the premises and by simple conversion on them, see the end of Sect. 4, to let now P become the minor term and S the major term, the previous syllogistic inference can be transformed into the one for the valid first-figure syllogism $\mathsf{E}_{MS}, \mathsf{I}_{PM} \models \mathsf{O}_{PS}$.

The invalid syllogisms can be rejected with the linear diagrams in [3] as well, but for the fact that one has to go through all of them, as explained there.

The following lemma lists the concatenations of pairs of syllogistic diagrams yielding a syllogistic diagram through a syllogistic inference.

Lemma 3.2 *A syllogistic inference applied to a concatenation of two syllogistic diagrams yields a syllogistic diagram in exactly the following cases*:

(i) $S \rightarrow M \rightarrow P$
(ii) $S \rightarrow \bullet \leftarrow M \leftarrow P$
(iii) $S \rightarrow M \rightarrow \bullet \leftarrow P$
(iv) $S \leftarrow M \leftarrow \bullet \rightarrow P$
(v) $S \leftarrow \bullet \rightarrow M \rightarrow P$
(vi) $S \leftarrow \bullet \rightarrow M \rightarrow \bullet \leftarrow P$
(vii) $S \leftarrow M \leftarrow \bullet \rightarrow \bullet \leftarrow P$
(viii) $S \leftarrow \bullet \rightarrow \bullet \leftarrow M \leftarrow P$

Proof Clearly, syllogistic inference applies to each of the diagrams listed in the statement and, making M disappear, yields a syllogistic diagram involving only S and P, as a conclusion. Conversely, by also keeping in mind Remark 3.1, we proceed by cases:

(a) the only way to obtain $S \rightarrow P$ as a conclusion of a syllogistic inference is by (i), since no bullet symbol is allowed to occur in the conclusion.
(b) the only way to obtain $S \rightarrow \bullet \leftarrow P$ as a conclusion of a syllogistic inference is by either (ii) or (iii), since exactly one bullet symbol must occur in the conclusion with two arrow symbols converging to it.
(c) similarly, the only way to obtain $S \leftarrow \bullet \rightarrow P$ as a conclusion of a syllogistic inference is by either (iv) or (v), since exactly one bullet symbol must occur in the conclusion with two arrow symbols diverging from it.
(d) the only way to obtain $S \leftarrow \bullet \rightarrow \bullet \leftarrow P$ as a conclusion of a syllogistic inference is by either (vi), (vii) or (viii), since exactly two bullet symbols must occur in the conclusion, together with three alternating arrow symbols. □

Notation 3.3 Henceforth, the concatenation of syllogistic diagrams will be denoted by $\sharp$ and the syllogistic inferences also written in line, so that for example the syllogistic inference represented by diagram (1) will be also written

$$(\mathbf{A}_{PM})\sharp(\mathbf{E}_{SM}) \models (\mathbf{E}_{SP})$$

The next theorem shows that the syllogisms in Table 2 are exactly those that are valid by the method of syllogistic inference.

Theorem 3.4 *A syllogism is valid if and only if there is a* (*necessarily unique*) *syllogistic inference from its premises to its conclusion.*

Proof On one hand it suffices to explicitly construct a syllogistic inference for each of the syllogisms in Table 2, as follows:

Figure 1:

$$(\mathbf{A}_{MP})\sharp(\mathbf{A}_{SM}) \models (\mathbf{A}_{SP})$$
$$(\mathbf{E}_{MP})\sharp(\mathbf{A}_{SM}) \models (\mathbf{E}_{SP})$$
$$(\mathbf{A}_{MP})\sharp(\mathbf{I}_{SM}) \models (\mathbf{I}_{SP})$$
$$(\mathbf{E}_{MP})\sharp(\mathbf{I}_{SM}) \models (\mathbf{O}_{SP})$$

Figure 2:

$$(\mathbf{E}_{PM})\sharp(\mathbf{A}_{SM}) \models (\mathbf{E}_{SP})$$
$$(\mathbf{A}_{PM})\sharp(\mathbf{E}_{SM}) \models (\mathbf{E}_{SP})$$
$$(\mathbf{E}_{PM})\sharp(\mathbf{I}_{SM}) \models (\mathbf{O}_{SP})$$
$$(\mathbf{A}_{PM})\sharp(\mathbf{O}_{SM}) \models (\mathbf{O}_{SP})$$

Figure 3:

$$(\mathbf{I}_{MP})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{SP})$$
$$(\mathbf{A}_{MP})\sharp(\mathbf{I}_{MS}) \models (\mathbf{I}_{SP})$$
$$(\mathbf{O}_{MP})\sharp(\mathbf{A}_{MS}) \models (\mathbf{O}_{SP})$$
$$(\mathbf{E}_{MP})\sharp(\mathbf{I}_{MS}) \models (\mathbf{O}_{SP})$$

Figure 4:

$$(\mathbf{A}_{PM})\sharp(\mathbf{E}_{MS}) \models (\mathbf{E}_{SP})$$
$$(\mathbf{I}_{PM})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{SP})$$
$$(\mathbf{E}_{PM})\sharp(\mathbf{I}_{MS}) \models (\mathbf{O}_{SP})$$

On the other hand, it suffices to construct syllogistic inferences to a given possible conclusion.

– By Lemma 3.2 (i), the only way to obtain $\mathbf{A}_{SP}$ as a conclusion is represented by the diagram

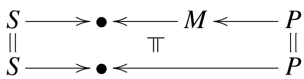
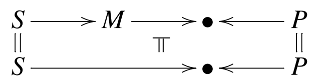
$$\begin{array}{ccccc} S & \longrightarrow & M & \longrightarrow & P \\ \| & & \top & & \| \\ S & & \longrightarrow & & P \end{array}$$

which exactly corresponds to the syllogistic inference

$$(\mathbf{A}_{MP})\sharp(\mathbf{A}_{SM}) \models (\mathbf{A}_{SP})$$

validating the mood **AAA** in the first figure.

– By Lemma 3.2 (ii) and (iii), the only ways to obtain $\mathbf{E}_{SP}$ as a conclusion, are represented by the two diagrams

The leftmost can be read as either the syllogistic inference

$$(\mathbf{A}_{PM})\sharp(\mathbf{E}_{SM}) \models (\mathbf{E}_{SP})$$

or

$$(\mathbf{A}_{PM})\sharp(\mathbf{E}_{MS}) \models (\mathbf{E}_{SP})$$

which validate the mood **AEE** in the second and fourth figures, respectively. The rightmost can be read as either the syllogistic inference

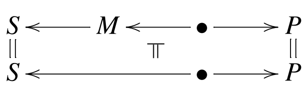
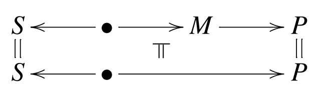
$$(\mathbf{E}_{MP})\sharp(\mathbf{A}_{SM}) \models (\mathbf{E}_{SP})$$

or

$$(\mathbf{E}_{PM})\sharp(\mathbf{A}_{SM}) \models (\mathbf{E}_{SP})$$

which validate the mood **EAE** in the first and second figures, respectively.

– By Lemma 3.2 (iv) and (v), the only ways to obtain $\mathbf{I}_{SP}$ as a conclusion is represented by the two diagrams

The leftmost can be read as either the syllogistic inference

$$(\mathbf{I}_{MP})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{SP})$$

or

$$(\mathbf{I}_{PM})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{SP})$$

which validate the mood **IAI** in the third and fourth figures, respectively. The rightmost can be read as either the syllogistic inference

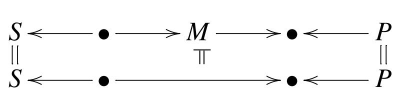
$$(\mathbf{A}_{MP})\sharp(\mathbf{I}_{SM}) \models (\mathbf{I}_{SP})$$

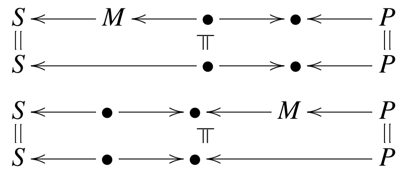
or

$$(\mathbf{A}_{MP})\sharp(\mathbf{I}_{MS}) \models (\mathbf{I}_{SP})$$

that validate the mood **AII** in the first and third figures, respectively.

– By Lemma 3.2 (vi), (vii) and (viii), the only ways to obtain $\mathbf{O}_{SP}$ as a conclusion is represented by the three diagrams





The first can be read as any of the syllogistic inferences

$$(\mathbf{E}_{MP})\sharp(\mathbf{I}_{SM}) \models (\mathbf{O}_{SP})$$
$$(\mathbf{E}_{PM})\sharp(\mathbf{I}_{SM}) \models (\mathbf{O}_{SP})$$
$$(\mathbf{E}_{MP})\sharp(\mathbf{I}_{MS}) \models (\mathbf{O}_{SP})$$
$$(\mathbf{E}_{PM})\sharp(\mathbf{I}_{MS}) \models (\mathbf{O}_{SP})$$

that validate the mood **EIO** in all the figures. The second can be read as the syllogistic inference

$$(\mathbf{O}_{MP})\sharp(\mathbf{A}_{MS}) \models (\mathbf{O}_{SP})$$

that validates the mood **OAO** in the third figure. The third can be read as the syllogistic inference

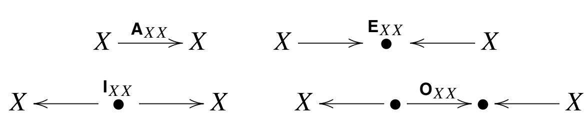
$$(\mathbf{A}_{PM})\sharp(\mathbf{O}_{SM}) \models (\mathbf{O}_{SP})$$

validating the mood **AOO** in the second figure. □

Remark 3.5 After Theorem 3.4, it is an easy exercise to read off the well-known *rules of syllogism* from the list in Lemma 3.2.

(1) From two negative premises nothing can be inferred.
(2) From two particular premises nothing can be inferred.
(3) If the first premise of a syllogism is particular, whereas its second premise is negative, then nothing can be inferred.
(4) If one premise is particular, then the conclusion is particular.
(5) The conclusion of a syllogism is negative if and only if so is one of its premises.

For every term-variable X, particularly interesting instances of syllogistic diagrams are the following:



which must be correspondingly read as

$\mathbf{A}_{XX}$: All X is X

$\mathbf{E}_{XX}$: No X is X

$\mathbf{I}_{XX}$: Some X is X

$\mathbf{O}_{XX}$: Some X is not X

The diagrams $\mathbf{A}_{XX}$ and $\mathbf{I}_{XX}$ represent the *laws of identity*. Diagram $\mathbf{I}_{XX}$ in particular represents existential import, whereas diagram $\mathbf{E}_{XX}$ affirms the emptiness of X. We look at diagram $\mathbf{O}_{XX}$ as expressing contradiction, for reasons that will be more clearly illustrated by Proposition 3.8.

Theorem 3.4 extends to comprise the syllogisms in Table 3, by taking into account existential imports of the forms $\mathbf{I}_{SS}$, $\mathbf{I}_{MM}$, $\mathbf{I}_{PP}$. The first step toward that direction is Lemma 3.6 below.

Lemma 3.6 *A syllogistic inference applied to a concatenation of two syllogistic diagrams and one existential import yields a syllogistic diagram in exactly the following cases*:

(i) $S \leftarrow \bullet \rightarrow S \rightarrow M \rightarrow P$
(ii) $S \leftarrow M \leftarrow \bullet \rightarrow M \rightarrow P$
(iii) $S \leftarrow M \leftarrow P \leftarrow \bullet \rightarrow P$
(iv) $S \leftarrow \bullet \rightarrow S \rightarrow M \rightarrow \bullet \leftarrow P$
(v) $S \leftarrow \bullet \rightarrow S \rightarrow \bullet \leftarrow M \leftarrow P$
(vi) $S \leftarrow M \leftarrow \bullet \rightarrow M \rightarrow \bullet \leftarrow P$

Proof On one hand, it is clear that syllogistic inference applies to the diagrams listed in the statement and, making M disappear, yields a syllogistic diagram involving only S and P as a conclusion. On the other hand, by also keeping in mind Remark 3.1, we proceed by cases:

(a) There is no way to obtain $S \rightarrow P$ as the conclusion of a syllogistic inference under any existential import of the form $\mathbf{I}_{SS}$, $\mathbf{I}_{MM}$ or $\mathbf{I}_{PP}$, because of the presence of one indelible bullet symbol in any of these.
(b) There is no way to obtain $S \rightarrow \bullet \leftarrow P$ as the conclusion of a syllogistic inference under any existential import of the form $\mathbf{I}_{SS}$, $\mathbf{I}_{MM}$ or $\mathbf{I}_{PP}$, because of the presence of one indelible bullet symbol in any of these, together with two morphisms diverging from it.
(c) The only ways to obtain $S \leftarrow \bullet \rightarrow P$ as the conclusion of a syllogistic inference, under any existential import of the form $\mathbf{I}_{SS}$, $\mathbf{I}_{MM}$ or $\mathbf{I}_{PP}$, is by either (i), (ii) or (iii), since exactly one bullet symbol must occur in the conclusion together with two morphisms diverging from it.
(d) There is no way to obtain $S \leftarrow \bullet \rightarrow \bullet \leftarrow P$ as the conclusion of a syllogistic inference under the existential import $\mathbf{I}_{PP}$, since such an assumption necessarily yields a conclusion like $S \cdots \bullet \rightarrow P$ which by no means can be $S \leftarrow \bullet \rightarrow \bullet \leftarrow P$. The only ways to obtain $S \leftarrow \bullet \rightarrow \bullet \leftarrow P$ as the conclusion of a syllogistic inference, under any existential import of the forms $\mathbf{I}_{SS}$ or $\mathbf{I}_{MM}$, is by either (iv), (v) or (vi), since exactly two bullet symbols must occur in the conclusion, together with three alternating morphisms. □

Theorem 3.7 *A syllogism with existential import is valid if and only if there is a syllogistic inference from its premises to its conclusion.*

Proof On one hand it suffices to construct explicitly a suitable syllogistic inference for each of the syllogisms in Table 3, as follows:

Figure 1:

$$(\mathbf{A}_{MP})\sharp(\mathbf{A}_{SM})\sharp(\mathbf{I}_{SS}) \models (\mathbf{I}_{SP})$$

$$(\mathbf{E}_{MP})\sharp(\mathbf{A}_{SM})\sharp(\mathbf{I}_{SS}) \models (\mathbf{O}_{SP})$$

Figure 2:

$$(\mathsf{A}_{PM})\sharp(\mathsf{E}_{SM})\sharp(\mathsf{I}_{SS}) \models (\mathsf{O}_{SP})$$
$$(\mathsf{E}_{PM})\sharp(\mathsf{A}_{SM})\sharp(\mathsf{I}_{SS}) \models (\mathsf{O}_{SP})$$

Figure 3:

$$(\mathsf{A}_{MP})\sharp(\mathsf{I}_{MM})\sharp(\mathsf{A}_{MS}) \models (\mathsf{I}_{SP})$$
$$(\mathsf{E}_{MP})\sharp(\mathsf{I}_{MM})\sharp(\mathsf{A}_{MS}) \models (\mathsf{O}_{SP})$$

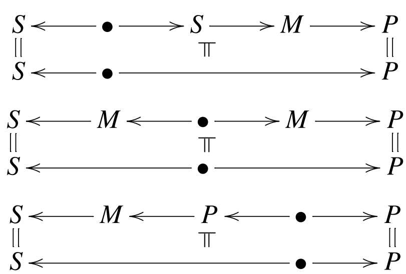
Figure 4:

$$(\mathsf{A}_{PM})\sharp(\mathsf{E}_{MS})\sharp(\mathsf{I}_{SS}) \models (\mathsf{O}_{SP})$$
$$(\mathsf{E}_{PM})\sharp(\mathsf{I}_{MM})\sharp(\mathsf{A}_{MS}) \models (\mathsf{O}_{SP})$$
$$(\mathsf{I}_{PP})\sharp(\mathsf{A}_{PM})\sharp(\mathsf{A}_{MS}) \models (\mathsf{I}_{SP})$$

On the other hand it suffices to construct a syllogistic inference to a given possible conclusion, by taking into account existential imports.

Because of points (a) and (b) in the proof of Lemma 3.6, there is no way to obtain the syllogistic diagrams A_{SP} and E_{SP} as the conclusion of a syllogistic inference under any existential import.

- By Lemma 3.6 (i), (ii), (iii), the only ways to obtain the conclusion I_{SP} under the existential import I_{SS}, I_{MM} or I_{PP} are represented by the three diagrams

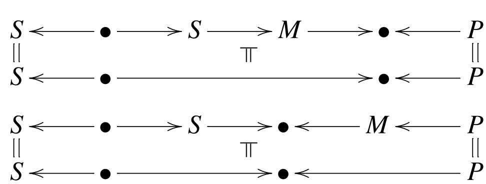


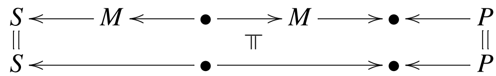
which from top to bottom can be read as the syllogistic inferences

$$(\mathsf{A}_{MP})\sharp(\mathsf{A}_{SM})\sharp(\mathsf{I}_{SS}) \models (\mathsf{I}_{SP})$$
$$(\mathsf{A}_{MP})\sharp(\mathsf{I}_{MM})\sharp(\mathsf{A}_{MS}) \models (\mathsf{I}_{SP})$$
$$(\mathsf{I}_{PP})\sharp(\mathsf{A}_{PM})\sharp(\mathsf{A}_{MS}) \models (\mathsf{I}_{SP})$$

respectively, validating the mood **AAI** in the first, third and fourth figure.

- By Lemma 3.6 (iv), (v), (vi), the only ways to obtain the conclusion O_{SP} is under the existential import I_{SS} or I_{MM}, as represented by the three diagrams





the first of which can be read as either the syllogistic inference

$$(\mathbf{E}_{MP})\sharp(\mathbf{A}_{SM})\sharp(\mathbf{I}_{SS}) \models (\mathbf{O}_{SP})$$

or

$$(\mathbf{E}_{PM})\sharp(\mathbf{A}_{SM})\sharp(\mathbf{I}_{SS}) \models (\mathbf{O}_{SP})$$

that validate the mood **EAO** in the first and second figures, respectively. The second diagram can be read as either the syllogistic inference

$$(\mathbf{A}_{PM})\sharp(\mathbf{E}_{SM})\sharp(\mathbf{I}_{SS}) \models (\mathbf{O}_{SP})$$

or

$$(\mathbf{A}_{PM})\sharp(\mathbf{E}_{MS})\sharp(\mathbf{I}_{SS}) \models (\mathbf{O}_{SP})$$

that validate the mood **AEO** in the second and fourth figures, respectively. The third diagram can be read as either the syllogistic inference

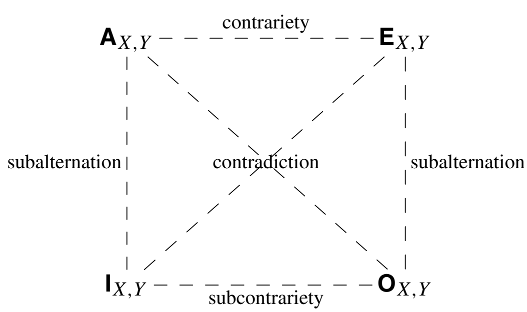
$$(\mathbf{E}_{MP})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{MS}) \models (\mathbf{O}_{SP})$$

or

$$(\mathbf{E}_{PM})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{MS}) \models (\mathbf{O}_{SP})$$

that validate the mood **EAO** in the third and fourth figures, respectively. □

We end the section by briefly discussing the existing connections between the so far described calculus of syllogisms and the laws of the *square of opposition*



in which

– $\mathbf{A}_{XY}$ and $\mathbf{O}_{XY}$, as well as $\mathbf{E}_{XY}$ and $\mathbf{I}_{XY}$, are *contradictory* because they cannot both be true, nor both false. In each pair, one proposition is true if and only if the other is false.
– Under existential import on X, $\mathbf{A}_{XY}$ and $\mathbf{I}_{XY}$ as well as $\mathbf{E}_{XY}$ and $\mathbf{O}_{XY}$, are *subaltern* because $\mathbf{I}_{XY}$ is true if $\mathbf{A}_{XY}$ is true, and $\mathbf{O}_{XY}$ is true if $\mathbf{E}_{XY}$ is true, but not the converse, in both cases.
– Under existential import on X, $\mathbf{A}_{XY}$ and $\mathbf{E}_{XY}$ are *contraries* because they cannot both be true, but can both be false. Each of them implies the negation of the other, but not the converse.

– Under existential import on X, $\mathbf{I}_{XY}$ and $\mathbf{O}_{XY}$ are *subcontraries* because they cannot both be false, but can both be true. The negation of each of them implies the other, but not the converse.

Proposition 3.8 *The laws of the square of opposition can be obtained by syllogistic inference.*

Proof The contradictory relations are calculated by the syllogistic inferences

$$(\mathbf{A}_{XY})\sharp(\mathbf{O}_{XY}) \models (\mathbf{O}_{XX})$$
$$(\mathbf{E}_{XY})\sharp(\mathbf{I}_{XY}) \models (\mathbf{O}_{XX})$$

and this is the reason why we look at $\mathbf{O}_{XX}$ as expressing contradiction, as previously hinted at. The remaining laws are calculated by the syllogistic inferences

$$(\mathbf{A}_{XY})\sharp(\mathbf{I}_{XX}) \models (\mathbf{I}_{XY})$$
$$(\mathbf{E}_{XY})\sharp(\mathbf{I}_{XX}) \models (\mathbf{O}_{XY})$$

Indeed, they immediately provide the subalternation laws. They provide evidence for the laws of contrariety because $\mathbf{I}_{XY}$ is the negation of $\mathbf{E}_{XY}$ and $\mathbf{O}_{XY}$ is the negation of $\mathbf{A}_{XY}$. They provide evidence for the laws of subcontrariety since $\mathbf{A}_{XY}$ is the negation of $\mathbf{O}_{XY}$ and $\mathbf{E}_{XY}$ is the negation of $\mathbf{I}_{XY}$. Moreover, both the syllogistic inferences cannot be reversed since one bullet symbol occurs in $\mathbf{I}_{XY}$ and no bullet symbols occur in $\mathbf{A}_{XY}$, two bullet symbols occur in $\mathbf{O}_{XY}$ and one bullet symbol occurs in $\mathbf{E}_{XY}$, see Remark 3.1. □

4 Extending the Calculus: *n*-Term Syllogisms

Whereas syllogisms, either with existential import or not, involve exactly 3 term-variables, n-term syllogisms involve exactly n term-variables $A_1, \dots, A_n$, $n \geq 1$, linked by n categorical propositions any two contiguous of which have exactly one term in common. The total number of valid n-term syllogisms is $3n^2 - n$, see [9], where such a formula was obtained by rejecting the not valid moods on the bases of the traditional rules of syllogism. The same formula has been reobtained in [12] and [4]. The aim of the present section is that of generalizing Theorems 3.4 and 3.7 to the case of n-term syllogisms and that of recalculating the previously cited formula through the employment of syllogistic inferences.

Lemma 4.1 *For every positive natural number* n, *a syllogistic inference yields a syllogistic diagram as a conclusion in exactly the following cases*:

(i) $A_1 \to A_2 \to \cdots \to A_i \to A_{i+1} \to \cdots \to A_{n-1} \to A_n$.
(ii) $A_1 \to A_2 \to \cdots \to A_i \to \bullet \leftarrow A_{i+1} \leftarrow \cdots \leftarrow A_{n-1} \leftarrow A_n$, *with* $1 \leq i \leq n-1$.
(iii) $A_1 \leftarrow A_2 \leftarrow \cdots \leftarrow A_i \leftarrow \bullet \to A_{i+1} \to \cdots \to A_{n-1} \to A_n$, *with* $1 \leq i \leq n-1$.
(iv) $A_1 \leftarrow A_2 \leftarrow \cdots \leftarrow A_i \leftarrow \bullet \to A_i \to \cdots \to A_{n-1} \to A_n$, *with* $1 \leq i \leq n$.
(v) $A_1 \leftarrow A_2 \leftarrow \cdots \leftarrow A_i \leftarrow \bullet \to \bullet \leftarrow A_{i+1} \leftarrow \cdots \leftarrow A_{n-1} \leftarrow A_n$, *with* $1 \leq i \leq n-1$.
(vi) $A_1 \leftarrow \cdots \leftarrow A_i \leftarrow \bullet \to A_{i+1} \to \cdots \to A_{j-1} \to \bullet \leftarrow A_j \leftarrow \cdots \leftarrow A_n$, *with* $1 \leq i < j \leq n$.

(vii) $A_1 \leftarrow \cdots \leftarrow A_i \leftarrow \bullet \rightarrow A_i \rightarrow \cdots \rightarrow A_{j-1} \rightarrow \bullet \leftarrow A_j \leftarrow \cdots \leftarrow A_n$, *with* $1 \leq i < j \leq n$.

Proof It is clear that a syllogistic inference applies to each of the diagrams listed in the statement yielding a syllogistic diagram involving the terms A_1 and A_n only. Conversely, we proceed by cases:

(a) the only way to obtain $A_1 \rightarrow A_n$ as a conclusion of a syllogistic inference is by (i), since no bullet symbol is allowed to occur in the conclusion.
(b) the only way to obtain $A_1 \rightarrow \bullet \leftarrow A_n$ as a conclusion of a syllogistic inference is by (ii), since exactly one bullet symbol must occur in the conclusion with two morphisms converging to it.
(c) the only way to obtain $A_1 \leftarrow \bullet \rightarrow A_n$ as a conclusion of a syllogistic inference is by (iii) or (iv), since exactly one bullet symbol must occur in the conclusion with two morphisms diverging from it.
(d) the only way to obtain $A_1 \leftarrow \bullet \rightarrow \bullet \leftarrow A_n$ as a conclusion of a syllogistic inference is by (v), (vi) or (vii), since exactly two bullet symbols must occur in the conclusion, with three alternating morphisms. □

Lemma 4.2 *For every positive natural number* n, *let* $\varphi(n)$ *and* $\psi(n)$ *be the number of diagrams like those in points* (vi) *and* (vii) *of Lemma* 4.1, *respectively. The following facts hold*

(i) $\varphi(n) = \frac{(n-1)(n-2)}{2}$.
(ii) $\psi(n) = \frac{n(n-1)}{2}$.

Proof (i) For every positive natural number n, $\varphi(n+1) = \varphi(n) + (n-1)$. Because, passing from n to $n+1$ is a matter of inserting one more arrow symbol $\rightarrow$ or $\leftarrow$, on the left, on the right or in the middle of the diagrams constructed at n, so to extend them with one more term-variable. There are exactly $n-1$ possibilities of doing this. Finally, by induction on the number of term-variables, the thesis is easily achieved.

(ii) The argument is completely similar to the previous but for the fact that for every positive natural number n, $\psi(n+1) = \psi(n) + n$. □

Theorem 4.3 *For every positive natural number* n, *an* n*-term syllogism is valid if and only if there is a syllogistic inference from its premises to its conclusion. Moreover, the number of valid* n*-term syllogisms is* $3n^2 - n$.

Proof Lemmas 4.1 and 4.2, permit to conclude that the n-term syllogisms in Table 4 are all valid. Moreover they are $3n^2 - n$. Conversely, we use Lemmas 4.1 and 4.2, to construct a syllogistic inference to a given possible conclusion:

– By Lemma 4.1 (i), the diagram

$$\begin{array}{ccccccccccccc}
A_1 & \longrightarrow & A_2 & \longrightarrow & \cdots & \longrightarrow & A_i & \longrightarrow & A_{i+1} & \longrightarrow & \cdots & \longrightarrow A_{n-1} \longrightarrow & A_n \\
\| & & & & & & & & \top\!\top & & & & \| \\
A_1 & \longrightarrow & & & & & & & & & & & A_n
\end{array}$$

represents the only way to produce evidence for the syllogistic inference

$$(\mathbf{A}_{A_{n-1}A_n})\sharp\cdots\sharp(\mathbf{A}_{A_1A_2}) \models (\mathbf{A}_{A_1A_n})$$

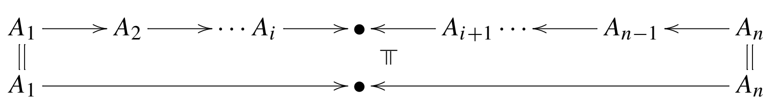
Table 4 The valid n-term syllogisms

Syllogism	Quantity
$\mathbf{A}_{A_{n-1}A_n}, \ldots, \mathbf{A}_{A_1A_2} \models \mathbf{A}_{A_1A_n}$	1
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_iA_{i+1}}, \ldots, \mathbf{A}_{A_1A_2} \models \mathbf{E}_{A_1A_n}$	$n-1$
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_{i+1}A_i}, \ldots, \mathbf{A}_{A_1A_2} \models \mathbf{E}_{A_1A_n}$	$n-1$
$\mathbf{A}_{A_{n-1}A_n}, \ldots, \mathbf{I}_{A_iA_{i+1}}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{I}_{A_1A_n}$	$n-1$
$\mathbf{A}_{A_{n-1}A_n}, \ldots, \mathbf{I}_{A_{i+1}A_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{I}_{A_1A_n}$	$n-1$
$\mathbf{A}_{A_{n-1}A_n}, \ldots, \mathbf{I}_{A_iA_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{I}_{A_1A_n}$	n
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{O}_{A_iA_{i+1}}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$	$n-1$
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_{j-1}A_j}, \ldots, \mathbf{I}_{A_iA_{i+1}}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$	$\frac{(n-1)(n-2)}{2}$
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_{j-1}A_j}, \ldots, \mathbf{I}_{A_{i+1}A_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$	$\frac{(n-1)(n-2)}{2}$
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_jA_{j-1}}, \ldots, \mathbf{I}_{A_iA_{i+1}}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$	$\frac{(n-1)(n-2)}{2}$
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_jA_{j-1}}, \ldots, \mathbf{I}_{A_{i+1}A_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$	$\frac{(n-1)(n-2)}{2}$
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_{j-1}A_j}, \ldots, \mathbf{I}_{A_iA_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$	$\frac{n(n-1)}{2}$
$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_jA_{j-1}}, \ldots, \mathbf{I}_{A_iA_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$	$\frac{n(n-1)}{2}$

validating the n-term syllogism

$$\mathbf{A}_{A_{n-1}A_n}, \ldots, \mathbf{A}_{A_1A_2} \models \mathbf{A}_{A_1A_n}$$

– By Lemma 4.1 (ii), the $n-1$ diagrams



represent the only way to produce evidence for the syllogistic inference

$$(\mathbf{A}_{A_nA_{n-1}})\sharp\cdots\sharp(\mathbf{E}_{A_iA_{i+1}})\sharp\cdots\sharp(\mathbf{A}_{A_1A_2}) \models (\mathbf{E}_{A_1A_n})$$

as well as for the syllogistic inference

$$(\mathbf{A}_{A_nA_{n-1}})\sharp\cdots\sharp(\mathbf{E}_{A_{i+1}A_i})\sharp\cdots\sharp(\mathbf{A}_{A_1A_2}) \models (\mathbf{E}_{A_1A_n})$$

validating the $n-1$ syllogisms

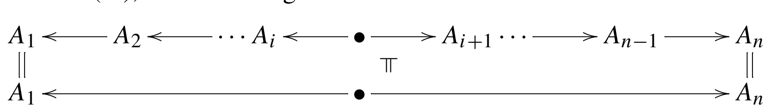
$$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_iA_{i+1}}, \ldots, \mathbf{A}_{A_1A_2} \models \mathbf{E}_{A_1A_n}$$

and the $n-1$ syllogisms

$$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_{i+1}A_i}, \ldots, \mathbf{A}_{A_1A_2} \models \mathbf{E}_{A_1A_n}$$

respectively. Thus, in total there are $2(n-1)$ valid n-term syllogisms with conclusion $\mathbf{E}_{A_1A_n}$.

– By Lemma 4.1 (iii), the $n-1$ diagrams



represent the only way to produce evidence for the $n-1$ syllogistic inferences

$$(\mathbf{A}_{A_{n-1}A_n})\sharp\cdots\sharp(\mathbf{I}_{A_iA_{i+1}})\sharp\cdots\sharp(\mathbf{A}_{A_2A_1}) \models (\mathbf{I}_{A_1A_n})$$

as well as for the $n-1$ syllogistic inferences

$$(\mathsf{A}_{A_{n-1}A_n})\sharp\cdots\sharp(\mathsf{I}_{A_{i+1}A_i})\sharp\cdots\sharp(\mathsf{A}_{A_2A_1}) \models (\mathsf{I}_{A_1A_n})$$

that validate the $n-1$ n-term syllogisms

$$\mathsf{A}_{A_{n-1}A_n},\ldots,\mathsf{I}_{A_iA_{i+1}},\ldots,\mathsf{A}_{A_2A_1} \models \mathsf{I}_{A_1A_n}$$

and the $n-1$ n-term syllogisms

$$\mathsf{A}_{A_{n-1}A_n},\ldots,\mathsf{I}_{A_{i+1}A_i},\ldots,\mathsf{A}_{A_2A_1} \models \mathsf{I}_{A_1A_n}$$

respectively.

By Lemma 4.1 (iv), the n diagrams

$$\begin{array}{lcr} A_1 \longleftarrow A_2 \longleftarrow \cdots A_i \longleftarrow & \bullet & \longrightarrow A_i \cdots \longrightarrow A_{n-1} \longrightarrow A_n \\ \| & \top\!\!\top & \| \\ A_1 \longleftarrow & \bullet & \longrightarrow A_n \end{array}$$

represent the only way to produce evidence for the n syllogistic inferences

$$(\mathsf{A}_{A_{n-1}A_n})\sharp\cdots\sharp(\mathsf{I}_{A_iA_i})\sharp\cdots\sharp(\mathsf{A}_{A_2A_1}) \models (\mathsf{I}_{A_1A_n})$$

that validate the n n-term syllogisms

$$\mathsf{A}_{A_{n-1}A_n},\ldots,\mathsf{I}_{A_iA_i},\ldots,\mathsf{A}_{A_2A_1} \models \mathsf{I}_{A_1A_n}$$

so that in total there are $2(n-1)+n$ valid n-term syllogisms with conclusion $\mathsf{I}_{A_1A_n}$.

– By Lemma 4.1 (v), the $n-1$ diagrams

$$\begin{array}{lcr} A_1 \longleftarrow A_2 \longleftarrow \cdots A_i \longleftarrow & \bullet \longrightarrow \bullet & \longleftarrow A_{i+1} \cdots \longleftarrow A_{n-1} \longleftarrow A_n \\ \| & \top\!\!\top & \| \\ A_1 \longleftarrow & \bullet \longrightarrow \bullet & \longleftarrow A_n \end{array}$$

represent the only way to produce evidence for the $n-1$ syllogistic inferences

$$(\mathsf{A}_{A_nA_{n-1}})\sharp\cdots\sharp(\mathsf{O}_{A_iA_{i+1}})\sharp\cdots\sharp(\mathsf{A}_{A_2A_1}) \models (\mathsf{O}_{A_1A_n})$$

that validate the $n-1$ n-term syllogisms

$$\mathsf{A}_{A_nA_{n-1}},\ldots,\mathsf{O}_{A_iA_{i+1}},\ldots,\mathsf{A}_{A_2A_1} \models \mathsf{O}_{A_1A_n}.$$

By Lemma 4.1 (vi) and Lemma 4.2 (i), the $\frac{(n-1)(n-2)}{2}$ diagrams

$$\begin{array}{lcr} A_1 \longleftarrow \cdots A_i \longleftarrow \bullet & \longrightarrow A_{i+1} \cdots \longrightarrow \cdots A_{j-1} \longrightarrow & \bullet \longleftarrow A_j \cdots \longleftarrow A_n \\ \| & \top\!\!\top & \| \\ A_1 \longleftarrow \bullet & \longrightarrow & \bullet \longleftarrow A_n \end{array}$$

represent the only way to produce evidence for the $4\cdot\frac{(n-1)(n-2)}{2}$ syllogistic inferences

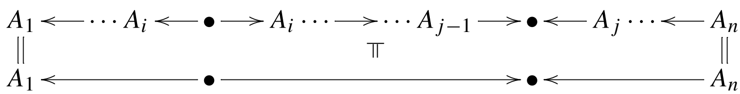
$$\begin{array}{l} (\mathsf{A}_{A_nA_{n-1}})\sharp\cdots\sharp(\mathsf{E}_{A_{j-1}A_j})\sharp\cdots\sharp(\mathsf{I}_{A_iA_{i+1}})\sharp\cdots\sharp(\mathsf{A}_{A_2A_1}) \models (\mathsf{O}_{A_1A_n}) \\ (\mathsf{A}_{A_nA_{n-1}})\sharp\cdots\sharp(\mathsf{E}_{A_{j-1}A_j})\sharp\cdots(\mathsf{I}_{A_{i+1}A_i})\sharp\cdots\sharp(\mathsf{A}_{A_2A_1}) \models (\mathsf{O}_{A_1A_n}) \\ (\mathsf{A}_{A_nA_{n-1}})\sharp\cdots\sharp(\mathsf{E}_{A_jA_{j-1}})\sharp\cdots\sharp(\mathsf{I}_{A_iA_{i+1}})\sharp\cdots\sharp(\mathsf{A}_{A_2A_1}) \models (\mathsf{O}_{A_1A_n}) \\ (\mathsf{A}_{A_nA_{n-1}})\sharp\cdots\sharp(\mathsf{E}_{A_jA_{j-1}})\sharp\cdots\sharp(\mathsf{I}_{A_{i+1}A_i})\sharp\cdots\sharp(\mathsf{A}_{A_2A_1}) \models (\mathsf{O}_{A_1A_n}) \end{array}$$

that validate the $4\cdot\frac{(n-1)(n-2)}{2}$ n-term syllogisms

$$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_{j-1}A_j}, \ldots, \mathbf{I}_{A_iA_{i+1}}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$$
$$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_{j-1}A_j}, \ldots, \mathbf{I}_{A_{i+1}A_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$$
$$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_jA_{j-1}}, \ldots, \mathbf{I}_{A_iA_{i+1}}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$$
$$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_jA_{j-1}}, \ldots, \mathbf{I}_{A_{i+1}A_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$$

respectively.

By Lemma 4.1 (vii) and Lemma 4.2 (ii), the $\frac{n(n-1)}{2}$ diagrams



represent the only way to produce evidence for the $2 \cdot \frac{n(n-1)}{2}$ syllogistic inferences

$$(\mathbf{A}_{A_nA_{n-1}})\sharp \cdots \sharp(\mathbf{E}_{A_{j-1}A_j})\sharp \cdots \sharp(\mathbf{I}_{A_iA_i})\sharp \cdots \sharp(\mathbf{A}_{A_2A_1}) \models (\mathbf{O}_{A_1A_n})$$
$$(\mathbf{A}_{A_nA_{n-1}})\sharp \cdots \sharp(\mathbf{E}_{A_jA_{j-1}})\sharp \cdots \sharp(\mathbf{I}_{A_iA_i})\sharp \cdots \sharp(\mathbf{A}_{A_2A_1}) \models (\mathbf{O}_{A_1A_n})$$

that validate the $2 \cdot \frac{n(n-1)}{2}$ n-term syllogisms

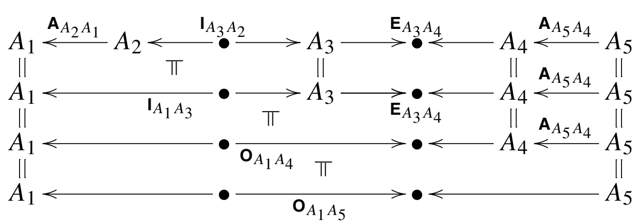
$$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_{j-1}A_j}, \ldots, \mathbf{I}_{A_iA_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$$
$$\mathbf{A}_{A_nA_{n-1}}, \ldots, \mathbf{E}_{A_jA_{j-1}}, \ldots, \mathbf{I}_{A_iA_i}, \ldots, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_n}$$

respectively. Thus in total there are $n - 1 + 4 \cdot \frac{(n-1)(n-2)}{2} + 2 \cdot \frac{n(n-1)}{2}$ valid n-term syllogisms with conclusion $\mathbf{O}_{A_1A_n}$.

In total, the valid n-term syllogisms are in number of

$$1 + 2(n-1) + 2(n-1) + n + (n-1) + 4 \cdot \frac{(n-1)(n-2)}{2} + 2 \cdot \frac{n(n-1)}{2} = 3n^2 - n \qquad \square$$

The existing syllogistic inference providing evidence for the validity of an n-term syllogism is not uniquely determined, in general. In order to see this consider the valid 5-term syllogism $\mathbf{A}_{A_5A_4}, \mathbf{E}_{A_3A_4}, \mathbf{I}_{A_3A_2}, \mathbf{A}_{A_2A_1} \models \mathbf{O}_{A_1A_5}$ for example. It can be obtained as the composite of subsequent syllogistic inferences in two different ways, as shown by the diagrams



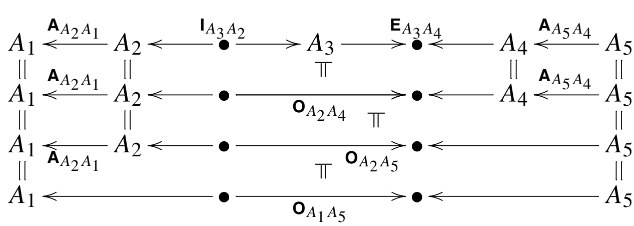


Table 5 Figures of the 2-term syllogisms

	Fig. 1	Fig. 2
premise	A_1A_2	A_1A_2
conclusion	A_1A_2	A_2A_1

which is also the reason why, differently from Theorem 3.4, in Theorem 3.7 the uniqueness condition on the existence of a syllogistic inference for the validity of a syllogism with existential import has been dropped.

We end with the explicit description of the valid n-term syllogisms for $n = 1$ and $n = 2$, respectively. In the first case, there is only one figure, that is A_1A_1 and only two valid moods for it, that is A and I so that, as observed in [8] and [9], the only valid 1-term syllogisms are $\mathsf{A}_{A_1A_1} \models \mathsf{A}_{A_1A_1}$ and $\mathsf{I}_{A_1A_1} \models \mathsf{I}_{A_1A_1}$, that is the laws of identity we hinted at in the previous section. In the second case there are two figures, as shown in Table 5, and ten valid 2-term syllogisms, six in the first figure and four in the second, as follows:

Figure 1: $\mathsf{A}_{A_1A_2} \models \mathsf{A}_{A_1A_2}$, $\mathsf{E}_{A_1A_2} \models \mathsf{E}_{A_1A_2}$, $\mathsf{I}_{A_1A_2} \models \mathsf{I}_{A_1A_2}$, $\mathsf{O}_{A_1A_2} \models \mathsf{O}_{A_1A_2}$, plus the *laws of subalternation* $\mathsf{A}_{A_1A_2}, \mathsf{I}_{A_1A_1} \models \mathsf{I}_{A_1A_2}$, $\mathsf{E}_{A_1A_2}, \mathsf{I}_{A_1A_1} \models \mathsf{O}_{A_1A_2}$.

Figure 2: $\mathsf{E}_{A_2A_1} \models \mathsf{E}_{A_1A_2}$, $\mathsf{I}_{A_2A_1} \models \mathsf{I}_{A_1A_2}$ which are the *laws of simple conversion*, and $\mathsf{I}_{A_2A_2}, \mathsf{A}_{A_2A_1} \models \mathsf{I}_{A_1A_2}$, $\mathsf{E}_{A_2A_1}, \mathsf{I}_{A_1A_1} \models \mathsf{O}_{A_1A_2}$ which are the *laws of conversion per accidents*.

5 Syllogisms with Complemented Terms

The investigation of the possibility of an algebraic calculus of syllogisms was in particular carried out by Augustus De Morgan in the nineteenth century. In [2], he introduced the so called *spicular notation* to manage the syllogistic reasoning in purely symbolic terms. Such a notation rises from a detailed analysis of the meaning of the categorical propositions. De Morgan argued that a symbolic formalism for their representation should simultaneously convey information about them being universal or particular, affirmative or negative, saving the possibility of pointing out these characteristics separately, or in his words:

> A *fundamental symbol* should not be of compound meaning: that is, should not expressly signify more than one thing.

Thus, De Morgan lets a term-variable X be enclosed by a parenthesis, as in $X)$ or $(X$, to express universal quantification, that is “all Xs”, whereas he lets a term-variable being excluded by a parenthesis, as in $)X$ or $X($, to mean particular quantification, namely “some Xs”. Furthermore, he lets an even number of dots, or none at all, between parentheses, express affirmation or agreement of terms, whereas he lets an odd number of dots express negation or disagreement of terms. The following are the fundamental categorical propositions as they appear in the spicular notation:

$$\mathsf{A}_{XY}\colon X\,)\,)\,Y \qquad \mathsf{E}_{XY}\colon X\,)\cdot(\,Y$$
$$\mathsf{I}_{XY}\colon X\,(\,)\,Y \qquad \mathsf{O}_{XY}\colon X\,(\cdot(\,Y$$

which accordingly should now be read

$$\mathbf{A}_{XY}: \text{ All } X\text{s are some } Y\text{s}$$

$$\mathbf{E}_{XY}: \text{ All } X\text{s are not all } Y\text{s}$$

$$\mathbf{I}_{XY}: \text{ Some } X\text{s are some } Y\text{s}$$

$$\mathbf{O}_{XY}: \text{ Some } X\text{s are not all } Y\text{s}$$

The possibility of doing formal inferences through the employment of the spicular notation is based on the *erasure rule*, that is the deletion of the middle term together with the pair of symbols around it. For example, the verification of the validity of the syllogism $\mathbf{A}_{PM}, \mathbf{O}_{SM} \models \mathbf{O}_{SP}$ appears as $S\,(\,\cdot\,(M\,(\,(P \models S\,(\,\cdot\,(P$, whereas the verification of the validity of the syllogism $\mathbf{A}_{MP}, \mathbf{I}_{SM} \models \mathbf{I}_{SP}$ appears as $S\,(\,)\,M\,)\,)\,P \models S\,(\,)\,P$.

With the spicular notation there is another rule involved, that is the *transformation rule*, strictly related to the introduction of *complements* of terms. The complement of a term X is the term that means non-X, which De Morgan was used to denote x. Of course X is the complement of x and the transformation rule for the spicular notation allows to substitute a term with its complement, by rewriting it with the corresponding lower/upper case letter, simultaneously changing the curvature of its parenthesis and adding or removing a negation dot. For example, with the possibility of mentioning complements of terms, the alternative reading of $X\,)\,)\,Y$ could be "All Xs are not all non-Ys", which can be symbolically obtained by applying the transformation rule just described, so to retrieve $X\,)\,\cdot\,(\,y$.

The possibility of making a distinction between a term being universally or particularly quantified, as well as between affirmative and negative modes of predication is supported by the diagrammatic formalism we discussed in the previous sections, together with the possibility of handling complements of terms. Indeed, we can look at the symbols $\bullet$, $\rightarrow$ and $\leftarrow$ as to fundamental ones. In a diagram built as a combination of such fundamental symbols, a term-variable X is universally quantified if it enters in it as $X \rightarrow$ or $\leftarrow X$, whereas it is particularly quantified if it enters in it as $X \leftarrow$ or $\rightarrow X$. The complement of X is represented as $X \rightarrow \bullet$ or $\bullet \leftarrow X$, both of which may be abbreviated as x. In our opinion, the giving of an explicit encoding of complements of terms through the fundamental symbols $\bullet$, $\rightarrow$ and $\leftarrow$, is an advantage of the diagrammatic formalism with respect to the spicular notation. The term X is affirmed if it enters in a diagram as $\bullet \rightarrow X$ or $X \leftarrow \bullet$.

By the employment of complements, De Morgan was able to introduce four more categorical propositions and also to let the particular and universal affirmative modes of predication be the fundamental ones. Table 6 is essentially the one contained in [2]. It lists the forms of predication now available, with the corresponding spicular and diagrammatic representations. We are assuming that a pair of accordingly oriented arrows separated by a bullet never compose, that is $\rightarrow \bullet \rightarrow$ for example does not reduces to $\rightarrow$. The double negation of a term X is represented by a diagram like

$$\bullet \leftarrow X \rightarrow \bullet$$

and whenever it will be the case, in performing a syllogistic inference we will refer to *double negation substitution* as to the operation of substituting that diagram with the sole

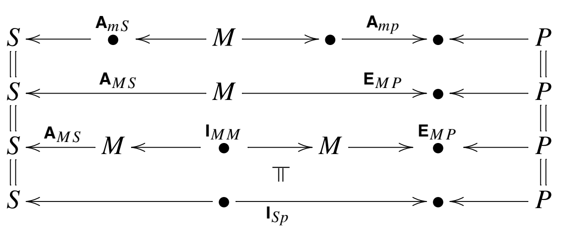
Table 6 De Morgan's forms of predication

	Spic. notation	Diagr. notation	Natural language
A_{XY}	$X\,)\,)\,Y$	$X \to Y$	All X is Y
A_{xy}	$x\,)\,)\,y$ or $X\,(\,(\,Y$	$X \to \bullet \to \bullet \leftarrow Y$	All Y is X
A_{Xy}	$X\,)\,)\,y$ or $X\,(\,\cdot\,(\,Y$	$X \to \bullet \leftarrow Y$	No X is Y
A_{xY}	$x\,)\,)\,Y$ or $X\,(\,\cdot\,)\,Y$	$X \to \bullet \to Y$	Everything is X or Y
I_{XY}	$X\,(\,)\,Y$	$X \leftarrow \bullet \to Y$	Some X is Y
I_{xy}	$x\,(\,)\,y$ or $X\,)\,(\,Y$	$X \to \bullet \leftarrow \bullet \to \bullet \leftarrow Y$	Some things are neither X nor Y
I_{Xy}	$X\,(\,)\,y$ or $X\,(\,\cdot\,(\,Y$	$X \leftarrow \bullet \to \bullet \leftarrow Y$	Some X is not Y
I_{xY}	$x\,(\,)\,Y$ or $X\,)\,\cdot\,)\,Y$	$X \to \bullet \leftarrow \bullet \to Y$	Some Y is not X

Table 7 The valid De Morgan's syllogisms

	1	2	3	4	5	6	7	8
S	*	*	*	*				
s					*	*	*	*
M	*	*			*	*		
m			*	*			*	*
P	*		*		*		*	
p		*		*		*		*

symbol X. For example, the syllogistic inference for the validity of the strengthened syllogism $\mathsf{A}_{mp}, \mathsf{A}_{mS} \models \mathsf{I}_{Sp}$ is



in which the first and second steps consist of a double negation substitution and of the insertion of the existential import I_{MM}, respectively, whereas the third one is the honest syllogistic inference.

From Table 6, De Morgan recovered thirty-two valid syllogisms with complements. Eight universal syllogisms whose scheme of inference is $)\,)\ \ )\,) =)\,)$. Sixteen syllogisms with one particular premise, precisely eight whose scheme of inference is $(\,(\ \ (\,) = (\,)$ and further eight whose scheme of inference is $(\,)\ \ )\,) = (\,)$. Finally, eight strengthened syllogisms whose scheme of inference is $(\,(\ \,)\,) = (\,)$. We listed them as indicated in Table 7 which we took from [6], with personal notation, and where each column corresponds to a syllogism in one of the previously described schemes of inference.

Theorem 5.1 *Each syllogism with complements in Table 7 is validated by a syllogistic inference and vice versa.*

Proof On one hand, we proceed by fixing one of the four schemes of inference previously described and going through the columns of Table 7, indicated by the numbers in the items below, showing that a syllogistic inference providing evidence for the validity of the corresponding syllogism with complements exists. The syllogistic inferences containing a double negation substitution are exactly those for a syllogism in which the middle term is complemented in the first as well as in the second premise. Existential import for the strengthened syllogisms is explicitly indicated.

(i) For the scheme $)\,)\;\;)\,) =)\,)$
- (i.1) $\mathbf{A}_{MP}, \mathbf{A}_{SM} \models \mathbf{A}_{SP}$ yields $(\mathbf{A}_{MP})\sharp(\mathbf{A}_{SM}) \models (\mathbf{A}_{SP})$
- (i.2) $\mathbf{A}_{Mp}, \mathbf{A}_{SM} \models \mathbf{A}_{Sp}$ yields $(\mathbf{A}_{Mp})\sharp(\mathbf{A}_{SM}) \models (\mathbf{A}_{Sp})$
- (i.3) $\mathbf{A}_{mP}, \mathbf{A}_{Sm} \models \mathbf{A}_{SP}$ yields $(\mathbf{A}_{mP})\sharp(\mathbf{A}_{Sm}) \models (\mathbf{A}_{SP})$
- (i.4) $\mathbf{A}_{mp}, \mathbf{A}_{Sm} \models \mathbf{A}_{Sp}$ yields $(\mathbf{A}_{mp})\sharp(\mathbf{A}_{Sm}) \models (\mathbf{A}_{Sp})$
- (i.5) $\mathbf{A}_{MP}, \mathbf{A}_{sM} \models \mathbf{A}_{sP}$ yields $(\mathbf{A}_{MP})\sharp(\mathbf{A}_{sM}) \models (\mathbf{A}_{sP})$
- (i.6) $\mathbf{A}_{Mp}, \mathbf{A}_{sM} \models \mathbf{A}_{sp}$ yields $(\mathbf{A}_{Mp})\sharp(\mathbf{A}_{sM}) \models (\mathbf{A}_{sp})$
- (i.7) $\mathbf{A}_{mP}, \mathbf{A}_{sm} \models \mathbf{A}_{sP}$ yields $(\mathbf{A}_{mP})\sharp(\mathbf{A}_{sm}) \models (\mathbf{A}_{sP})$
- (i.8) $\mathbf{A}_{mp}, \mathbf{A}_{sm} \models \mathbf{A}_{sp}$ yields $(\mathbf{A}_{mp})\sharp(\mathbf{A}_{sm}) \models (\mathbf{A}_{sp})$

(ii) For the scheme $(\,(\;\;(\,) = (\,)$
- (ii.1) $\mathbf{I}_{MP}, \mathbf{A}_{MS} \models \mathbf{I}_{SP}$ yields $(\mathbf{I}_{MP})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{SP})$
- (ii.2) $\mathbf{I}_{Mp}, \mathbf{A}_{MS} \models \mathbf{I}_{Sp}$ yields $(\mathbf{I}_{Mp})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{Sp})$
- (ii.3) $\mathbf{I}_{mP}, \mathbf{A}_{mS} \models \mathbf{I}_{SP}$ yields $(\mathbf{I}_{mP})\sharp(\mathbf{A}_{mS}) \models (\mathbf{I}_{SP})$
- (ii.4) $\mathbf{I}_{mp}, \mathbf{A}_{mS} \models \mathbf{I}_{Sp}$ yields $(\mathbf{I}_{mp})\sharp(\mathbf{A}_{mS}) \models (\mathbf{I}_{Sp})$
- (ii.5) $\mathbf{I}_{MP}, \mathbf{A}_{Ms} \models \mathbf{I}_{sP}$ yields $(\mathbf{I}_{MP})\sharp(\mathbf{A}_{Ms}) \models (\mathbf{I}_{sP})$
- (ii.6) $\mathbf{I}_{Mp}, \mathbf{A}_{Ms} \models \mathbf{I}_{sp}$ yields $(\mathbf{I}_{Mp})\sharp(\mathbf{A}_{Ms}) \models (\mathbf{I}_{sp})$
- (ii.7) $\mathbf{I}_{mP}, \mathbf{A}_{ms} \models \mathbf{I}_{sP}$ yields $(\mathbf{I}_{mP})\sharp(\mathbf{A}_{ms}) \models (\mathbf{I}_{sP})$
- (ii.8) $\mathbf{I}_{mp}, \mathbf{A}_{ms} \models \mathbf{I}_{sp}$ yields $(\mathbf{I}_{mp})\sharp(\mathbf{A}_{ms}) \models (\mathbf{I}_{sp})$

(iii) For the scheme $(\,)\;\;)\,) = (\,)$
- (iii.1) $\mathbf{A}_{MP}, \mathbf{I}_{SM} \models \mathbf{I}_{SP}$ yields $(\mathbf{A}_{MP})\sharp(\mathbf{I}_{SM}) \models (\mathbf{I}_{SP})$
- (iii.2) $\mathbf{A}_{Mp}, \mathbf{I}_{SM} \models \mathbf{I}_{Sp}$ yields $(\mathbf{A}_{Mp})\sharp(\mathbf{I}_{SM}) \models (\mathbf{I}_{Sp})$
- (iii.3) $\mathbf{A}_{mP}, \mathbf{I}_{Sm} \models \mathbf{I}_{SP}$ yields $(\mathbf{A}_{mP})\sharp(\mathbf{I}_{Sm}) \models (\mathbf{I}_{SP})$
- (iii.4) $\mathbf{A}_{mp}, \mathbf{I}_{Sm} \models \mathbf{I}_{Sp}$ yields $(\mathbf{A}_{mp})\sharp(\mathbf{I}_{Sm}) \models (\mathbf{I}_{Sp})$
- (iii.5) $\mathbf{A}_{MP}, \mathbf{I}_{sM} \models \mathbf{I}_{sP}$ yields $(\mathbf{A}_{MP})\sharp(\mathbf{I}_{sM}) \models (\mathbf{I}_{sP})$
- (iii.6) $\mathbf{A}_{Mp}, \mathbf{I}_{sM} \models \mathbf{I}_{sp}$ yields $(\mathbf{A}_{Mp})\sharp(\mathbf{I}_{sM}) \models (\mathbf{I}_{sp})$
- (iii.7) $\mathbf{A}_{mP}, \mathbf{I}_{sm} \models \mathbf{I}_{sP}$ yields $(\mathbf{A}_{mP})\sharp(\mathbf{I}_{sm}) \models (\mathbf{I}_{sP})$
- (iii.8) $\mathbf{A}_{mp}, \mathbf{I}_{sm} \models \mathbf{I}_{sp}$ yields $(\mathbf{A}_{mp})\sharp(\mathbf{I}_{sm}) \models (\mathbf{I}_{sp})$

(iv) For the scheme $(\,(\;\;)\,) = (\,)$
- (iv.1) $\mathbf{A}_{MP}, \mathbf{A}_{MS} \models \mathbf{I}_{SP}$ yields $(\mathbf{A}_{MP})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{SP})$
- (iv.2) $\mathbf{A}_{Mp}, \mathbf{A}_{MS} \models \mathbf{I}_{Sp}$ yields $(\mathbf{A}_{Mp})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{Sp})$
- (iv.3) $\mathbf{A}_{mP}, \mathbf{A}_{mS} \models \mathbf{I}_{SP}$ yields $(\mathbf{A}_{mP})\sharp(\mathbf{A}_{mS}) \models (\mathbf{A}_{MP})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{SP})$
- (iv.4) $\mathbf{A}_{mp}, \mathbf{A}_{mS} \models \mathbf{I}_{Sp}$ yields $(\mathbf{A}_{mp})\sharp(\mathbf{A}_{mS}) \models (\mathbf{A}_{MP})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{MS}) \models (\mathbf{I}_{Sp})$
- (iv.5) $\mathbf{A}_{MP}, \mathbf{A}_{Ms} \models \mathbf{I}_{sP}$ yields $(\mathbf{A}_{MP})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{Ms}) \models (\mathbf{I}_{sP})$
- (iv.6) $\mathbf{A}_{Mp}, \mathbf{A}_{Ms} \models \mathbf{I}_{sp}$ yields $(\mathbf{A}_{Mp})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{Ms}) \models (\mathbf{I}_{sp})$
- (iv.7) $\mathbf{A}_{mP}, \mathbf{A}_{ms} \models \mathbf{I}_{sP}$ yields $(\mathbf{A}_{mP})\sharp(\mathbf{A}_{ms}) \models (\mathbf{A}_{MP})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{Ms}) \models (\mathbf{I}_{sP})$
- (iv.8) $\mathbf{A}_{mp}, \mathbf{A}_{ms} \models \mathbf{I}_{sp}$ yields $(\mathbf{A}_{mp})\sharp(\mathbf{A}_{ms}) \models (\mathbf{A}_{Mp})\sharp(\mathbf{I}_{MM})\sharp(\mathbf{A}_{Ms}) \models (\mathbf{I}_{sp})$

Conversely, we let the reader convince herself that it suffices to proceed as already done in Theorems 3.4 and 3.7, that is by constructing suitable syllogistic inferences to a fixed possible conclusion by taking into account both double negation substitutions and existential imports in the premises. It does not suffices to exclusively do this for $\mathbf{A}_{sp}$, $\mathbf{A}_{sP}$, $\mathbf{I}_{sp}$

and I_{sP}, since for example A_{SP} can be now obtained as a possible conclusion through the syllogistic inference $(\mathsf{A}_{MP})\sharp(\mathsf{A}_{SM}) \models (\mathsf{A}_{SP})$ as well as through the syllogistic inference $(\mathsf{A}_{mP})\sharp(\mathsf{A}_{Sm}) \models (\mathsf{A}_{SP})$, in which a double negation substitution occurs. □

References

1. Corcoran, J.: Completeness of an ancient logic. J. Symb. Log. **37**, 696–702 (1972)
2. De Morgan, A.: On the symbols of logic, the theory of the syllogism, and in particular of the copula, and the application of the theory of probabilities, to some questions of evidence. Trans. Camb. Philos. Soc. **9**, 79–127 (1851)
3. Englebretsen, G.: Linear diagrams for syllogisms (with relationals). Notre Dame J. Form. Log. **33**(1), 37–69 (1992)
4. Głazowska, K.: The structure of valid *n*-term syllogisms. Stud. Log. **8**, 249–257 (1958)
5. Hammer, E.: Reasoning with sentences and diagrams. Notre Dame J. Form. Log. **35**(1), 73–87 (1994)
6. Hobart, M.E., Richards, J.L.: De Morgan's logic. In: Handbook of the History of Logic, vol. 4, British Logic in the Nineteenth Century, pp. 283–329. Elsevier, Amsterdam (2008)
7. Lemon, O., Pratt, I.: On the insufficiency of linear diagrams for syllogisms. Notre Dame J. Form. Log. **39**(4), 573–580 (1998)
8. Łukasiewicz, J.: Aristotle's Syllogistic from the Standpoint of Modern Formal Logic. Clarendon, Oxford (1951)
9. Meredith, C.A.: The figures and moods of the *n*-term Aristotelian syllogism. Domin. Studies **6**, 42–47 (1953)
10. Shin, S.-J.: Situation-theoretic account of valid reasoning with Venn diagrams. In: Logical Reasoning with Diagrams. Stud. Logic Comput., vol. 6, pp. 81–108. Oxford University Press, New York (1996)
11. Smiley, T.J.: What is a syllogism? J. Philos. Log. **2**, 136–154 (1973)
12. Smyth, M.B.: A diagrammatic treatment of syllogistic. Notre Dame J. Form. Log. **12**(4), 483–488 (1971)

R. Pagnan (✉)
DISI, University of Genoa, Genoa, Italy
e-mail: ruggero.pagnan@disi.unige.it

Beyond Syllogisms: Carroll's (Marked) Quadriliteral Diagram

Amirouche Moktefi

Abstract The logician Lewis Carroll (1832–1898) invented a diagrammatic scheme for syllogisms and described how it could be used for logic problems involving more than 3 terms. Curiously, he never provided in print any diagrammatic solution for such a complex problem. The aim of this paper is to make sense of a manuscript where Carroll attempts to solve a sorite using his quadriliteral diagram. In this problem, three propositions are offered as premises. The purpose is to look for what information can be gathered as to the relation between two given terms involved in the argument. This case study provides some insights about the use of diagrams to solve elimination problems that were highly considered by early symbolist logicians.

Keywords Syllogism · Sorite · Logic diagram · Logic problem · Visual reasoning · Elimination · Premise · Conclusion · Lewis Carroll · Symbolic logic · Carroll diagram · Venn diagram

Mathematics Subject Classification 00A66 · 01A55 · 97E30

1 Introduction

It is well known that the logician Lewis Carroll invented in the 1880s an original diagrammatic scheme to solve syllogistic problems [6]. It is also known that he later described a series of logic diagrams that could be used to solve logic problems involving more than 3 terms [12, pp. 176–179]. Carroll published plenty of such complex problems known as sorites. Curiously however, he never printed any diagrammatic solution of them. A simple explanation would be that he might have intended to do so in the second part of his *Symbolic Logic* which he never completed. Unfortunately the surviving fragments collected and published by William W. Bartley III in 1977 do not contain any diagrammatic solution to logic problems for more than 3 terms, while they contain plenty of examples solved symbolically or using the method of trees [4].

However, a recent collection of Carroll's logic pamphlets edited by Francine Abeles, reproduced a manuscript where a complex logic problem involving four terms is solved by Carroll with the help of his diagrammatic method [3, p. 59]. That manuscript is reproduced as Fig. 1. It is part of 4 sheets of logic notes that are preserved in the *Houghton Collection* (Pierpont Morgan Library, New York), one of which contains a text dated on 13 September 1892 [10]. This manuscript shows a logic problem at its top-right side, a 4-term logic diagram at its top-left side, and several small diagrams and formulae at its

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic, DOI 10.1007/978-3-0348-0600-8_4, © Springer Basel 2013

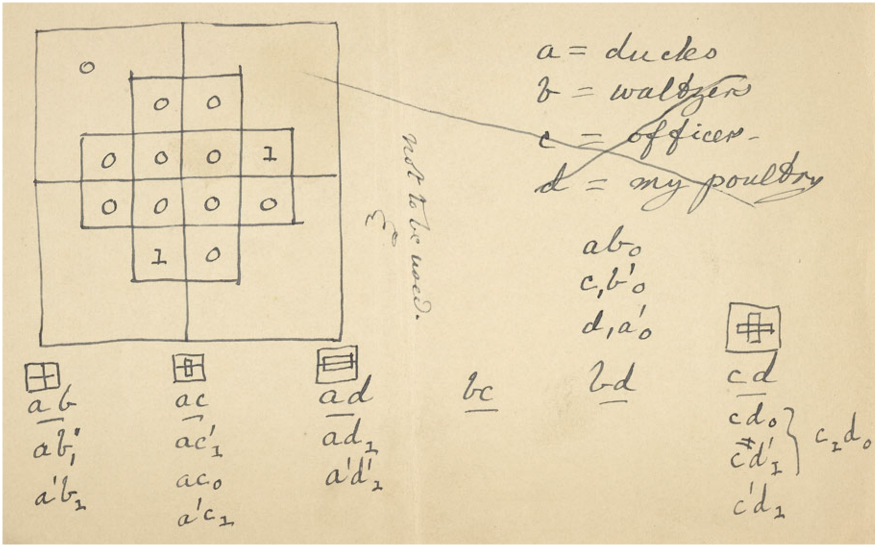


Fig. 1 Quadriliteral diagram by Charles L. Dodgson. Reprinted with kind permission from © The Pierpont Morgan Library, New York. AAH 545, 2013

bottom. Carroll handles here the problem in an unusual way, even for those familiar with his method of diagrams. The aim of this paper is to make sense of this manuscript and to explain how Carroll solved this 4-term problem with his diagrams. As such, this case study might provide some lights on the use of diagrams to solve sorites.

2 Carroll's Visual Logic

2.1 Biliteral and Triliteral Diagrams

Carroll's diagrams, invented in 1884 and first published in 1886, are Venn-type diagrams where the universe is represented with a square. However, it is not clear whether Carroll worked his diagrams independently or as a modification of John Venn's. Still, Carroll's scheme looks like a "mature" method summing up several improvements that have been introduced by his predecessors and contemporaries [2, 14]. For 2 terms x and y, Carroll divides the square into 4 compartments, and obtains the so-called biliteral diagram, as shown in Fig. 2 (where x' stands for *not*-x and y' for *not*-y). For 3 terms x, y and m, Carroll adds a smaller square in order to get 8 compartments and obtain the triliteral diagram, as shown in Fig. 3 (where m' stands for *not*-m).

In order to represent propositions, one has to add marks. A compartment is marked with a '0' if it is empty and is marked with a 'I' if it is occupied. For instance, in order to represent the proposition "All x are y", one has to put a '0' on the x *not*-y compartment and a 'I' on the xy compartment, in accordance with Carroll's interpretation of A-propositions, as shown in Fig. 4. Finally, suppose that one wants to represent the proposition "Some x are m" on a triliteral diagram. This means that either xym or $xy'm$ is occupied ("or" is understood here inclusively). To represent this uncertainty, Carroll

Fig. 2

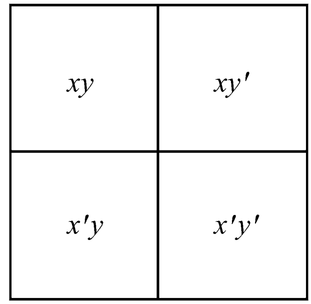


Fig. 3

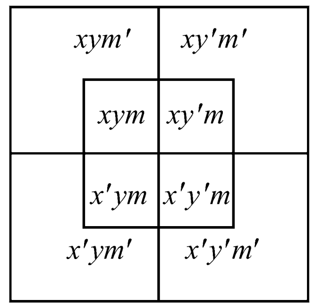


Fig. 4

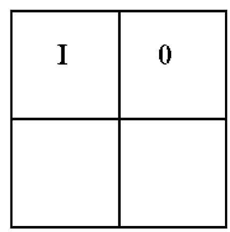
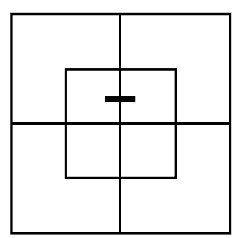


Fig. 5



puts the symbol 'I' (for occupation) on the boundary between those two compartments, as shown in Fig. 5 [12, p. 26].

In order to find the conclusion of a syllogism, Carroll first represents the data expressed by the two premises on a triliteral diagram. Let the premises be: "No m is y" and "Some x are m". Their representation is shown in Fig. 6. Contrary to Venn who extracts the conclusion directly from his 3-term diagram, Carroll transfers the data shown by the triliteral diagram into a biliteral diagram, involving only the 2 terms that should appear in the conclusion (here x and y) and consequently eliminating the middle term (here m).

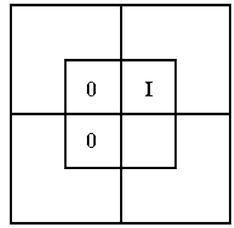


Fig. 6

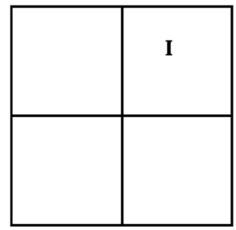


Fig. 7

This transfer is made thanks to two rules that Carroll applies on the 4 quarters of the triliteral and biliteral diagrams [12, p. 53]:

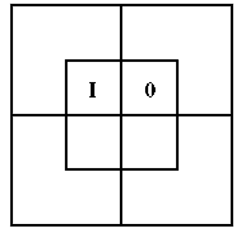
– Rule A: If the quarter of the triliteral diagram contains a 'I' in *either* Cell, then it is certainly occupied, and one may mark the corresponding quarter of the biliteral diagram with a 'I' to indicate that it is occupied.
– Rule B: If the quarter of the triliteral diagram contains two '0's, one in *each* cell, then it is certainly empty, and one may mark the corresponding quarter of the biliteral diagram with a '0' to indicate that it is empty.

The application of these rules here gives Fig. 7, which holds the conclusion: "Some x are *not*-y", that one draws from the given pair of premises. The importance of Carroll's method of transfer, unknown to Venn, should not be underestimated. It alone shows how to extract the conclusion from the premises of a syllogism [16, pp. 641–644].

2.2 The Elimination Problem

In the example above, we departed from traditional syllogistic problems where both premises and their conclusion are given and one is asked to check the validity of the trio. In our example, we were rather given a pair of premises and were then asked what conclusion(s) is/are to be drawn. This new approach reflects how Carroll and most symbolist logicians of his time handled the problem of syllogisms. This move has already been made by George Boole who worked on a general method for finding the conclusion that is to be drawn from any number of propositions given as premises containing any number of terms [5, p. 8]. For this purpose, one has to eliminate undesired terms in order to get the relation between the terms one wants to keep in the conclusion. This problem, known as the elimination problem, occupied the mind of nineteenth-century logicians who developed symbolic, visual and sometimes mechanical devices to solve it. Carroll was no exception, and his work should be understood within this historical

Fig. 8



context. Naturally, when one is offered a logic problem involving 3 terms, with two 2-term propositions given as premises, the elimination problem one faces is simply a traditional syllogism. In that case, the premises express the relation of the major and minor terms with the middle term, so that the only missing information is the relation between the major and minor terms. This means that one has to necessarily eliminate the middle term.

Now suppose that we were given two 3-term propositions such as "Some xy are m" and "No xm is *not*-y" and were not told what term was to be eliminated. The representation of those premises on the triliteral diagram is shown in Fig. 8. If we were asked what conclusion (in the form of a 2-term proposition) is to be drawn, it all depends on what terms we are interested in.

There are three possible cases, depending on whether we eliminate m, y or x:

– Eliminating m requires transferring information to the standard biliteral diagram we are used to, as shown in Fig. 9 which gives the conclusion "Some x are y".
– Eliminating y means that we have to transfer information to a new 2-term diagram, which is obtained by removing, from the triliteral diagram, the vertical line that divides the universe into subdivisions y and *not*-y. We thus obtain Fig. 10 which tells that "Some x are m".
– Finally, we proceed similarly to eliminate x, by the transfer of information to Fig. 11 which gives the conclusion: "Some y are m".

Reading conclusions on these new diagrams might require some practice. A solution to make it easy is to transfer information again from these new figures into a standard biliteral diagram, as we will see later. The point is that, contrary to the example in the previous section, we have eliminated here one term in each case in order to check what the relation between the two remaining terms is. Of course, it would also have been possible to look for relations between compound terms. For instance, one might also conclude from Fig. 8 that "All xm are y". However, for the purpose of this paper, we will limit ourselves to one-to-one relations.

In his published works, Carroll used merely diagrams of the form shown in Fig. 9. The reason is that when one applies Carroll's diagrammatic technique to traditional syllogisms, one already knows what term is the middle term. A Carrollian manuscript in Princeton University Library shows however two diagrams similar to Fig. 10 and Fig. 11, on the reverse of a circular dated on 1890, though the date of the diagrams themselves is unknown [9]. The fact that Carroll numbered those two diagrams as cases (2) and (3) shows that he used them the same way we did in the example above. There, the missing case (1) must have been Carroll's standard biliteral diagram.

Fig. 9

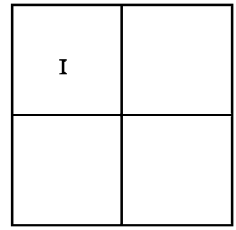


Fig. 10

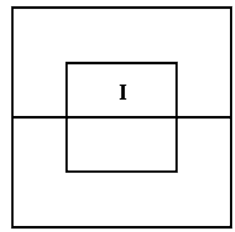
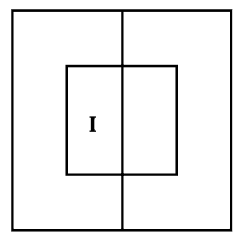


Fig. 11



2.3 The Quadriliteral Diagram

As far, we discussed solving diagrammatically problems involving 3 terms. However, nothing prevents from handling more complex problems. Contrary to old syllogistic where such diagrams were hardly desired, the new Boolean logic made logicians care about designing diagrammatic schemes for more than 3 terms. One particular difficulty was to keep the curves continuous and still make the diagrams provide the visual aid that one would expect from such devices [15]. Venn, whose three-circle diagram fits perfectly for syllogisms, abandoned circles in favor of ellipses in order to represent 4 terms, as shown in Fig. 12. In this diagram, all classes are continuous and easy to locate [18, p. 116]. For instance, the star indicates the compartment *not-x* y z w. For more than 4 terms, Venn unhappily made use of a non-continuous class, thus privileging regularity at that stage. Allan Marquand introduced new rectilinear diagrams where he made no attempt to keep his figures continuous at all [13]. Even for just 4 terms, his diagram represents classes C, not-C, D, not-D with disconnected regions as shown in Fig. 13, where a stand for not-A, b for not-B, etc. Carroll, though he used tabular diagrams like Marquand, shared Venn's concern about the continuity of the figures up to 4 terms, and similarly failed to provide a satisfactory diagram for 5 terms [12, p. 177].

For 4 terms, Carroll simply changed the small square that was inside the triliteral diagram into a rectangle, then added another rectangle that intersects with the first one in the

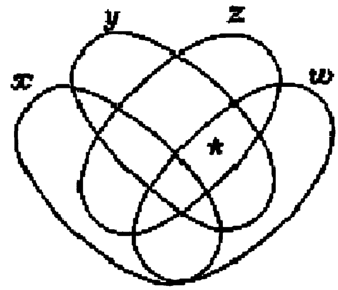


Fig. 12 Taken from Venn's book of 1894 [18]

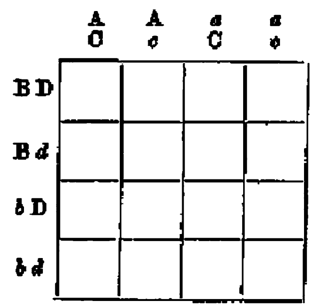


Fig. 13 Taken from Marquand (1881, [13])

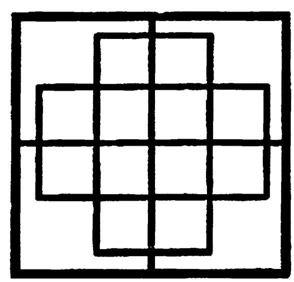


Fig. 14

desired manner, so that to obtain the quadriliteral diagram shown in Fig. 14. Here is how Carroll describes it in his *Symbolic Logic*:

> For four letters (which I call a, b, c, d) I use this diagram; assigning the North Half to a (and of course the rest of the diagram to a'), the West half to b, the Horizontal Oblong to c, and the Upright Oblong to d. We have now got 16 Cells. [12, p. 177]

Solving diagrammatically a logic problem involving more than 3 terms requires the same rules as for syllogisms: one represents first the information contained in the premises on the appropriate diagram (depending on the number of terms), then one has to transfer the information into a smaller diagram showing the specific relation (or relations) between the term (or terms) that might interest him, simply by eliminating the undesired terms. For instance, let us work on the following set of propositions offered as premises [12, p. 113]:

No c is d,
No *not-d* is a,
No *not-c* is b.

Fig. 15

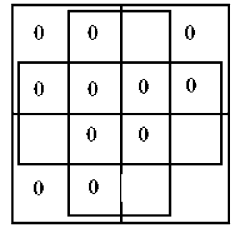
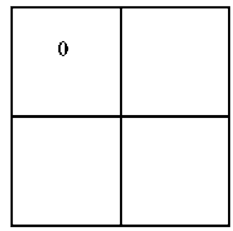


Fig. 16



Representing these premises on a quadriliteral diagram gives Fig. 15. Terms c and d appear twice (once affirmed and once denied), while a and b appear just once. So, one expects the former (c and d) to be eliminated, and the latter (a and b) to appear in the conclusion. Transferring information into a biliteral diagram, whose terms are a and b, gives Fig. 16 which provides the conclusion "No a is b". This is the same conclusion that Carroll arrived at symbolically [12, p. 158].

Now suppose we did not have any specific expectations as to what terms should appear in the conclusion, and were thus led to discuss all possible relations between any two terms. For instance, in the above example, we might want to investigate what the premises tell about the relation between terms a and c, or between b and d, none of which is stated in either premise considered alone. Of course, the data might say nothing at all as to the relation between those terms, but that itself would be new information that was not known until one has discussed all possible relations between terms one-to-one. For this purpose, we should proceed the same way we did with 3-term problems in Sect. 2.2. That's very precisely what Carroll's manuscript (Fig. 1) is about, as we shall see in the next sections.

3 Making Sense of the (Marked) Quadriliteral Diagram

3.1 Reading the Data

As far, we exposed Carroll's diagrammatic scheme and discussed how he made (and might have made) use of it to solve syllogisms and sorites. Let's now return to our manuscript and examine what information can be gathered there. The top-left side of the manuscript shows a marked quadriliteral diagram that does already represent the premises of the logic problem. The dictionary and the symbolic formulae on the top-right side of the manuscript make it possible to tell what the premises of the logic problem were, even if no concrete propositions are given. Finally, the bottom of the manuscript contains six columns of formulae. Four columns are headed by 2-term diagrams, among which some have already been discussed in Sect. 2.2. Let us in the following reconstruct the original problem that Carroll was working on, check the correctness of the diagrammatic representation he provided, and finally extract the conclusions that should be drawn from those premises.

The main diagram is a quadriliteral one, which means that we have a 4-term problem. These terms are listed as a, b, c and d on the top-right side of the manuscript. We are also told that a refers to "ducks", b to "waltzers", c to "officer", and d to "my poultry". Under the dictionary, three 2-term propositions appear in subscript forms:

1. ab_0
2. $c_1b'_0$
3. $d_1a'_0$

These propositions, given as premises, can easily be interpreted by any reader familiar with Carroll's logical notation. The introduction of symbolism in logic from the eighteenth century forwards made several notations compete [17]. Early notations were mostly equational as can be seen in the writings of Boole and his immediate followers William S. Jevons and Venn. Others such as Charles S. Peirce, Ernst Schröder and Hugh MacColl appealed rather to inclusional or implicational notations. Carroll explored a different path as he introduced subscripts to indicate the state of a class: "0" for emptiness and "1" for existence [1, 14].

For instance, the proposition "No x is y" tells that the class xy is empty. So, one simply represents it as: "xy_0". Similarly, the proposition "Some x are y" is represented as "xy_1". Propositions as "All x are y" are more complex because Carroll considered them to assert both the existence of x and the emptiness of xy' (where y' stand for *not*-y). Consequently, he represented them as "$x_1y'_0$", with subscripts taking effect back to the beginning of the formula [12, p. 72]. Thanks to these conventions, one can easily transform the trio of premises given in the manuscript into the following abstract forms:

1. No a is b
2. All c are b
3. All d are a

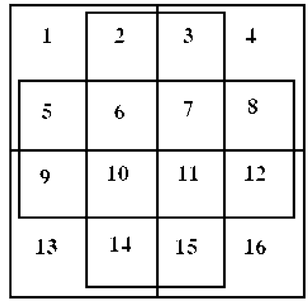
Again, if we replace letters a, b, c and d as indicated in the dictionary, we obtain the original problem as it must have been in concrete form:

1. No *duck* is *a waltzer*
2. All *officers* are *waltzers*
3. All *my poultry* are *ducks*

3.2 Representing Premises

Surprisingly, the representation of the premises we got on a quadriliteral diagram, as we described it in Sect. 2.3, would not lead to a figure similar to the one drawn by Carroll in his manuscript. The explanation is quite simple: in Sect. 2.3, we quoted the only passage where Carroll described his quadriliteral diagram. That was in the first part of his *Symbolic Logic*, first published in 1896, with a fourth edition in 1897. There, he explained that the horizontal rectangle in the square was for class c while the vertical stands for d. In the

Fig. 17



manuscript however, probably dated in 1892, the horizontal rectangle is for class d while the vertical stands for c. This is made clear by the small diagrams on top of columns 2 and 3, corresponding to combinations ac and ad respectively. We do not know whether Carroll departed here exceptionally from his regular use or whether he switched rectangles c and d in his quadriliteral diagram between 1892 and 1896.

The point is that once we put class c vertically and d horizontally, the same way Carroll did in the manuscript, the representation of the three premises gives the same marked diagram provided by Carroll, as we will see hereafter. Carroll divided his square as shown in Fig. 17:

Class a covers compartments 1, 2, 3, 4, 5, 6, 7 and 8.
Class a' covers compartments 9, 10, 11, 12, 13, 14, 15 and 16.
Class b covers compartments 1, 2, 5, 6, 9, 10, 13 and 14.
Class b' covers compartments 3, 4, 7, 8, 11, 12, 15 and 16.
Class c covers compartments 2, 3, 6, 7, 10, 11, 14 and 15.
Class c' covers compartments 1, 4, 5, 8, 9, 12, 13 and 16.
Class d covers compartments 5, 6, 7, 8, 9, 10, 11 and 12.
Class d' covers compartments 1, 2, 3, 4, 13, 14, 15 and 16.

The representation of the three premises on this quadriliteral diagram requires the same rules used for the biliteral and triliteral diagrams. The mark 'I' indicates the occupation of a compartment while '0' indicates its emptiness. A mark might be put on a border between two compartments when it is not clear which one it belongs to. It follows that:

I. The first premise "No a is b" tells that compartments 1, 2, 5, and 6 are empty.

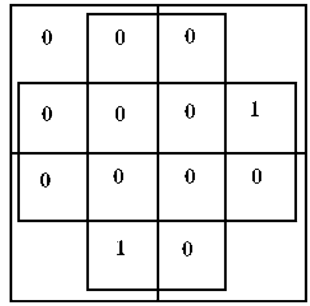
II. The second premise "All c are b" tells that:

- Compartments 3, 7, 11, and 15 are empty.
- At least one among compartments 2, 6, 10, and 14 is occupied. But we know that 2 and 6 are empty (see entry *I* above). So, we infer that either compartment 10 or 14 (or both) is (are) occupied.

III. The third premise "All d are a" tells that:

- Compartments 9, 10, 11, and 12 are empty. But we know that either compartment 10 or 14 is occupied (see entry *II* above). So, we infer that compartment 14 is occupied.
- At least one of compartments 5, 6, 7 and 8 is occupied. But we know that compartments 5, 6 and 7 are empty (see entries *I* and *II* above). So, we infer that compartment 8 is occupied.

Fig. 18



Introducing the appropriate marks on the diagram in accordance with what has been stated above gives the complete quadriliteral diagram (Fig. 18), which is exactly the same that Carroll produced in his manuscript.

3.3 Drawing Conclusions

Finding the conclusion that is to be drawn from the premises depends on what terms one wants to have in the conclusion. Carroll attempts to discuss all possible cases: *ab*, *ac*, *ad*, *bc*, *bd* and *cd*, as long as one is concerned with 2-term propositions. These are the 6 columns that one can see in the bottom of the manuscript. In order to get the conclusions, Carroll proceeds the same way he did with syllogisms. For each case, he transfers information to a new 2-term diagram where are kept only the 2 terms that appear in the conclusion. Those 2-term diagrams are obtained simply by removing the other classes from the quadriliteral diagram, as is shown in Fig. 19. Note that Carroll did not represent in his manuscript the 2-term diagrams in the fourth and fifth columns, corresponding to cases *bc* and *bd* respectively.

Transferring information from the quadriliteral diagram into each of these 2-term diagrams is made according to the same rules that we previously described in the working of syllogisms. The idea is that a compartment is empty only if *all* its subdivisions are known to be empty, and is occupied if *at least one* of its subdivisions is known to be occupied. Once, the information is transferred into the appropriate 2-term diagram, one has just to "read" it there. This last step, as simple as it might look, could still prove to be quite difficult to the beginner who is not acquainted with those diagrams yet. Not only the various 2-term diagrams do have different shapes, but also one single diagram can hold more than just one conclusion. A solution that could be pursued, though we have no evidence that Carroll ever used it, is to transfer again information from each

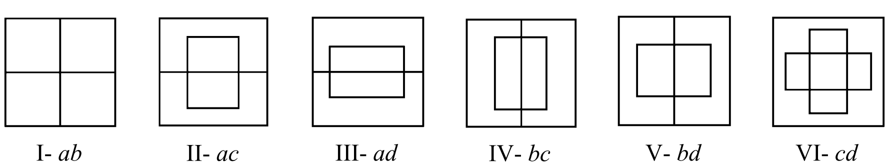
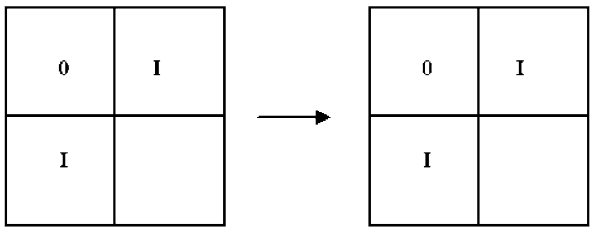


Fig. 19

2-term diagram into the standard (better-looking) biliteral diagram as we described it in Sect. 2.1.

In the following, we will discuss the six cases in the same order as Carroll did. For each case, we will first indicate what terms are kept in the conclusion (i.e. the retinends) and which ones have been eliminated (i.e. the eliminands). We will represent the conclusion on the appropriate 2-term diagram (on the left side), then we will transfer information into a standard biliteral diagram (on the right side). Finally, we will reproduce Carroll's conclusions in subscript and abstract forms, as he listed them in his manuscript, and will complete them when needed.

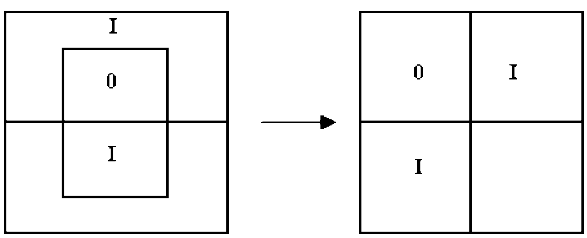
1st Case Retinends: a, b; Eliminands: c, d.



Carroll's conclusions: ab'_1 (i.e. "Some a are *not*-b") and $a'b_1$ (i.e. "Some *not*-a are b").

Carroll overlooks the third conclusion: ab_0 (i.e. "No a is b"), which in combination with the previous ones give final conclusions: a_1b_0 (i.e. "All a are *not*-b") and b_1a_0 (i.e. "All b are *not*-a").

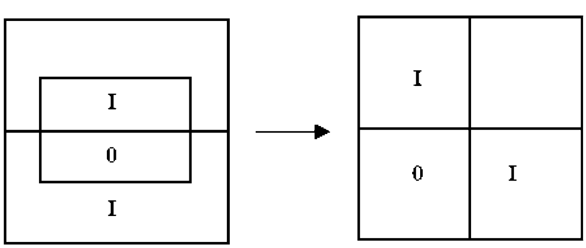
2nd Case Retinends: a, c; Eliminands: b, d.



Carroll's conclusions: ac'_1 (i.e. "Some a are *not*-c"), ac_0 (i.e. "No a is c"), and $a'c_1$ (i.e. "Some *not*-a are c").

Carroll overlooks the final (combined) conclusions: a_1c_0 (i.e. "All a are *not*-c") and c_1a_0 (i.e. "All c are *not*-a").

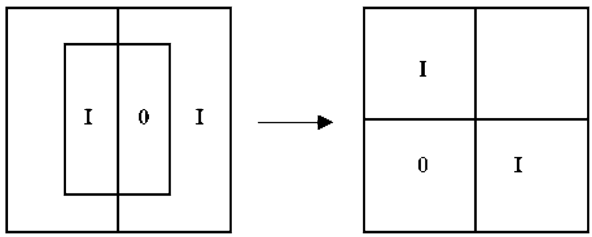
3rd Case Retinends: a, d; Eliminands: b, c.



Carroll's conclusions: ad_1 (i.e. "Some *a* are *d*") and $a'd'_1$ (i.e. "Some *not-a* are *not-d*").

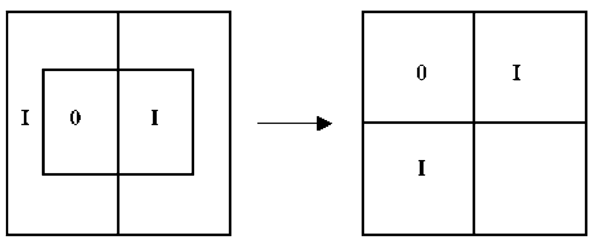
Carroll overlooks the third conclusion: $a'd_0$ (i.e. "No *not-a* is *d*"), which in combination with the previous ones gives final conclusions: a'_1d_0 (i.e. "All *not-a* are *not-d*") and $d_1a'_0$ (i.e. "All *d* are *a*").

4th Case Retinends: b, c; Eliminands: a, d.



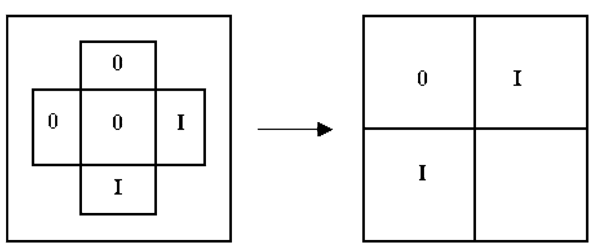
Carroll does not list any conclusion. As such, he overlooks the two (combined) conclusions: $c_1b'_0$ (i.e. "All *c* are *b*") and b'_1c_0 (i.e. "All *not-b* are *not-c*").

5th Case Retinends: b, d; Eliminands: a, c.



Carroll does not list any conclusion. As such, he overlooks the two (combined) conclusions: b_1d_0 (i.e. "All *b* are *not-d*") and d_1b_0 (i.e. "All *d* are *not-b*").

6th Case Retinends: c, d; Eliminands: a, b.



Carroll's conclusions: cd_0 (i.e. "No *c* is *d*"), cd'_1 (i.e. "Some *c* are *not-d*"), and $c'd_1$ (i.e. "Some *not-c* are *d*").

Carroll noted that the combination of the first two conclusions gives c_1d_0 (i.e. "All *c* are *not-d*"). However, he overlooks the fact that the combination of the first and third conclusions also gives d_1c_0 (i.e. "All *d* are *not-c*").

3.4 Symbolic Variations

Now that we gave a complete solution to the logic problem under consideration, we observe that Carroll didn't go so far, and that many of his conclusions were incomplete. In addition, a look at the manuscript shows that Carroll stroke out the problem in the top-right side and added the phrase "not to be used" in the center. All these indications suggest that Carroll never completed solving his logic problem and must have abandoned it at this stage.

The phrase "not to be used" is enigmatic still. A first thought would be that Carroll designed that logic problem at this occasion, and that being unhappy with it, he decided not to use it anymore. However, we know that he did use it both before and after working on it in this manuscript. Indeed that problem does already appear in Carroll's *Fifth Paper on Logic* ([7], [3, p. 208]), a collection of problems that he privately printed in May 1887, probably for use in his logical teaching. The first problem in that paper was:

1. No ducks waltz
2. All officers waltz
3. All my poultry are ducks

This is exactly the problem we discussed lengthily above. Of course, the date of our manuscript is uncertain, and might well be dated prior to 1892 as we assumed. However, we know from Carroll's private diaries that it was on 24 November 1888 that he invented his quadriliteral diagram, at least in the shape it has in the manuscript [19, p. 434]. So, the logic problem under study, already known in 1887, must have been designed by Carroll prior to its use in this manuscript, whatever its date is. This issue is important because it suggests that Carroll first asked his students to solve this problem without his diagrammatic method. As such, they would have appealed to one of the symbolic methods that Carroll designed for sorites. No solutions were provided in the *Fifth Paper on Logic* however.

Interestingly, Carroll also included this logic problem, with some variations that we will mention afterwards, few years later in the first part of his *Symbolic Logic* [12, p. 112]. There, he also provided a solution using the symbolic method of underscoring [12, p. 158]. We do not intend to discuss that method here. It suffices for our purpose to explain that the main idea is to eliminate terms that appear with unlike signs in the premises given in subscript form. For instance, in our manuscript, term b is affirmed in the first premise and denied in the second. Similarly, term a is affirmed in the first premise and denied in the third. The elimination of a and b means that one has only to discuss the sixth case above where c and d appear in the conclusion. As it has been shown in the previous section, in that case, there are two combined conclusions:

(1) "All c are *not*-d" ("All officers are not my poultry");
(2) "All d are *not*-c" (All my poultry are not officers").

Surprisingly, a look at Carroll's symbolic solution in *Symbolic Logic* shows that he didn't reach these conclusions. Indeed, the solution there reads: "My poultry are not officers" which is merely our conclusion (2). Amusingly, Carroll gets here the conclusion which he didn't list in the manuscript and omits the one which he did list alone in the manuscript.

The reason why our conclusion (1) disappears from Carroll's solution is that he slightly changed the expression of the second premise "All officers waltz" (as it appears in the

Fifth Paper on Logic and in the manuscript) into "No officers ever decline to waltz". This change might look trivial, but actually it is not. Indeed, though both propositions are universal, the second alone is negative. As such, in accordance with Carroll's theory of existential import, this new premise does not assert the existence of "officers" (i.e. term *c*) anymore. Thus, none of the premises does assert the existence of *c*, while conclusion (1) does. It follows that conclusion (1) is invalid. Dropping existential import from conclusion (1) would turn it into:

(3) "No *c* is *d*" ("No officer is my poultry").

However, listing conclusion (3) is superfluous because its information is already contained in conclusion (2).

Actually, there is a second minor modification that Carroll made in the expression of the problem as it appears in *Symbolic Logic*. Indeed, he switched letters *b* and *d* in his dictionary, making *b* stand for "my poultry" and *d* for "willing to waltz". This change doesn't seem to involve any consequence in the symbolic solution of the problem. Of course, Carroll arrived at a conclusion containing letters *b* and *c*, rather than *c* and *d* as we did. But that's purely anecdotic because once one applies the dictionary, we get similar concrete terms.

It is noteworthy however that this switch between letters *b* and *d* has a practical consequence when it comes to diagrammatic solving. Indeed, the selection of the 2-term diagram to use depends on what terms we are going to keep in the conclusion. Hence, looking for the relation between *b* and *c*, rather than *c* and *d*, prevents from using the diagram in the sixth column, which happened to be the only one that doesn't have continuous classes. Indeed, contrary to the other 2-term diagrams that divide the universe into four subdivisions each, the sixth diagram has 6 subdivisions. The reason is that compartments *c not-d* and *d not-c* are formed by two discontinuous areas each, which makes that diagram more difficult to use.

4 Conclusion

In this paper, we attempted to make sense of a manuscript where Carroll worked a logic problem with his quadriliteral diagram. Another possible, though unlikely, reading of the manuscript would suggest that Carroll didn't really solve that problem diagrammatically. Indeed, he didn't represent the two 2-term diagrams that should have been drawn on top of columns 4 and 5. This objection does not hold because Carroll did not provide any conclusions for those cases, neither diagrammatic nor symbolic. So, it is likely that he did abandon the problem there. A stronger objection to our reading is that Carroll did not transfer any information at all from the quadriliteral diagram to the 2-term diagrams in columns 1, 2, 3 and 6. Then, conclusions are represented diagrammatically nowhere, while they are explicitly expressed symbolically in each column. Hence, it is possible that Carroll did use those diagrams only as heuristic tools at some stage before proceeding symbolically, rather than carrying out a diagrammatic reasoning proper. Actually, Carroll's logic notebook in the *Parrish Collection* (Princeton University Library), which he apparently used around 1890, contains indeed some marked quadriliteral diagrams that seem to accompany merely solutions that were rather symbolic in their own [8]. However, contrary to that notebook, our manuscript does not show any symbolic method. So, how did Carroll proceed *in practice* to get those conclusions?

An obvious possibility is that he used other (missing) sheets to make calculations. However that holds for both diagrammatic and symbolic methods, and there is no way here to privilege one method over the other. Another possibility is that Carroll represented the premises on the quadriliteral diagram, listed the six cases, draw for each case the corresponding 2-term diagram to visualize what areas should be considered, and then extracted the conclusions *mentally*. This might look uncomfortable for the reader who is not acquainted with those diagrams. However, Carroll was used to working logic problems with his diagrams, and additionally, he was also trained to solving complex problems mentally as he published a full set of such problems to be thought out during "sleepless nights" [11]. If Carroll did work this problem in the way we just described, nothing prevents us from regarding it as diagrammatic reasoning still, even if those diagrams were just mental images not ink on paper. These speculations on the true diagrammatic status of the solution in our manuscript should not make us forget that Carroll certainly knew how to proceed, and very likely did proceed, the way we did in our discussion. This is evidenced by the two diagrams on a 1890 circular that we alluded to in Sect. 2.2.

Now, a look at Carroll's solution of our problem in his *Symbolic Logic* shows that it needs only one line to be solved symbolically, while the diagrammatic solution in the manuscript requires much more time, space and work. We already explained in our discussion of this manuscript that Carroll must have abandoned his work at some stage and never finished his diagrammatic solution. Also, we observed that Carroll provided later a symbolic solution to this problem in his *Symbolic Logic*, while he never published any diagrammatic solution to any problem involving more than 3-terms. An easy shortcut might lead one to think that this would illustrate the superiority of symbolic methods over diagrammatic ones. That would be misleading because Carroll never abandoned diagrammatic methods for complex problems. In the appendix of the first part of *Symbolic Logic*, he described several diagrams for problems involving up to 10 terms [12, p. 179]. As we previously explained, Carroll never managed to finish and publish the second part, so we do not know whether he would have made use of them there or not. It is true that when the number of terms increases, diagrams become more complex and difficult to grasp. Hence, it is understandable that logicians, Carroll included, might prefer other methods for problems involving more than, let us say, 6 or 7 terms. However, our manuscript is not about this issue because the problem it discusses has just 4 terms, which is still workable diagrammatically.

We made the preliminary remarks above in order to dispel some misconceptions that might divert us in this conclusion from what we think is the main issue here. Besides the use of a diagrammatic method to solve the problem in the manuscript, the interesting point is precisely Carroll's idea of what a logic problem is. Indeed, Carroll looks there for all possible conclusions as to the relation between any 2 terms involved in the argument. As far as we know, Carroll never reworked this way in his *Symbolic Logic*, not even with his symbolic methods. In the subscript solution we described briefly in Sect. 3.4, Carroll discussed only one case (the sixth), while he did explore six cases in the manuscript. Hence, it does not make sense to compare the two methods, because Carroll was pursuing there two different paths. If we were to discuss one case merely, after determining what terms should be in the conclusion as has been done in the example in Sect. 2.3, the diagrammatic method would be perfectly appropriate, easy, efficient and reliable.

Acknowledgements This work benefited greatly from conversations with many people, notably Francine Abeles, Anthony Edwards, Mark Richards and Edward Wakeling, to whom I express my gratitude. This paper draws upon work supported by a research grant from *The Friends of the University* *of Princeton Library*. Finally, I express grateful acknowledgements to the archivists who helped me to consult the material used in Pierpont Morgan Library (New York) and Princeton University Library.

References

1. Abeles, F.F.: Lewis Carroll's formal logic. Hist. Philos. Logic **26**(1), 33–46 (2005)
2. Abeles, F.F.: Lewis Carroll's visual logic. Hist. Philos. Logic **28**(1), 1–17 (2007)
3. Abeles, F.F. (ed.): The Logic Pamphlets of Charles Lutwidge Dodgson and Related Pieces. Lewis Carroll Society of North America, New York (2010)
4. Bartley, W.W. III (ed.): Lewis Carroll's Symbolic Logic. Clarkson N. Potter, New York (1977)
5. Boole, G.: An Investigation of the Laws of Thought. Walton & Maberly, London (1854)
6. Carroll, L.: The Game of Logic. Macmillan, London (1887)
7. Carroll, L.: Fifth Paper on Logic. The Parrish Collection, Princeton University Library, Dodgson 431 (1887)
8. Carroll, L.: Algebraical Geometry. Notebook. The Parrish Collection, Princeton University Library, box 1, folder 1 (1890?)
9. Carroll, L.: Two Diagrams. The Parrish Collection, Princeton University Library, Dodgson 224 & 225 (1890?)
10. Carroll, L.: Notes and Calculations on Problems in Symbolic Logic. The Houghton Collection, Pierpont Morgan Library, New York, AAH545 (1892?)
11. Carroll, L.: Pillow Problems. Macmillan, London (1893)
12. Carroll, L.: Symbolic Logic—Part I, 4th edn. Macmillan, London (1897)
13. Marquand, A.: Logical diagrams for *n* terms. Philos. Mag. **12**, 266–270 (1881)
14. Moktefi, A.: Lewis Carroll's logic. In: Gabbay, D.M., Woods, J. (eds.) British Logic in the Nineteenth-Century, pp. 457–505. North-Holland, Amsterdam (2008)
15. Moktefi, A., Edwards, A.W.F.: One more class: Martin Gardner and logic diagrams. In: Burstein, M. (ed.) A Bouquet for the Gardener: Martin Gardner Remembered, pp. 160–174. Lewis Carroll Society of North America, New York (2011)
16. Moktefi, A., Shin, S.-J.: History of logic diagrams. In: Gabbay, D.M., Pelletier, F.J., Woods, J. (eds.) Logic: A History of Its Central Concepts, pp. 611–682. North-Holland, Amsterdam (2012)
17. Venn, J.: On the various notations adopted for expressing the common propositions of logic. Proc. Camb. Philos. Soc. **4**, 36–47 (1880)
18. Venn, J.: Symbolic Logic, 2nd edn. Macmillan, London (1894)
19. Wakeling, E. (ed.): Lewis Carroll's Diaries, vol. 8. Lewis Carroll Society, Clifford (2004)

A. Moktefi (✉)
IRIST, Université de Strasbourg, 7 Rue de l'Université, 67000 Strasbourg, France
e-mail: moktefi@unistra.fr

A Diagrammatic Bridge Between Standard and Non-standard Logics: The Numerical Segment

Ferdinando Cavaliere

Abstract The system of 'Distinctive' Predicate Calculus (DPC) used here is part of a more general enterprise to bring logic closer to natural language and natural intuitions. This system works on the central concept of "middle" or "intermediate" and finds a synthesis in the diagram of the 'Numerical Segment' (NS). We intend to show how this diagrammatic tool and its logic can be theoretically used as an interesting link up with non-standard logics (fuzzy, polyvalent, paraconsistent) and applied to new form of diagrammatic reasoning.

Keywords Quantifiers · Natural languages · Distinctive logic · Logical diagram · Numerical syllogisms · Non-standard logic

Mathematics Subject Classification (2010) 03B20 · 03B52 · 03B65 · 03B60 · 03B52 · 05C99

1 Distinctive Segment from Classical Logic to Numerical Systems

The system of 'Distinctive' Predicate Calculus (DPC) used here is part of a more general enterprise to bring logic closer to natural language and natural intuitions. This system works on the central concept of "middle" or "intermediate" predicate. Categorical predicates are interpretable as points or subsegments of a 'Distinctive Segment' (DS). In the 'Numerical Distinctive Segment' (NS) every point represents a precise numerical predication: a diagrammatic representation of a distinctive numerical calculus is set out.

1.1 The Concept of 'Middle'

In classical logic the contraries are always two in number: they come in pairs, either of terms (predicates) or universal predications. By definition it is impossible for contraries to subsist together, but 'intermediate' or 'middle' between two contraries are possible. Thus, *gray* can be predicated as an intermediate between *black* and *white*. There may be more than one intermediate, such as *pallid*, between *black* and *white*. Such intermediates always represent a partial privation of the extreme properties, in that the two corresponding contraries represent the maximal (perfect or complete) difference. Thus, by using a metaphor, the *extremes of a line segment* are the *two points maximally distant* from each other.

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic, DOI 10.1007/978-3-0348-0600-8_5,

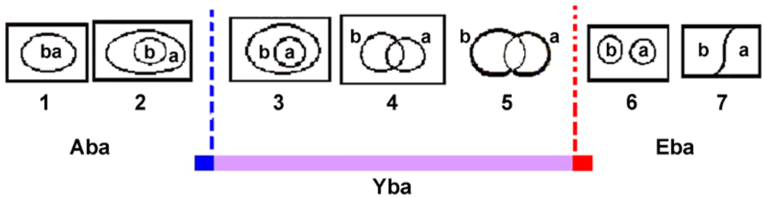


Fig. 1 The distinctive segment

For some pairs of contraries there is no intermediate: they are *nongradable* predicates. Examples are *even* and *odd* (for numbers), and also, *true* and *false*. Every predication is either true or false, and from the truth of the one the falsity of its negation is derived, and vice-versa. This defines a *bivalent logic*. In classical logic, the affirmative and negative existential categoricals are defined without any reference to intermediates. Intermediates are excluded between contradictories, but as investigated by R. Blanché [5, pp. 36–39], on the one hand, Aristotle depicts the existential categorical *Some b is a* (traditionally symbolized by **Iba**) as being 'partial' ('a part but not all'), excluding all; on the other hand, when developing his predicate logic and his syllogistic, the very same existential form is interpreted as 'at least one and possibly all'.[1] The term that has become accepted in the tradition (since Apuleius) for *Some b is a* or *Some b is not-a* (the latter traditionally symbolized by **Oba**) is 'particulars'.

By Blanché's interpretation, we can presume that the two universals, being contraries, represent the extremes of the opposition between *All b is a* (in symbols: **Aba**) and *No b is a* (in symbols: **Eba**), which can thus admit intermediates. If *Every man is white* and *No man is white* are contraries, what sort of intermediate situation can we imagine?

1.2 The Distinctive Segment (DS)

As we can deduce from Blanché [4, 5], it is the *conjunction* (here symbolized by "*") *of the two particulars* that forms the intermediate between the two contrary universals. The 'partial', hinged on the quantifier 'only some' (or 'exclusive existential' symbolized by **Y**), is the natural intermediate between the contrary universals. Yet Aristotelian syllogistic, and with it the standard classic syllogistic of the Western world, would take a different course. In fact, the predicate 'partial' will not be taken into consideration until the eighteenth century, with Ploucquet or von Holland.[2]

In Fig. 1, next to the formulas one finds the seven corresponding Venn-diagrams for all possible relations between (contextually defined) pairs of sets that make the associated

[1]Blanché tells the story of a "mortal combat" between the partial and the particular in the evolution of Aristotle's thinking [5, p. 39, note 27]. A forerunner of Blanché was Sesmat [21].

[2]See Lenzen [16]. J.H. Lambert writes: "...we have these three cases: 1. *Every b is a*, 2. *Only some b is a*, 3. *No b is a*. Of the three propositions only one can be true at any given time" [15, pp. 76–77] (translation in English by the author). The falsity of one does not allow one to conclude to the truth of either other.

formula true.[3] ***Yba*** represents precisely the cases that are intermediate between *Aba* and *Eba*. One could say that Y is 'spread' over the segment running from A to E, with the exclusion of the two extremes.

We call this "simplified square of opposition" as "*Distinctive Segment"* of opposition or DS. In this simple diagram the purple represent a middle color between the blue and red of the extreme points, by analogy with corresponding predicates.[4] The predication *Only some b is a*, is affirmative only in the grammatical sense, but we distinguish in the subject-class b a subset of elements that a may be predicated of, besides a subset of elements that a cannot be predicated of, so that a' ($=$ *not* a) is predicable of them. For this reason we call the segment "distinctive".[5] Thus a geometrical metaphor is transformed in a logical structure.

1.3 The Numerical Distinctive Segment (NS)

After the first attempts by Lambert, the numerical syllogisms were developed in an algebraic context, from De Morgan, Peirce and Keynes [9, pp. 38, 42, 135], to the epigones of the last century such as Hacker and Parry [10], Carnes and Peterson [6] or Murphree [18, 19] (see Pfeiffer [20]).

One restrictive feature of the system based on a tripartite DS is the fact that it captures all intermediate situations into one single undifferentiated predication, namely the partial (or particular). The major advantage of DS is that it permits the refining of numerical quantification. The DS model is subdivided into as many intervals as are needed, depending upon the number of elements in question. Not only the vertices, but also every point on the Numerical Distinctive Segment or Numerical Segment (NS) represents a precise numerical predication, and conversely. This is the reason why our diagrammatic scheme does not make use of circles or other figures that do not have 2 extremes and a uni-dimensional (rectilinear) order for all their points.

---

[3] All sets in question are distinct from the universe of objects u and from the null set [3, pp. 53–55]. In the cases 5 and 7, u coincides with the union of the two sets b and a, the difference being that, in case 7, b is the complement of a in u. In the other five cases, u is represented by the rectangular frame. One notes that such a mutually exclusive distribution over the proposition types is not possible in the classic square of opposition. The 7 types of propositions can already be found in De Morgan [8, pp. 65–67] that called them complex propositions and symbolized by $D - D$, $-D' - P - C' - C$, $-C$ (respectively for our cases from 1 to 7). In the same work, De Morgan makes use of segments (or letters arranged in rows) to illustrate propositions [8, pp. 61, 79, 81–82, etc.]. After De Morgan, Keynes and other logicians identified the 7 cases. None of them (De Morgan included) used the "partial" quantificator.

[4] The same goes for the next Figs. 2 and 3, while in Figs. 4, 5 and 6, the colors have the function of underscoring.

[5] On this basis we have constructed the Distinctive Triangular and Hexagonal Calculi [7, pp. 242–246]. The known laws of immediate inference have now been extended with the equivalence of the two Y-obverses: $Yba = Yba'$ (where a' stands for the complement of a in u). The Russian logician N.A. Vasil'ev developed a truly tripartite logic like the one just sketched. In 1910, he called it the Logic of Notions. Besides the proposition types that he called 'general' (that is, our 'universal'), he posited the 'accidental' type, and the 'judgments' could be not only affirmative or (internally) negative, but also indifferent. See Seuren [22] and Suchon [23] about triangular schemata, Blanché [4] and Beziau [2] for hexagonal schemata developments.

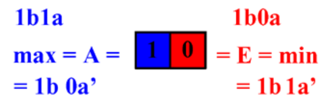


Fig. 2 NS of a singular predicate

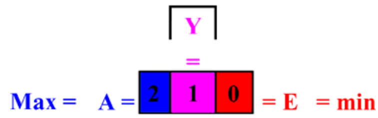


Fig. 3 NS for a subject with 2 elements

In NS the predication is specified according to the cardinality of the subject term referent: “*Four out of nine dogs escaped*” or “*9d4e*”. In traditional syllogistic, predications like *Socrates is a politician* or *Socrates is not a politician* (or *Socrates is a non-politician*) are called, respectively, *affirmative* and *negative singulars*. In modern logic they are known as (Singular) Definite Descriptions. When numerical quantifiers are used, the first case becomes “*1s1p”*: “precisely one *s* is *p*”, while the second becomes “*1s0p”*: “precisely none [zero] of *s* is *p*” or “exactly one *s* is *non-p*”. To represent the two predications using a model analogous to NS we must resort to two distinct adjacent points, without intermediates (Fig. 2).

The minimal number of elements of the subject that can give rise to a “partial” is 2, as is shown in Fig. 3. From this figure we gather that the affirmative universal *all* coincides with the *maximal numerical quantifier* (*“two out of two b are a”*); the negative universal *none* with the *null numerical quantifier* (“*of 2 b 0 is a*”), while the distinctive particular *only some* becomes “*of two b one is a*”.

With a subject of 4 elements (Fig. 4)[6] the distinctive particular *only some* becomes the disjunction of all intermediate numerical quantifiers: *“of 4 b [3 v 2 v 1] are a*”.

The traditional quantifiers and those known in the literature as “intermediate” [9], or “quasi-numerical”, are translated into numerical ones by means of adjunctive symbols. Thus, $>$ stands for more, $<$ for less, $\geq$ for (more or equal) or at least, $\leq$ for (less or equal) or at most. Such symbols are abbreviations of disjunctions of numerical quantifiers, as shown in Fig. 5, where the subject class has cardinality 6.

Below the extremes of NS are placed in a scalar way, and the expressions of intervals/quantifiers preceded by the expression “at least” or “at most”, which are considered primitive by those who have so far looked at numerical syllogisms. By contrast, we consider more intuitive the expressions preceded by “exactly”, the other being considered as derived. The same holds for the intermediate quantifiers.

Every expression of the left or right laterals implies those below it. Figure 5 shows how the possible oppositive square has at the left or right laterals the projections of the progressive intervals traceable to NS. We can associate with these laterals not only the numericals of the authors just mentioned, but also the scalar or gradual linguistic expressions studied by L. Horn [13, pp. 236–238]. The obversion of a numerical predication comes

[6] The acronym “alo” is taken from a talk of J.-Y. Béziau (at the Congress “Logic Now and Then”, Brussels, November 5–7, 2008).

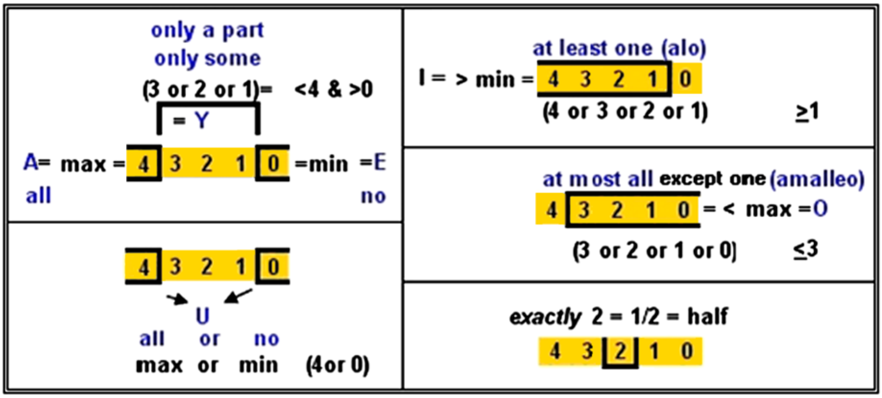


Fig. 4 NS with 4 elements

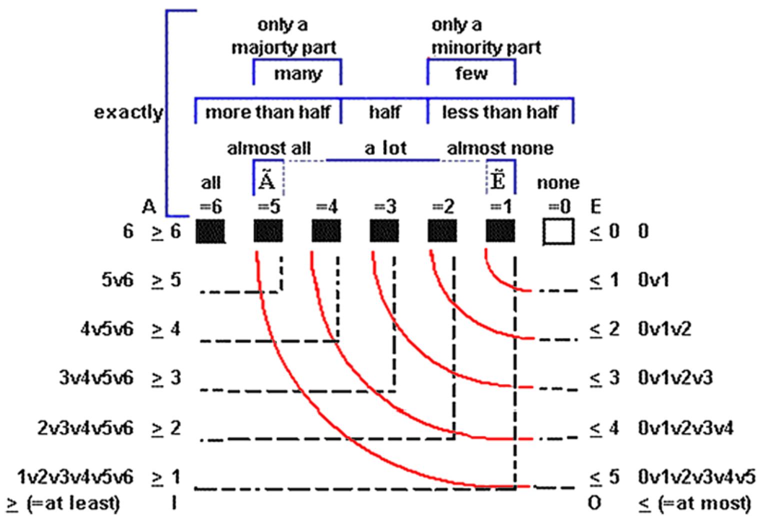


Fig. 5 NS with expressions of natural language

about when symmetrical quantifiers are interchanged with regard to the central point of NS.[7]

Just as the preceding Distinctive systems, the obversion of a numerical predication is obtained by swopping the quantifiers that are symmetrical with respect to the vertical axis. For example, "*Of 6 b*, *exactly 2 are a* = *Of 6 b*, *exactly 4 are non-a*". Here it becomes clear that the intermediates are a mixture of contraries, inversely proportional to each other.

[7]The square does not exhaust the possible combinations. For example, it does not contain those with discontinuous quantifiers/intervals, or those that lack extremes, such as have various intermediate quantifiers between them.

Fig. 6 Double NS

1.4 The Double Segment and Mediate Inferences

With the help of the numerical quantifiers we can quantify intervals, not just of the subject but also of the predicate; thereby realizing the dream of a syllogistic sought in vain by W. Hamilton [11], C.E. Stanhope (see Harley [12]) and other proponents of this idea. For example, a formula like "(6 S) $\geq 3 \leq 5$ ($P \geq 7$)" can be read as "Among the S's, which are exactly 6 in number, from 3 to 5 are P's, which are at least 7 in number".

If we want a complete diagrammatic representation of the entire universe, we must add a second NS to the first for the complement of the subject (or of the two terms). That way we can represent composite categoricals [7, p. 253]. For example: "*4b 3a4 b*′*6* = (*4b3a* & *6b*′*4a*)" (see Fig. 6).

This diagram is inspired by similar (without numerical attributes) schemata of Leibniz, Lambert, De Morgan, Stanhope, Macfarlane and other logicians. The density of information in this diagram, as regard all the relations among the classes involved (including the u), and its geometrical shape, makes it suitable for the numerical mediate inference [9, pp. 38, 42]. We only need to draw a second double segment, parallel to the first, in which a new variable appears. The third double segment, for the conclusion, rises by confrontation between the two above. This method would be simpler than the equivalent operations based on the rules for complex numerical predicates, but it is still at an experimental stage [7, pp. 252–253].

2 Numerical Segment and Non-standard Logics

The diagrammatic tool of NS can be applied to new form of diagrammatic reasoning, with n-level of numerical segments, which transcend the theoretical antinomy between logics of Non-contradiction and Non-standard Logics.

2.1 Fuzzy Segment (FS) and 'Interbivalence' of Natural Language

The underlying structure of the Distinctive Logic reveals an isomorphism with regard to a natural logic that we call "interbivalent" and in which we define, besides the values true and false, also the values partially true and partially false. The NS becomes "fuzzy" (in the sense of Zadeh [24] or Kosko [14]) when discrete segments are replaced with continuous ones—that is, with segments that are infinitely subdivisible.

We could let the points of NS represent the elements of the subject class, while ordering them from the point with the highest value on the fuzzy scale to the predicate assigned (membership function) to the one with the lowest value. The gradability or scalarity thus turns from discrete into continuous or variegated, to generate a "Fuzzy" Segment (FS) (see Fig. 7).

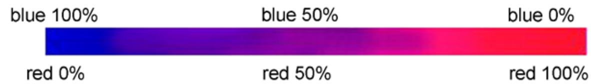


Fig. 7 The fuzzy segment

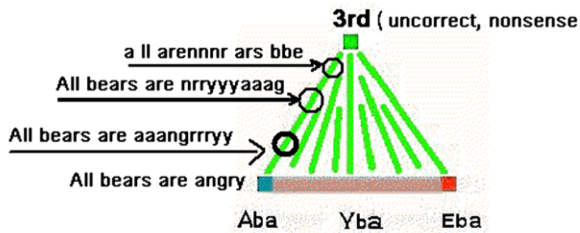


Fig. 8 NSs for predicates between meaning and nonsense (3rd value)

2.2 NS and Trivalence

NS can be positioned in *n*-dimensional space. Fuzzy systems that are generally named polyvalent, should be renamed Interbivalent, in that the value of a proposition is true or false or a calibrated mixture of these two. In our classification, an authentic trivalence requires a third value that is heterogeneous to the true and the false, such as "devoid of sense", as in *All bears are yyygggrrraaaa*, which is not well-formed. Here too it is possible to specify situations that are intermediate between any two values in the system. As an example we might quote a sentence like *All bears are aaangrryy* (that's *almost* well-formed) or other sentences of that nature (Fig. 8, where numerical values are not explicit, but graspable by analogical way).

2.3 NS and Paraconsistent Logics

The Aristotelian formula: "the same property cannot at the same time belong and not belong to the same entity from the same point of view" [1, pp. 142–143][8] is, of course, well-known. So we ask ourselves: and what if we change the point of view? Or the level of analysis? In the latter case we can use two or more DS in parallel, one for every level or sub-level. For example (where useful = not-dangerous and vice versa):

(All) knives are useful (sub-level 1A): if used properly;
(All) knives are dangerous (sub-level 1B): if used inexpertly;
(All) knives are useful and dangerous (level 2): without taking their use into account (synthetic).

Here we have paraconsistent or dialectical predications [17, pp. 163–175]. We can handle paraconsistent situations using the following technique. Construct a table of the parallel truths, showing for each point of view the interval that satisfies it on the pertinent NS.

This yields, as shown on the left hand side, all possible NS for the case at hand.[9] Now decide which kind of logic you wish to adopt at second-level (the right hand side of

[8]Translation in English by the author.

[9]N/M is an intermediate numerical quantifier.

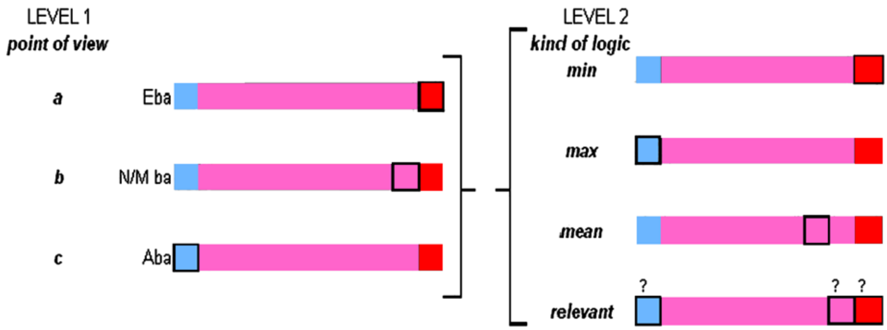


Fig. 9 Diagrammatic reasoning with the points of view *a*, *b*], *c*

Fig. 9). If you consider only the minimal result from among the results given, the final NS will be the first one. If only the maximal, it will be the second one. If you choose a result in between (mean), it will be the third one. Finally, you can decide to take one specific point of view as the only relevant one, eliminating any other point of view (the lower NS).

3 Conclusion

The NS is a simple diagram, a graphical transposition of some concepts of the Distinctive Logic elaborated by the author. It's a structure that can represent many kinds of natural predicates: categorical, singular, numerical, fuzzy and paraconsistent ones. Its iconicity simplifies the analysis of many composed proposition and open the way to further visual methods for standard or non standard reasoning.

References

1. Aristotle: Metafisica (Italian translation by G. Reale, with the Greek text in front). Bompiani, Milan (2004)
2. Béziau, J.-Y.: The power of the hexagon. Logica Univers. **6**(1–2), 1–43 (2012)
3. Bird, O.: Syllogistic and Its Extensions. Prentice Hall, Englewood Cliffs (1964)
4. Blanché, R.: Structures Intellectuelles: Essai sur l'Organisation Systematique des Concepts. Vrin, Paris (1966)
5. Blanché, R.: La logica e la sua storia da Aristotele a Russell (Italian translation by A. Menzio). Astrolabio-Ubaldini, Rome (1973)
6. Carnes, R.D., Peterson, P.L.: Intermediate quantifiers vs percentages. Notre Dame J. Form. Log. **32**(2), 294–306 (1991)
7. Cavaliere, F.: Fuzzy syllogisms, numerical square, triangle of contraries, interbivalence. In: Beziau, J.Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition, pp. 241–260. Springer, Basel (2012)
8. De Morgan, A.: Formal Logic, or the Calculus of Inference, Necessary and Probable. Taylor and Walton, London (1847)
9. Gardner, M.: Logic Machines and Diagrams, 2nd edn. University of Chicago Press, Chicago (1982)
10. Hacker, E.A., Parry, W.T.: Pure numerical Boolean syllogisms. Notre Dame J. Form. Log. **8**(4), 321–324 (1967)
11. Hamilton, W.: Lectures on Metaphysics and Logic, vol. 4. Blackwood, Edinburgh (1866)

12. Harley, R.: The Stanhope demonstrator. Mind **4**, 192–210 (1879)
13. Horn, L.R.: A Natural History of Negation. University of Chicago Press, Chicago (1989)
14. Kosko, B.: Fuzzy Thinking: The New Science of Fuzzy Logic. Hyperion, New York (1993)
15. Lambert, J.H.: Nuovo Organo (Italian translation by R. Ciafardone). Laterza, Bari (1977)
16. Lenzen, W.: Ploucquet's 'refutation' of the traditional square of opposition. Logica Univers. **2**, 43–58 (2008)
17. Moretti, A.: The geometry of logical opposition. PhD thesis, Université de Neuchatel, Switzerland (2009)
18. Murphree, W.A.: Numerically Exceptive: A Reduction of the Classical Syllogism. Peter Lang, New York (1991)
19. Murphree, W.A.: Numerical term logic. Notre Dame J. Form. Log. **39**(3), 346–362 (1998)
20. Pfeifer, N.: Contemporary syllogistics: comparative and quantitative syllogisms. In: Kreuzbauer, G., Dorn, G. (eds.) Argumentation in Theorie und Praxis: Philosophie und Didaktik des Argumentierens, pp. 57–71. LIT, Vienna (2006)
21. Sesmat, A.: Logique. Hermann, Paris (1951)
22. Seuren, P.A.M.: The Logic of Language. Language from Within, vol. 2. Oxford University Press, Oxford (2010)
23. Suchon, W.: Vasil'iev: what did he exactly do? Log. Log. Philos. **7**, 131–141 (1999)
24. Zadeh, L.A.: A computational approach to fuzzy quantifiers in natural languages. Comput. Math. Appl. **9**(1), 149–184 (1983)

F. Cavaliere (✉)
V. della Libertà 5, 47042 Cesenatico (FC), Italy
e-mail: cavaliere.ferdinando@gmail.com

Diagrammatic Reasoning with Classes and Relationships

Jørgen Fischer Nilsson

Abstract We present and discuss a diagrammatic visualization and reasoning language coming about by augmenting Euler diagrams with higraphs. The diagrams serve (hierarchical as well as trans-hierarchical) classification and specification of various logical relationships between classes. The diagrams rely on a well-defined underlying class-relationship logic, called CRL, being a fragment of predicate logic. The inference rules at the level of diagrams take form of simple diagrammatic *ipso facto* rules. The diagrams are intended for computerization by offering navigation and zooming facilities as known from road maps. As such they may facilitate ontological engineering, which often involves larger amounts of data. The underlying inference process is expressible in function-free definite clauses, DATALOG. We also discuss the relationship to similar diagram and logic proposals.

Keywords Logic diagrams · Diagrammatic reasoning · Relationships between classes · Logical visualization principles and tools

Mathematics Subject Classification (2010) Primary 03B80; Secondary 68U35

1 Introduction

Diagrammatic visualization and reasoning is becoming increasingly important due to the rise of applications which require large scale ontologies and specifications. Diagrammatic visualization is also spurred by availability of computer screen dynamics.

We consider diagrammatic visualization and reasoning for a class relationship logic, CRL, achieved visually by extending Euler diagrams with higraphs [18], a tradition surveyed in [19]. The diagrams afford intuitive appealing inference principles inherent in the visual formalism. The diagrams are intended for facilitating computer assisted reasoning in particular also for large scale ontological engineering.

The main pragmatic motivation and desiderata are as follows:

- Formal ontology engineering and domain modelling call for appropriately adapted fragments of predicate logic.
- Specifications may preferably be rendered as diagrams as an alternative to sentences. The accompanying deductive reasoning capabilities should be reflected in the diagrams in an intuitive manner.
- The opportunities possessed by the computer screen for flexible and dynamic visualization should be exploited for management of large specifications.

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic,
DOI 10.1007/978-3-0348-0600-8_6,

- The dynamic visualization capabilities are to be integrated with the visual inference abilities for querying and browsing purposes.

The backbone of formal ontologies is made up of class inclusion relationships (class subsumption). Often ontologies are augmented with ascription of properties to the specified classes. Formal ontologies traditionally are depicted as Hasse diagrams, that is directed acyclic graphs representing the inclusion relationships between classes. It is tempting to replace the Hasse diagrams with Euler diagrams. However, to this end Euler diagrams should preferably be extended with facilities for expressing arbitrary relationships between classes, in addition to inclusion relationships. This is a key issue in the present paper. In order to ensure rigor, we apply predicate logic as basis in our specification. However, the intention is that the diagram proposal can be applied without knowledge of the predicate logical underpinning.

The paper is organized as follows: Section 2 presents the key ideas of the class relationship logic dialect CRL. Section 3 introduces to CRL diagram reasoning, whereas Sect. 4 explains the informal language understanding of diagrams. Section 5 explains diagrams more formally, followed by formal elucidation of the logical inference principles in Sect. 6.

Section 7 discusses some key structures in diagrams, and Sect. 8 introduces the notion of analytical relationships. Finally is CRL compared with similar logic proposals in Sect. 9, followed by a concluding section.

2 Class Relationship Logic at a Glance

Class relationship logic, CRL [9], is concerned with classes $c, d, \ldots$ from a finite collection of classes $\mathcal{C}$ and binary relationships between classes $r, \ldots$ from a finite collection $\mathcal{R}$. Classes are dealt with intensionally, but may be understood extensionally as named sets of individuals standing *inter alia* in class inclusion relationships.

The prime CRL logical relationship form is the so-called ∀∃-form, explicated in predicate logic as

$$\forall x\big(c(x) \rightarrow \exists y\big(r(x, y) \wedge d(y)\big)\big)$$

This logical sentence form is of fundamental interest in domain modelling and in particular formal ontologies for two reasons:

- It encompasses as distinguished, important case the extensional class inclusion relationship *isa*

$$\forall x\big(c(x) \rightarrow d(x)\big)$$

  coming about by letting r be identity "=" in the ∀∃-form.
- It admits attribution of class properties to classes in that the form attributes the class d via attribution with r to the class c. This attributed property is inherited to subclasses of c by logical entailment.

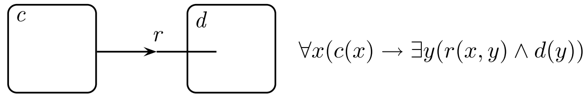


Fig. 1 Class-class relationship

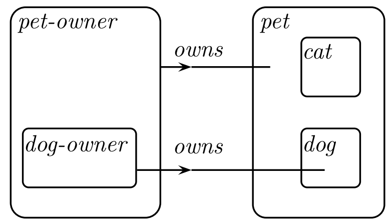


Fig. 2 Class-class relationship with subclass

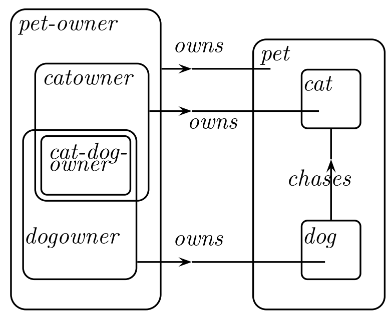


Fig. 3 Class-class relationship with class overlap

2.1 Forms of CRL-Diagrams for the ∀∃-Relationship

The various CRL relationship sentences have accompanying diagram forms. Let us first consider the form ∀∃, for which we apply the diagram in Fig. 1.

The arrow tells the direction of the relation and does not imply functionality. The diagram is to visualize that all elements in c are r-related to some element(s) in d. This convention for rendition is intended to reflect the logical inferences admitted with the CRL sentences as to be accounted for in Sect. 6.

Figure 2 shows a sample (incomplete) CRL diagram, telling that dog owners are pet owners, that all pet owners own a pet, and that cat and dogs are disjoint classes of pets—in turn being disjoint with the class of pet owners.

Figure 3 elaborates on this example. It introduces a class "cat-dog-owner" in the overlap between cat owner and dog owner, which intuitively is to inherit the relationships of its super classes by visual and logical inference rules. There is also a so called ∀∀ relationship, cf. Fig. 5, between dogs and cats, expressing that all dogs chase all cats.

3 Diagrammatic *Ipso Facto* Reasoning

We argue that those logical consequences of CRL specifications which themselves take form of CRL sentences are implicitly present in the CRL diagram by simple diagrammatic considerations. In this sense CRL diagrams offer *ipso facto* reasoning, cf. also the elaborate discussion of iconic logics in [24].

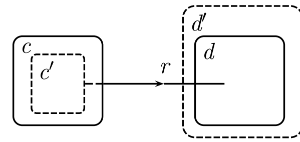


Fig. 4 Diagrammatic inference principle

3.1 Inference Principle for ∀∃ Relationships

The partial order properties of the class inclusion relationship are reflected directly in the topological properties of diagrammatic inclusion.

The Eulerian topological inclusion principle for class inclusion is extended with rules concerning relationships. The diagrammatic inference principles for relationships is illustrated for a ∀∃ relationship with r for classes c and d in Fig. 4. The figure tells that the shown r-relationship holds also for subclasses of the left class c, as well as for superclasses of the right class d.

The vertices of relationship arcs are situated either at the border of a box or in the interior of a box. The former case is called ∀-end point, because the pertinent relationship applies to the entire inner of the box. The latter case is called an ∃-end point because it applies to some part of the interior of the box.

The ∀∃ sentence form $\forall x(c(x) \rightarrow \exists y(r(x, y) \wedge d(y)))$ may be abstracted as the combinator term

$$\forall\exists(c, r, d)$$

where ∀∃ is then conceived of as the name of a ternary combinator. We refer to this form of logical sentence as the Peirce form since it is the predicate logical counterpart of the Peirce product in Boolean modules, cf. [4].

3.2 Inventory of Logical Relationship Forms

At present we consider 4 different CRL relationship forms

$$\forall\forall) \quad \forall x\big(c(x) \rightarrow \forall y\big(d(y) \rightarrow r(x, y)\big)\big)$$
$$\forall\exists) \quad \forall x\big(c(x) \rightarrow \exists y\big(d(y) \wedge r(x, y)\big)\big)$$
$$\exists\forall) \quad \exists x\big(c(x) \wedge \forall y\big(d(y) \rightarrow r(x, y)\big)\big)$$
$$\exists\exists) \quad \exists x\big(c(x) \wedge \exists y\big(d(y) \wedge r(x, y)\big)\big)$$

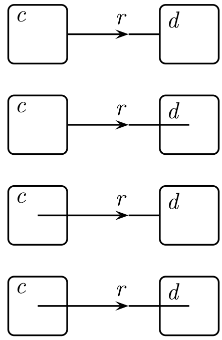
These sentence forms may be abstracted as the 4 combinator forms

$$Q_1 Q_2(c, r, d)$$

where $Q_i \in \{\forall, \exists\}$. The ∀∃ sentence form $\forall x(c(x) \rightarrow \exists y(r(x, y) \wedge d(y)))$ may thus be abstracted as the combinator term $\forall\exists(c, r, d)$.

These forms are reminiscent of De Morgan's schemas of categorial propositions, cf. [23]. For the special case $\forall x(c(x) \rightarrow d(x))$ we use the combinator form $isa(c, d)$.

Fig. 5 Diagrams for the 4 atomic relationships $Q_1 Q_2(c, r, d)$



∀∀) $\forall x(c(x) \rightarrow \forall y(d(y) \rightarrow r(x, y)))$

∀∃) $\forall x(c(x) \rightarrow \exists y(d(y) \wedge r(x, y)))$

∃∀) $\exists x(c(x) \wedge \forall y(d(y) \rightarrow r(x, y)))$

∃∃) $\exists x(c(x) \wedge \exists y(d(y) \wedge r(x, y)))$

3.3 Diagram Forms for the Four Basic CRL Relationships

Figure 5 shows the diagrams for the 4 available atomic relationships $Q_1 Q_2(c, r, d)$.

For a relationship arc r from c to d, in short $Q_1 Q_2(c, r, d)$, c is called the start box and d the end box. The edges initiating at a box c are called the outlets. Analogously, the arcs ending in a box c are called inlets.

An ∃-point has one inlet and no outlets for $Q_1\exists(c, r, d)$ or one outlet and no inlets for $\exists Q_2(c, r, d)$ as confirmed in the diagrams of Fig. 5.

An arc named r is unique for two given (not necessarily distinct) classes c and d and given logical type and direction; but arcs labelled with r may connect other pairs of classes.

3.4 Reasoning with Relationships

The classes in $\mathcal{C}$ are assumed non-empty. Thus there is no notion of empty class in the present form of CRL. Insistence on non-emptiness of classes, as in Aristotelian logic, that is, adoption of existential import, cf. e.g. [28], simplifies the inference system.

Furthermore, the assumption of non-empty classes makes it possible to specify intersection (overlap) of two classes c and d simply by positing presence of a class (thus non-empty) being included in c and d. Complementarily, disjointness (exclusion) of two classes c and d is expressed by absence of any common inclusion class together with appeal to the closed world assumption.

In our logical set up the closed world assumption may be exploited computationally at the metalogical level by non-monotonic appeal to failure to prove. Altogether, in this way the inclusion form of the ∀∃-relationship, that is

$$\forall x\big(c(x) \rightarrow d(x)\big)$$

enables expressing of the three ontologically fundamental class-class relationships, viz. class inclusion, overlap, and exclusion as known from Euler diagrams.

The reasoning principles at the diagrammatic level are coined into the following rules for class-class relationships:

- *Inheritance*: A ∀-end point extends as ∀ end point to inner boxes. This applies recursively for nested boxes.

- *Generalization*: An ∃-end point of a boxes act also as ∃ end point for all including boxes.
- *Weakening* of ∀ to ∃: A ∀-end point can be extended inwards to become an internal, that is ∃, end point.

In Sect. 6 we state the inference rules explicitly in their underlying logical form.

The combinator forms $\forall\exists(c, r, d)$ etc. are used directly in the supporting metalogic, where ∀∃ is re-conceived as a plain first order predicate taking three arguments with c, r and d being reified as individuals in the first order predicate logic applied as metalogic for CRL. Still, however, they function as classes and relationships at the CRL logical level. Accordingly, as elaborated in Sect. 6, the computerized reasoning with CRL relationships takes place at the metalogical level.

4 Natural Language Reading of Diagrams

The introduced logical sentence forms are fundamental in ontological domain modelling, as evidenced by their coverage of sentences of the principal linguistic form:

(*all* | *some*) *common-noun verb* (*all* | *some*) *common-noun*

In particular, the versatile ∀∃ relationship is read as "All *common-noun verb* some *common-noun*", as in the sample sentences (with implicit determiners)

dog-owners own dog (for *all dog-owners own some dog*)

cat isa pet

dog isa pet

where the latter ones are understood as class inclusion derivatives from ∀∃. It follows implicitly that cats and dogs are disjoint by absence of a common subclass. By contrast, there would presumably be a joint subclass for cat-owner and dog-owner, cf. Fig. 3.

The actual extension sets of classes are considered to be of no ontological concern. However, distinguished individuals may be lifted to become additional singleton classes.

In a bio-medical ontology one might have sample ∀∃ relationships

beta-cell produces insulin

pancreas has-part beta-cell

beta-cell isa cell

In spite of their apparent quantificational complexity in predicate logic, the four considered CRL relationship forms possess simple natural stylized language forms as stated in Fig. 6. Determiners that may be omitted without ambiguity arising appear in brackets.

These stylized forms supplement the CRL diagrams for users who are unfamiliar with the underlying predicate logical explication.

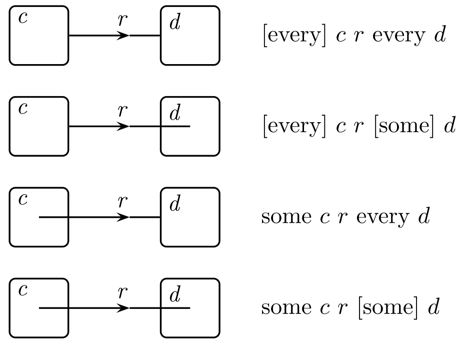
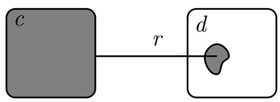


Fig. 6 The four language forms



$$\{x|c(x)\} \subseteq \{x|\exists y(r(x,y) \land d(y))\}$$

Fig. 7 Relationship between sets

5 Class-Relationship Diagrams

This section presents CRL diagrams as a formal system. A CRL diagram comprises a finite number of boxes, $\mathcal{B}$ (contours), depicting classes, and a finite number of arcs (directed edges connecting boxes), $\mathcal{A}$, depicting class-class relationships.

Figure 7 shows the rationale for the CRL diagram in Fig. 1 by a set-oriented, topological rendition with inverse image formation. Figure 7 is not itself a CRL diagram.

It should be observed that presence of additional CRL sentences may r-relate elements in c also to elements outside d. Thus, the present sentence would act as a constraint in this respect only via appeal to the closed world assumption, unlike the case for constraint diagrams [14], see further Sect. 9.

5.1 Boxes

Boxes are uniquely labelled with their class name. The name may be omitted in the rendition if irrelevant in the considered diagram.

Fundamentally, topological inclusion of box c in box d in a diagram expresses the subclass (class inclusion) relationship corresponding to the proposition $\forall x(c(x) \rightarrow \exists y(x = y \land d(y)))$, that is $\forall x(c(x) \rightarrow d(x))$. Thus boxes resemble contours in Euler diagrams. Accordingly, any pair of boxes possesses the topological properties of being disjoint, or partially overlapping, or one included in the other one.

However, a topological/geometrical overlap between two boxes unlike the contours of Euler diagrams is not a box (class) itself in CRL. Therefore such an overlap is assumed to embrace one or more boxes for classes belonging to the class overlap. Dually, the union of the boxes does not itself form a box. This is in accord with the ontological tenet that cross categories are acceptable, unlike would-be universals formed by disjunction or complementation, cf. [2], as discussed also in [7]. As such the diagram use differs from common use of Euler-Venn diagrams; for the latter see e.g. [19].

Fig. 8 Class inclusion

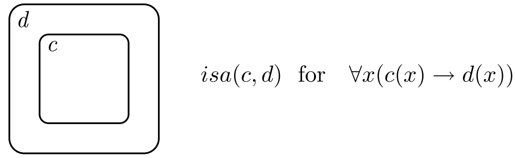


Fig. 9 Nested subclasses

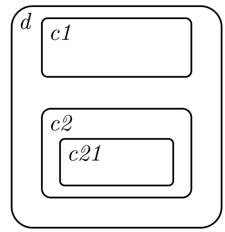
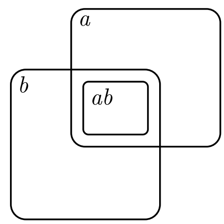


Fig. 10 Partial class overlap



5.2 The Case of Class Inclusion

The case of class inclusion (corresponding to r being identity in $\forall\exists(c, r, d)$, written $isa(c, d)$) obtains as in Euler diagrams, cf. Fig. 8.

As stated, by default convention classes are disjoint unless they have a common named subclass, all classes being non-empty.

5.3 CRL-Diagrams for Hierarchies and Trans-hierarchies

Boxes are by convention depicted in "menu style" (with sub-menus) as it appears in Fig. 9.

The underlying inclusion relation, *isa*, is a partial order. However it is not a lattice since the union and intersection (corresponding to lattice supremum and infimum) are not CRL classes in general. For intersecting classes there must be at least one common, named subclass in any overlap, cf. Fig. 10, since by convention, as mentioned, disjointness is assumed unless there is a joint subclass.

5.4 Relationship Arcs

Relationship arcs are directed edges (arcs) stretching from boxes to boxes or $\exists$-points. They are (not necessarily uniquely) labelled with the pertinent relation, so that a relation

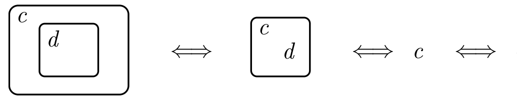


Fig. 11 Zooming

r may label any number of arcs. An arc labelled r from a box or ∃-point to a box or ∃-point represents a class-class relationship of one of the 4 basic forms. There may be any number of arcs between a pair of boxes. The logical type of a relationship is determined by the drawing conventions in Fig. 5. These conventions serve to facilitate diagram reasoning.

An ∃-point takes the name of its embracing box. It only serves in relationships with ∃-quantifiers. As such it represents some (one or more) unspecified member(s) of its class. The quantifier type of a relationship arc follows the conventions in Fig. 5, where an existential quantifier is indicated by an ∃-end point within a box.

5.5 CRL-Diagram Plasticity and Dynamics

CRL diagrams may be made more useful for large scale ontologies with many classes by provision of diagram visualization dynamics. This is achieved by metaphorically considering CRL diagrams as logical counterparts of road maps—stretching outside the screen. Accordingly, the focal point can be moved left and right and up and down in order to visualize relevant diagram parts, supplemented with zooming in and out.

The zooming facility implies that all boxes in a diagram are made simultaneously expandable/collapsible for browsing purposes, cf. Fig. 11. A box may be opened/blown-up to reveal inner boxes with concomitant arcs. Conversely, a box may be shrunk to just its name—or nothing. Relation arcs disappear according to appropriate conventions when boxes become diminished (cf. the road map principle). The dynamics must comply with the diagrammatic inference principles, which in turn reflect the admissible logical inferences.

6 Formal Logical Reasoning with CRL Diagrams

Diagrams have logical counterparts taking form of a finite collection of predicate logical ground factual sentences referred to as the metalogical forms $isa(\ ,\ )$ and $Q_1 Q_2(\ ,\ ,\ )$. In this section we describe how reasoning with the 4 basic class relationship sentences comprising 2 quantifiers can be conducted in predicate logic using these metalogical forms. This metalogical approach is inspired by our proposal for combinatory logic programming in [16] pursued in [7, 10]. This section serve to specify logically the diagrammatic reasoning rules introduced in Sect. 3.

All classes are assumed non-empty: $\exists x c(x)$ for all $c \in \mathcal{C}$. This non-emptiness assumption serves to streamline the reasoning rules, and to make provision for class overlap by introduction of a common subclass. It has no other bearing within CRL since classes are conceived of intensionally in their relationships to other classes, so that their extension is irrelevant in the ontological perspective.

6.1 Formalizing the Diagrammatic Ipso Facto Reasoning

As explained informally in Sect. 3, Fig. 4 tells that the shown r-relationship holds also for subclasses of the left class c, as well as for superclasses of the right class d. This conforms with a metalogic level inference rule

$$\frac{isa(c', c)\ \forall\exists(c, r, d)\ isa(d, d')}{\forall\exists(c', r, d')}$$

reflecting an object level rule

$$\frac{\forall x(c'(x) \rightarrow c(x))\ \forall x(c(x) \rightarrow \exists y(d(y) \wedge r(x, y)))\ \forall x(d(x) \rightarrow d'(x))}{\forall x(c'(x) \rightarrow \exists y(d'(y) \wedge r(x, y)))}$$

The diagram supports this inference in that an arc fixed at the contour of a box by convention can be extended to contours of all inner boxes. Moreover, it can be extended to a new ∃-point of inner boxes according to an inference rule of weakening

$$\frac{\forall\exists(c, r, d)}{\exists\exists(c, r, d)}$$

A relationship arc to or from an ∃-point by the very nature of the diagram function also as ∃-point for embracing boxes.

The diagrammatic rules for the remainder relationships follow the same principles as for ∀∃ relationships:

- ∀-end points (inlets and outlets) extend to the inner boxes (∃-points).
- ∃-end points (inlets and outlets) appear as such also to all embracing boxes.

6.2 Diagrammatic Reasoning at the Meta-logic Level

The inference rules needed for CRL are themselves formalized in predicate logic. Accordingly, there are two levels of logic: The CRL level and a metalogic level of inference rules. The inference rules are here expressed in the definite clause syntactic subset of function free predicate logic known as DATALOG, see e.g. [15]. Accordingly, DATALOG are definite clauses devoid of compound terms

$$p_0(t_{01}, \ldots, t_{0n_0}) \leftarrow p_1(t_{11}, \ldots, t_{1n_1}) \wedge \cdots \wedge p_m(t_{m1}, \ldots, t_{mn_m})$$

where the predicate argument terms t_{ij} are either constants or variables. All variables are implicitly ∀-quantified with prefixed quantifiers.

In a few cases resort may be taken to DATALOG extended with non-monotonic negation as failure to prove, DATALOG$^{\nvdash}$, cf. [15], e.g. for confirming absence of a common subclass of two classes.

6.3 Inference Rules Stated at the Meta-logic Level

In the meta-logic a relationship $\forall\exists(c, r, d)$ is conceived of as an atomic formula with a predicate symbol ∀∃. Similarly for the other relationships. The inference rules pertaining to CRL in the metalogic are re-shaped as definite clauses. Upper case indicates universally quantified variables.

For the inclusion relation between classes further we stipulate reflexivity and transitivity

$$isa(X, X)$$
$$isa(X, Z) \leftarrow isa(X, Y) \wedge isa(Y, Z)$$

Both way inclusion for a pair of classes is ruled out in diagrams for topological reasons. Accordingly, the *isa* relation forms a partial order, it being reflexive, anti-symmetric, and transitive.

The following inheritance, generalization, and weakening rules from Sect. 3 express the principle that classes can be specialized to subclasses for $\forall$, and generalized to superior classes for $\exists$.

Relationship inheritance

$$\forall\exists(C', R, D) \leftarrow \forall\exists(C, R, D) \wedge isa(C', C)$$
$$\exists\forall(C, R, D') \leftarrow \exists\forall(C, R, D) \wedge isa(D', D)$$
$$\forall\forall(C', R, D) \leftarrow \forall\forall(C, R, D) \wedge isa(C', C)$$
$$\forall\forall(C, R, D') \leftarrow \forall\forall(C, R, D) \wedge isa(D', D)$$

Relationship generalization:

$$\forall\exists(C, R, D') \leftarrow \forall\exists(C, R, D) \wedge isa(D, D')$$
$$\exists\exists(C', R, D) \leftarrow \exists\exists(C, R, D) \wedge isa(C, C')$$
$$\exists\exists(C, R, D') \leftarrow \exists\exists(C, R, D) \wedge isa(D, D')$$
$$\exists\forall(C', R, D) \leftarrow \exists\forall(C, R, D) \wedge isa(C, C')$$

The devised weakening rules further express that $\forall$ can be weakened to $\exists$, the classes being non-empty.

Weakening of quantifier:

$$\forall\exists(X, R, Y) \leftarrow \forall\forall(X, R, Y)$$
$$\exists\forall(X, R, Y) \leftarrow \forall\forall(X, R, Y)$$
$$\exists\exists(X, R, Y) \leftarrow \forall\exists(X, R, Y)$$
$$\exists\exists(X, R, Y) \leftarrow \exists\forall(X, R, Y)$$

We notice in passing that the above inference patterns reflect the notion of monotonicity in natural logic discussed in [28].

For cases of explicitly introduced inverse relations, introduced, say, by $inv(r, r')$, we posit:

$$\forall\forall(D, R', C) \leftarrow \forall\forall(C, R, D) \wedge inv(R, R')$$
$$\exists\exists(D, R', C) \leftarrow \exists\exists(C, R, D) \wedge inv(R, R')$$
$$\forall\exists(D, R', C) \leftarrow \exists\forall(C, R, D) \wedge inv(R, R')$$

These rules for inverses are not naturally reflected in CRL diagram rules, but calls for *ad hoc* diagram conventions for adding extra arcs for introduced inverses.

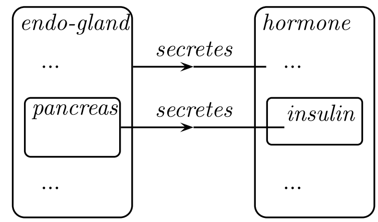


Fig. 12 Example from bio-ontology

Class overlap:

$$overlap(Y, Z) \leftarrow isa(X, Y) \wedge isa(X, Z)$$

Conversely, disjointness of classes in CRL diagrams is expressed by absence of any common subclass. Logically this is achieved by appealing to the closed world principle

$$disjoint(Y, Z) \leftarrow \nvdash overlap(Y, Z)$$

where $\nvdash$ represents negation as failure to prove. The formulation within DATALOG$^{\nvdash}$ offers the advantage that computation with the inference rules can be emulated as querying of a relational database where the CRL metalogic representations as ground atomic facts are stored, cf. [9].

The DATALOG fragment of predicate logic falls within the Bernays-Schönfinkel subclass and is thus effectively propositional and hence decidable. The propositional form is achievable by systematically instantiating the variables in the definite clauses with the available finite set of constants in $\mathcal{C} \cup \mathcal{R}$, and then re-conceive of the resulting ground atomic formulae as propositional symbols.

6.4 Example: Inheritance of Ascribed Properties

The tiny example (Fig. 12) from a bio-ontology is represented in the underlying logic as

$$isa(pancreas, endogland)$$

$$isa(insulin, hormone)$$

$$\forall\exists(endogland, secretes, hormone)$$

$$\forall\exists(pancreas, secretes, insulin)$$

As sample deducibles there are

$$\forall\exists(pancreas, secretes, hormone)$$

obtained from the above inference rules either by property inheritance or property generalization, and

$$\exists\exists(endogland, secretes, insulin)$$

obtained by weakening of a universal quantifier.

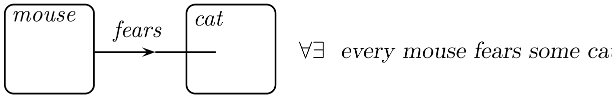


Fig. 13 Diagram for sample active voice

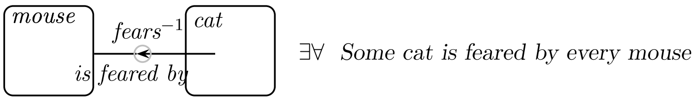
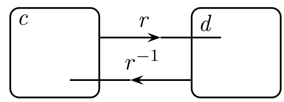


Fig. 14 Diagram for sample passive voice

Fig. 15 A tight relationship



7 Inverse Relations and Reciprocal Relationships

This section discusses some common CRL diagram patterns.

A binary relation r over individuals is logically bound to come with an inverse relation r^{-1}. By contrast inverse *class* relationships, called reciprocals, may or may not be present.

As an example compare the active voice form sentence with the corresponding passive voice sentence form in the pair of diagrams shown in Figs. 13 and 14. It is easy to verify by model checking that these sentences are not equivalent logically. Also, Skolemization yields different logical sentences.

7.1 Reciprocals and Tight Relationships

As mentioned, inverse *class* relationships (reciprocals) are usually absent. The case of co-presence of $\forall\exists(c, r, d)$ and $\forall\exists(d, r^{-1}, c)$ for r yields the diagram in Fig. 15. Such a configuration we dub a tight relationship.

As an example one may posit Fig. 16. But not the reciprocal Fig. 17.

The reciprocal configuration obtains, however, with Fig. 18.

Reciprocals play a special role for partonomic relationships where a pair of relationships, say, *partfor* and *haspart* arising from a binary relation *part* and its inverse, may form a tight partonomic relationship, cf. [26].

8 Analytic *Versus* Empirical Relationships

For a given class c in a CRL ontology diagram the outlet relationships $\forall\exists(c, r, d)$ (whether given explicitly or being deducible) for the various r and d constitute the attributed properties of c, cf. [7, 8]. A class c property ascription $\forall\exists(c, r, d)$ is said to constitute a specialization of $\forall\exists(c, r, d')$ when $isa(d, d')$.

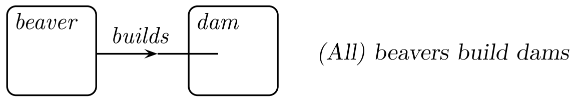


Fig. 16 Sample diagram for active voice

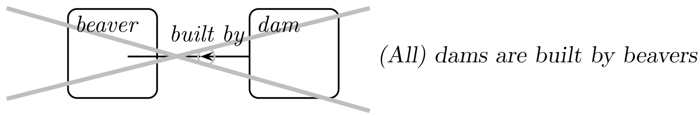


Fig. 17 The reciprocal of Fig. 16 does not hold

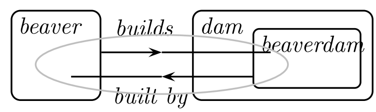


Fig. 18 Tight relationship by added subclass

Consider a class c' possessing all of the attributed properties of a class c albeit possibly in specialized form. This does not logically entail *per se* that c' is a subclass of c, $isa(c', c)$. However, if the entire bundle of attributed properties of a class c in the form of outlet relationships is taken as (if as well as only if) definition, then c' is bound to be a subclass of c in what may be termed intensional inclusion. What is at stake here is the distinction between analytic (definitional) and synthetic (i.e. empirical, observational) propositions, a distinction coming about here by way of an added if-definition with the bundle of attributed properties as premise.

Logically this principle of analyticity, if wished, might be achieved with an inference rule of analyticity stated with a class c as pivot and c' as candidate subclass.

As an example, let “pet owner” be taken to be an analytical concept defined as humans owning a pet animal, and similarly for “dog owner”. Given that dogs are pets, then “dog owner” would have to be a subclass of “pet owner” as in the diagram in Fig. 2.

Unlike the other inference rules, such an analyticity rule is not (at least not *prima facie*) expressible as a definite clause rule. Moreover, this rule contributes to the *isa* class subsumption relationships, thereby introducing a cyclic inference rule dependency between *isa* and the other relationships potentially initiating an inferential regression. We therefore suggest that an analyticity rule is only put at disposal as an option to be imposed on CRL diagrams for relationships that are tagged as contributing to the intensional definition of a class.

9 Similar Logics and Diagram Proposals

Numerous proposals have been put forward for diagrammatic representations of knowledge and reasoning. A number of these are discussed and compared in [27]. In the history of logic already Peirce’s existential graphs, see [24] for a comprehensive, contemporary exposition, introduce means of quantification as in predicate logic.

9.1 Proposals Applying Euler Diagrams

Many diagram proposals use as basis Euler diagrams, see e.g. [19, 20], which provides diagrammatic means of representing set inclusion, intersection, and exclusion topologically in an intuitive manner. Euler diagrams correspond logically to finite Boolean algebras: Given two sets as regions, their derivatives in the form of set union, intersection, and relative complement are also granted existence (albeit possibly as empty sets), at least in principle. This set-theoretic view (extending to a model-theoretic view in logic) is referred to as "Booleanism" in [25]. The Boolean view tends to introduce a multitude of sets which are irrelevant from an ontological point of view. The set-theoretic Boolean view is therefore commonly rejected in the context of formal ontologies in favour of a lattice oriented view, or, even sparser, a partial ordering with the inclusion relation. The favoured diagrams are then the Hasse diagrams known from lattice theory. The discarding of union and complement formation is bound up with the vexed question of negative properties in ontologies, see further [2], and [21] for a contemporary discussion.

In [13] the so called conceptual spaces are spanned by quality dimensions yielding a form of n-dimensional Venn diagrams. It is argued that natural kinds correspond to contiguous or even convex regions in such spaces. An algebraic logic for performing symbolic reasoning with such spaces is proposed in [6].

9.2 Euler Diagrams Extended with Relations

Let us now focus on some diagram proposals related to the CRL diagrams discussed here.

Logics similar to CRL with supporting diagram forms have been studied by Hammer in [17] and by Barwise and Hammer in [3]. However, [17] studies $\forall x(c(x) \to \forall y(d(y) \to r(x, y)))$ (corresponding to our $\forall\forall$) rather than the ontologically prevailing $\forall\exists$-form, and [3] considers $\exists x(c(x) \wedge \forall y(d(y) \to r(x, y)))$ (corresponding to our $\exists\forall$).

Hammer proposes diagrammatic reasoning rules for the considered $\forall\forall$ relationships. Hammer accepts configurations where a region (curve) is included in the set union, only, of two regions. This is not possible in CRL diagrams, since they convey inclusion, only, for given classes. As such CRL diagrams correspond to a partial ordering of given classes like Hasse diagrams.

In CRL Euler diagrams may be viewed as used merely for presenting certain predicate logical sentences in a convenient manner rather than representing sets topologically. Thus a logical sentence $\forall x(c(x) \to d(x))$ is depicted in CRL by inclusion of box named c in d. However, overlap of boxes a and b does not represent a sentence $\exists x(a(x) \wedge b(x))$, since the set elements are not recognized as such in CRL. This conforms with an intensional conception of the classes, where classes are characterized by the properties possessed jointly by the member individuals, rather than by enumeration of the individuals. Therefore, overlap of classes is explained in a diagram by presence of a common subclass, that is as presence of a box included in both boxes, together with the omnipresent assumption that classes are non-empty. Absence of any common subclass (witness common included box) is taken to mean that the two classes are disjoint, appealing to the principle of closed world assumption (CWA). The absence of explicit or entailed sentences expressing denials means that CRL has no means of expressing or generating inconsistency, unlike constraint diagrams.

9.3 Spider and Constraint Diagrams

Spider diagrams, [5], are Venn diagrams extended with diagrammatic facilities for introducing named or anonymous individuals. The diagrams enable expression of presence of a certain named individual in the various regions by means of so-called spiders. For instance it is possible to express presence of a certain constant within one or more of the sets described by the Venn diagram.

Constraint diagrams, [11, 14, 22], extend Spider diagrams by introducing (many-many) relations between the sets in a spider diagram. The relations are rendered as arrows with end points at the contours.

Constraint diagrams as presented in [14, 22] also appeal to Euler diagrams extended with relationship arcs. In constraint diagrams as presented in [14, 22] a relation f from a contour A to a contour B, written as $A.f = B$, is understood as follows

$$\forall x\bigl(A(x) \rightarrow \exists y\bigl(f(x,y) \wedge B(y)\bigr)\bigr)$$
$$\forall x\bigl(A(x) \rightarrow \neg\exists y\bigl(f(x,y) \wedge \neg B(y)\bigr)\bigr)$$

The former sentence corresponds to a CRL ∀∃ relationship from A to B. The latter sentence stipulates that no individual x in set A is f-related to something outside the set B. As such it expresses a constraint, which might be violated in a constraint diagram, with ensuing logical inconsistency. This is in contrast to CRL, where only the former sentence is expressed, with no means of achieving inconsistency.

Moreover, $A.f = B$ requires that

$$\forall y\bigl(B(y) \rightarrow \exists x\bigl(f(x,y) \wedge A(x)\bigr)\bigr)$$

corresponding to a reverse CRL arc.

Constraint diagrams bear resemblance to CRL diagrams when used for ontology specification. However, CRL diagrams differ in the available repertoire of logical class relationship forms as well as in the diagrammatic inference principles available in CRL. Fundamentally, constraint diagrams serve to express constraints by way of the logical inconsistency potential, whereas CRL logic and diagrams express only positive (ontological) knowledge, precluding thereby *per se* the risk or possibility of inconsistency.

9.4 Relation to Description Logic

The ∀∃-form $\forall\exists(c, r, d)$ has the counterpart sentence $c \sqsubseteq \exists r.d$ in description logic (DL), cf. e.g. [4, 15] for logico-algebraic and predicate logical accounts of DL.

More generally DL offers a fragment of FOL comprising classical negation supported by consistency (satisfiability) check facilities. DL comprise operators for forming new classes corresponding to the boolean operators in FOL. One may notice that the CRL relationship $\forall\forall(c, r, d)$ is different from the DL sentence $c \sqsubseteq \forall r.d$, the latter being $\forall x(c(x) \rightarrow \forall y(r(x, y) \rightarrow d(y)))$.

DL is biased towards a set-theoretic, extensional understanding of classes supported by the usual set operations on classes as sets. By contrast CRL supports a graph conception of class relationships in the vein of semantic networks. As such CRL is biased towards an intensional understanding of classes as objects possessing and inheriting properties via class-class relationships. Furthermore the CRL metalogic level supports classes and relationships as first class citizens which can be quantified over as it appears in the devised inference rules.

10 CRL Limitations and Extensions

In its present form CRL does not offer facilities for forming new classes by algebraic composition from the given ones. Nor does CRL offer facilities for composing relations to form chains between classes. However, a chain of relations might be provided visually in the diagram as highlighting of (shortest) connecting relational paths between classes. This may pertain to relationally homogeneous paths, say with transitive relations such as "causes" and "affects". This is achieved in principle with the following rule (specialized to homogeneous length 2 paths with relation R)

$$\forall\exists(C, R, R, D2) \leftarrow \forall\exists(C, R, C1) \wedge \forall\exists(C1, R, D2)$$

and more generally for homogeneous paths of length n:

$$\forall\exists(C, \underbrace{R, \ldots, R}_{n}, D_n) \leftarrow \forall\exists(C, R, C_1) \wedge \cdots \wedge \forall\exists(C_{n-1}, R, D_n)$$

being straightforwardly generalizable to heterogeneous relational paths. It should be noticed that the path inference would draw implicitly on the relational inference rules involving the class inclusion relationship. Such a computing of shortest paths between given classes c and d calls for application of standard search algorithms.

11 Concluding Summary

We have devised a proposal for diagrammatic visualization and reasoning for a fragment of predicate logic comprising various forms of logical relationships between classes. This sub-language is aimed at ontological domain modelling where taxonomies are enriched with *ad hoc* relationships between classes such as partonomic and causal relationships. The proposal is further distinguished by its dynamic appealing to the road map metaphor combined with intuitive visual reasoning principles, including the mentioned logical path finding. A prototype system with dynamic use of screen zooming has been developed in order to assess whether CRL offers a feasible compromise between on one hand logic diagram expressivity and on the other hand decidability and computational tractability.

Acknowledgements Many thanks to Sun-Joo Shin and Bartlomiej Szymczak, and to the anonymous reviewers for helpful comments and suggestions.

References

1. Allwein, G., Barwise, J.: Logical Reasoning with Diagrams. Oxford University Press, London (1996)

2. Armstrong, D.: A Theory of Universals. Cambridge University Press, Cambridge (1978)
3. Barwise, J., Hammer, E.: Diagrams and the concept of a logical system. In: [1]
4. Brink, C., et al.: Peirce algebras. Form. Asp. Comput. **6**(3), 339–358 (1994)
5. Dau, F., Fish, A.: Conceptual spider diagrams. In: Proceedings of the 16th International Conference on Conceptual Structures. Lecture Notes in Computer Science, vol. 5113, pp. 104–118. Springer, Berlin (2008)
6. Fischer Nilsson, J.: A conceptual space logic. In: Kawaguchi, E., et al. (eds.) Information Modelling and Knowledge Bases XI. 9th European-Japanese Conferences on Information Modelling and Knowledge Bases, Iwate, Japan, May 24–28, 1999, pp. 26–40. IOS Press, Amsterdam (2000)
7. Fischer Nilsson, J.: Ontological constitutions for classes and properties. In: Int. Conference on Conceptual Structures. Lecture Notes in Computer Science, vol. 4068, pp. 37–53. Springer, Berlin (2006)
8. Fischer Nilsson, J.: On reducing relationships to property ascriptions. In: Kiyoki, Y., et al. (eds.) Information Modelling and Knowledge Bases XX. Frontiers in Artificial Intelligence and Applications, vol. 190, pp. 245–252. IOS Press, Amsterdam (2008)
9. Fischer Nilsson, J.: Querying class-relationship logic in a metalogic framework. In: Flexible Query Answering Systems FQAS 2011. Lecture Notes in Computer Science, vol. 7022 (2011)
10. Fischer Nilsson, J., Palomäki, J.: Towards computing with intensions and extensions of concepts. In: Charrel, P.-J., et al. (eds.) Information Modelling and Knowledge Bases IX, pp. 100–114. IOS Press, Amsterdam (1998)
11. Fish, A., Flower, J., Howse, J.: The semantics of augmented constraint diagrams. J. Vis. Lang. Comput. **16**(6), 541–573 (2005)
12. Gabbay, D.M., Woods, J. (eds.): Handbook of the History of Logic, vol. 3, The Rise of Modern Logic: From Leibniz to Frege. Elsevier, Amsterdam (2004)
13. Gärdenfors, P.: Conceptual Spaces: On the Geometry of Thought. MIT Press, Cambridge (2000)
14. Gil, J., Howse, J., Kent, S.: Towards a formalization of constraint diagrams. In: Proceedings of the IEEE 2001 Symposia on Human Centric Computing Languages and Environments (HCC'01) (2001)
15. Grosof, B.N., Horrocks, I., Volz, R., Decker, S.: Description logic programs: combining logic programs with description logic. In: Proceedings of the Twelfth International World Wide Web Conference, WWW2003, Budapest, Hungary, 2003, pp. 48–57. ACM, New York (2003)
16. Hamfelt, A., Fischer Nilsson, J.: Towards a logic programming methodology based on higher-order predicates. New Gener. Comput. **15**(4), 421–448 (1997)
17. Hammer, E.M.: Logic and Visual Information. CSLI, Stanford (1995)
18. Harel, D.: On visual formalisms. Commun. ACM **31**(5), 514–530 (1988)
19. Howse, J.: Diagrammatic reasoning systems. In: Proceedings of the 16th International Conference on Conceptual Structures: Knowledge Visualization and Reasoning. Lecture Notes in Artificial Intelligence, vol. 5113, pp. 1–20 (2008)
20. Howse, J., et al.: Euler diagram-based notations. University of Brighton and University of Kent, UK
21. Merrill, G.H.: Ontological realism: methodology or misdirection? Appl. Ontol. **5**(2), 79–108 (2010)
22. Oliver, I., Howse, J., Stapleton, G., Nuutila, E., Törmä, S.: A proposed diagrammatic logic for ontology specification and visualization. In: 8th International Semantic Web Conference (Posters and Demos) (2009)
23. Sánchez Valencia, V.: The algebra of logic. In: [12]
24. Shin, S.-J.: The Iconic Logic of Peirce's Graphs. MIT Press, Cambridge (2002)
25. Smith, B.: Against fantalogy. In: Reicher, M.E., Marek, J.C. (eds.) Experience and Analysis, pp. 153–170 (2005)
26. Smith, B., Rosse, C.: The Role of Foundational Relations in the Alignment of Biomedical Ontologies, MEDINFO 2004, pp. 444–448. IOS Press, Amsterdam (2004)
27. Sowa, J.: Knowledge Representation: Logical, Philosophical and Computational Foundations. Brooks Cole, Pacific Grove (2000)
28. van Benthem, J.: Essays in Logical Semantics. Reidel, Dordrecht (1986)

J. Fischer Nilsson (✉)
DTU Informatics, Technical University of Denmark, Building 322, 2800 Lyngby, Denmark
e-mail: jfn@imm.dtu.dk

On the Completeness of Spider Diagrams Augmented with Constants

Gem Stapleton, John Howse, Simon Thompson, John Taylor, and Peter Chapman

Abstract Diagrammatic reasoning can be described formally by a number of diagrammatic logics; spider diagrams are one of these, and are used for expressing logical statements about set membership and containment. Here, existing work on spider diagrams is extended to include constant spiders that represent specific individuals. We give a formal syntax and semantics for the extended diagram language before introducing a collection of reasoning rules encapsulating logical equivalence and logical consequence. We prove that the resulting logic is sound, complete and decidable.

Keywords Spider diagrams · Constants · Soundness · Completeness · Monadic first-order logic · Diagrammatic reasoning

Mathematics Subject Classification (2010) Primary 68R02; Secondary 03B02

1 Introduction

Diagrams have been used for centuries in the visualization of mathematical concepts and to aid the exploration and formalization of ideas. This is not the place to survey that history; however, we give a brief overview of the background to the development of spider diagrams now.

One of the most successful visual notations is the Venn diagram for sets and their relationships; indeed, it is taught in the elementary school curriculum in many countries. While Venn diagrams contain all possible intersection regions between the sets, Euler diagrams [4] allow set intersection, disjointness and containment to be represented visually. The Euler diagram d_1 in Fig. 1 asserts that A and B are disjoint and C is a subset of A. The relative placement of the curves gives, for free, that C is disjoint from B. This 'free ride' is one of the areas where diagrams are thought to be superior to symbolic languages [20]. This example also illustrates the concept of 'well-matchedness' [8] since the visual representation of assertions mirrors those at the semantic level: for example, the containment of one curve by another mirrors the interpretation that the enclosed curve, C, represents a subset of the set represented by the enclosing curve, A. Moreover, this has the added benefit that the subset relation is mirrored by the transitive property of syntactic containment.

Various extensions to Euler diagrams have been proposed, such as including syntax to represent named individuals [27], or assert the existence of arbitrary finite numbers of elements [12]. The Euler diagram d_2 in Fig. 1 is augmented with shading, which asserts

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic,
DOI 10.1007/978-3-0348-0600-8_7,

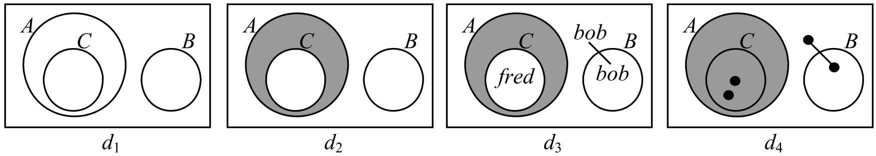


Fig. 1 Extended Euler diagrams

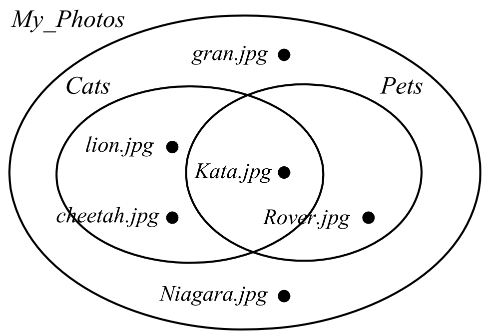


Fig. 2 Non-hierarchical file systems

the emptiness of the set $A - C$ and the Euler/Venn diagram d_3 tells us, in addition, that *fred* is in the set C and *bob* is not in the set A.

Spider diagrams [12] are also based on Euler diagrams. The spider diagram d_4 in Fig. 1 asserts the existence of two elements in the set C and at least one element outside of the set A; this is accomplished through the use of *existential spiders*. A spider is a tree which denotes a single element that can occupy one of the positions given by the nodes of the tree. The shading in d_4 is used to place an upper bound on the cardinality of A, limiting it to two: in a set represented by a shaded region, all elements must be denoted by spiders. Using a model-theoretic argument, it has been shown that spider diagrams are equivalent to Monadic First-Order Logic with equality [23].

Constant spiders [21, 25], corresponding to given spiders in [11], were introduced to provide users of spider diagrams with an explicit way to write constraints involving named individuals. There are a number of examples of spider diagrams being used in practice, such as assisting with the task of identifying component failures in safety critical hardware designs [2]. Equivalent notations have been used for representing non-hierarchical computer file systems [3], in a visual semantic web editing environment [16, 28] and for viewing clusters which contain concepts from multiple ontologies [9]. Each of these applications uses constants to represent specific objects, thus motivating the utility of augmenting spider diagrams with constants. To take a particular example, the VennFS system [3], is used to represent visually non-hierarchical files systems. The example in Fig. 2 provides information about the folder location of certain files stored on a computer: the labeled dots are files—or constant spiders—and the curves represent folders.

In [25], it was established that constants in spider diagrams could be simulated by a shaded contour containing a single (non-constant) spider. This translation gave a diagram that was expressively equivalent to the original, in the sense that it had the same model set as the spider diagram with a constant. As with many notations—both symbolic and diagrammatic—it is worthwhile adding a notation even though it might be dismissed as mere 'syntactic sugar'. The additional notation makes clear the *intention* of the user, and allows that intention to be preserved in reasoning, for instance. In a visual notation it makes it much easier to preserve the 'free ride' and 'well matchedness' properties; in the particular case of constants there is a direct naming of a constant, rather than an indirect naming through the name of the representing contour, for instance. Further discussion and motivation can be found in [21, 25].

Earlier work formalized the syntax and semantics of spider diagrams and specified a logic for the diagrams which was proved to be sound, complete and decidable; in this paper we do the same for spider diagrams with constants. Specifically, in Sect. 2, we give the syntax of spider diagrams extended to include constant spiders and, in Sect. 3, present formal semantics. In Sect. 4, we provide a collection of reasoning rules for spider diagrams with constants and, in Sect. 5, we present sketches of soundness, completeness and decidability results.

2 Syntax

In diagrammatic systems, we can distinguish two levels of syntax: concrete (or token) syntax and abstract (or type) syntax [10]. Concrete syntax captures the physical representation of a diagram. Abstract syntax is independent of the semantically unimportant spatial relations between syntactic elements in a concrete diagram. We do not include the concrete syntax in this discussion since we work at the abstract level here.

The closed curves in a spider diagram are called *contours* and each contour is identified by a label chosen from a countably infinite set, $\mathcal{CL}$. A *zone*[1] is defined to be a pair (in, out) of disjoint finite subsets of $\mathcal{CL}$. The set *in* contains the labels of the contours that include the zone (in, out) whereas *out* is the set of labels of the contours that do not include (in, out). So, in a unitary diagram, *in* and *out* form a partition of the contour label set. In diagram d_1 in Fig. 3 the zone that is inside contour A but outside B and C has abstract representation $(\{A\}, \{B, C\})$. A *region* is a set of zones. We define $\mathcal{Z}$ and $\mathcal{R} = \mathbb{P}\mathcal{Z}$ to be the sets of all zones and regions respectively. As noted earlier, in a Venn diagram, d, every possible zone—that is every element of $\mathbb{P}L$ for the set L of contour labels in d—is represented in d. This is not the case for spider diagram, and a zone is said to be *missing* if it is not a member of the possible zone set for the diagram.

A spider without a label is called an *existential spider*. A spider with a label is called a *constant spider*. A spider *touches* a zone if that zone is in its habitat, and a spider is said to *inhabit* the region in which it is placed, which is termed its *habitat*. To describe the existential spiders in a particular diagram, it is sufficient to say how many existential spiders there are in each region. We will use a bag of regions, called *existential spider descriptors*, with the number of occurrences of each region in the bag giving the number of existential spiders in the region. For example, the region

$$\big\{\big(\{A, C\}, \{B\}\big), \big(\emptyset, \{A, B, C\}\big), \big(\{B\}, \{A, C\}\big), \big(\{B, C\}, \{A\}\big)\big\}$$

in diagram d_2 in Fig. 3 contains two existential spiders. We must also specify which constant spider labels appear and, for each spider label, the habitat of the spider with that

[1] Since all constructs discussed here are abstract, we will use the terminology 'zone' rather than 'abstract zone' throughout.

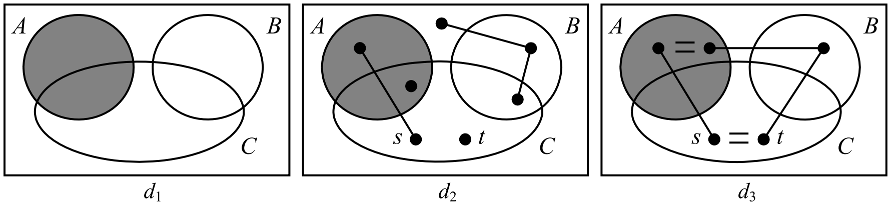


Fig. 3 Examples of unitary spider diagrams

label. At the abstract level, a unitary diagram will contain a finite set of constant spider labels together with a habitat function, mapping each constant spider label to a region in the diagram. The habitat of the constant spider labeled s in diagram d_2 in Fig. 3 is $\{(\{A\}, \{B, C\}), (\{C\}, \{A, B\})\}$.

We will assume that all of the constant spider labels come from a finite set $\mathcal{CS}$. An alternative choice would be to have a countably infinite set of constant spider labels. With this alternative choice, the work below on reasoning rules, soundness and completeness remains identical. However, the approach taken in [23] to prove that augmenting the spider diagram language with constants does not increase expressiveness would need to be modified.

Given two distinct constant spiders, each with a habitat sharing some zone z, a *tie*, represented by an 'equals' sign, can be placed between them in z. The *web* of a pair of constant spiders is the set of zones that contain a tie between those two spiders. The diagram d_3 in Fig. 3 contains two constant spiders, labeled s and t, connected by two ties. The web of s and t is the region made up of the zone inside contour A but outside B and C and the zone inside C but outside A and B.

General spider diagrams are a logical combination of diagrams; a single diagram is called *unitary*. The formal definition of an abstract unitary spider diagram with constants extends that given in [12] for unitary spider diagrams without constants. We assume that the sets $\mathcal{CS}$, $\mathcal{CL}$, $\mathcal{Z}$ and $\mathcal{R}$ are all pairwise disjoint.

Definition 2.1 An **abstract unitary spider diagram with constants**, d (with contour labels in $\mathcal{CL}$ and constant spider labels in $\mathcal{CS}$), is a 7-tuple

$$\langle L, Z, Z^*, ESD, CS, \theta, \omega\rangle$$

whose components are defined as follows.

1. $L = L(d) \subset \mathcal{CL}$ is a finite set of contour labels.
2. $Z = Z(d) \subseteq \{(in, L - in) : in \subseteq L\}$ is a set of zones such that
   (i) for each label $l \in L$ there is a zone $(in, L - in) \in Z(d)$ such that $l \in in$ and
   (ii) the zone $(\emptyset, L)$ is in $Z(d)$.
   We define $R(d) = \mathbb{P}Z - \{\emptyset\}$ to be the set of regions in d. We further define $MZ(d) = \{(in, L - in) : in \subseteq L\} - Z(d)$ to be the **missing zones** of d.
3. $Z^* = Z^*(d) \subseteq Z$ is a set of **shaded zones** and we define $R^*(d) = \mathbb{P}Z^*(d)$ to be the set of shaded regions in d. A region, $r \in R(d) - R^*(d)$, is **completely non-shaded** if and only if $r \cap Z^*(d) = \emptyset$.
4. $ESD = ESD(d) \subset \mathbb{Z}^+ \times R(d)$ is a finite set of **existential spider descriptors** such that

$$\forall (n_1, r_1), (n_2, r_2) \in ESD \; (r_1 = r_2 \Rightarrow n_1 = n_2).$$

If $(n, r) \in ESD$ we say there are n existential spiders with **habitat** r.

5. $CS = CS(d) \subseteq \mathcal{CS}$ is a finite set of constant spider labels.
6. $\theta = \theta_d : CS \to R(d)$, is a function which maps each constant spider label to a region in d. If $\theta_d(s_i) = r$ we say s_i has **habitat** r in d.
7. $\omega = \omega : CS(d) \times CS(d) \to \mathbb{P}Z$ is a function which returns the **web** of each pair of constant spiders where $z \in \omega(s_i, s_j)$ means that there is a **tie** between s_i and s_j in the zone z. Further, ω must ensure that the following hold for all s_i, s_j, s_k in $CS(d)$:
   (a) given two constant spiders there can only be ties in zones common to their habitat: $\omega(s_i, s_j) \subseteq \theta_d(s_i) \cap \theta_d(s_j)$,
   (b) each constant spider is joined by ties to itself (this simplifies the formalization of the semantics below): $\omega(s_i, s_i) = \theta_d(s_i)$,
   (c) if there is a tie between constant spiders s_i and s_j in zone z, then there is a tie between s_j and s_i in z: $\omega(s_i, s_j) = \omega(s_j, s_i)$, and
   (d) given any zone z, if s_i and s_j are joined by a tie in z and so are s_j and s_k, then s_i and s_k are joined by a tie in z: $z \in \omega(s_i, s_j) \cap \omega(s_j, s_k) \Rightarrow z \in \omega(s_i, s_k)$.

Some remarks about the above definition are in order, before we illustrate it with an example.

- Every contour in a concrete diagram contains at least one zone as captured by condition 2 (i).
- In any concrete diagram, the zone inside the boundary rectangle but outside all the contours is present and this is captured by condition 2 (ii).
- Being joined by a tie is interpreted transitively. In fact, ties give rise to an equivalence relation on the spiders in each zone, as specified by conditions 7 (b), (c) and (d).
- Therefore, in a zone z, taking the constant spiders in z as a set of vertices and the ties in that zone as a set of edges, we would have a graph whose components formed complete graphs with loops at each vertex. However, in our concrete syntax we will only draw a spanning forest in each zone so as to avoid unnecessary clutter in diagrams.
- We note that ties could also be used to connect existential spiders. Indeed, they could also be used to connect an existential spiders to constant spiders.[2]

Example The diagram d_1 in Fig. 4 has the following abstract description.

1. Contour label set $L(d_1) = \{A, B\}$.
2. Zone set

$$Z(d_1) = \big\{\big(\emptyset, \{A, B\}\big), \big(\{A\}, \{B\}\big), \big(\{B\}, \{A\}\big), \big(\{A, B\}, \emptyset\big)\big\}.$$

3. Shaded zone set $Z^*(d_1) = \{(\{B\}, \{A\})\}$.
4. The set of existential spider descriptors

$$ESD(d_1) = \big\{\big(1, \big\{\big(\{B\}, \{A\}\big)\big\}\big), \big(1, \big\{\big(\{A\}, \{B\}\big), \big(\{B\}, \{A\}\big)\big\}\big)\big\}.$$

5. Constant spider label set $CS(d_1) = \{s, t\}$.

[2] However, for any diagram that incorporated such ties it is possible to define a semantically equivalent diagram that does not contain such ties. This is not the case for ties between constant spiders. It is straightforward to extend the work in this paper to the case where these additional types of tie are permitted.

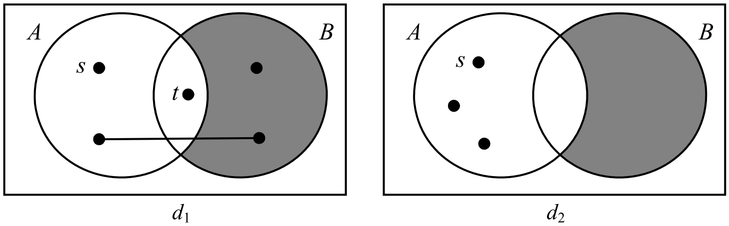


Fig. 4 Two spider diagrams with constants

6. The function $\theta_{d_1} : \{s,t\} \rightarrow R(d_1)$ where $\theta_{d_1}(s) = \{(\{A\},\{B\})\}$ and $\theta_{d_1}(t) = \{(\{A,B\},\emptyset)\}$.
7. The function $\omega_{d_1} : CS(d_1) \times CS(d_1) \rightarrow \mathbb{P}Z(d_1)$ where $\omega_{d_1}(s,s) = \theta_{d_1}(s)$, $\omega_{d_1}(t,t) = \theta_{d_1}(t)$ and $\omega_{d_1}(s,t) = \omega_{d_1}(t,s) = \emptyset$.

Now we introduce some terminology and notation on top of the concepts formalized in the definition. An existential spider descriptor (n, r) is intended to mean that there are precisely n existential spiders placed in the zones in the region r, and we can think of these being numbered from 1 to n. A typical such spider will be spider i, which we denote by $e_i(r)$, to avoid confusion with the notation (i, r) used for existential spider descriptors. The set of **existential spiders** in a unitary diagram d is given by

$$ES(d) = \left\{e_i(r) : \exists(n,r) \in ESD(d) \wedge 1 \leq i \leq n\right\}.$$

We also define $S(d) = ES(d) \cup CS(d)$ to be the set of **spiders** in d. We assume that the sets $ES(d)$ and $\mathcal{CS} \cup \mathcal{CL} \cup \mathcal{Z} \cup \mathcal{R}$ are disjoint. We also define a function

$$\eta : ES(d) \rightarrow R(d)$$

which returns the **habitat** of each existential spider, so that $\eta(e_i(r)) = r$.

Spiders represent the existence of elements and regions represent sets—thus we need to know how many elements are represented in each region. Note here that, in a unitary diagram, a constant spider and an existential spider represent the existence of distinct elements. For example, in Fig. 4, the diagram d_2 asserts that the set represented by the zone $(\{A\},\{B\})$ contains at least three elements, including the individual represented by s. The set of existential spiders contained by region r in d is denoted by $ES(r,d)$. More formally,

$$ES(r,d) = \left\{e \in ES(d) : \eta(e) \subseteq r\right\}.$$

Similarly, the set of constant spiders contained by region r in d is

$$CS(r,d) = \left\{s \in CS(d) : \theta_d(s) \subseteq r\right\}$$

and we also define

$$S(r,d) = ES(r,d) \cup CS(r,d).$$

So, any spider in d whose habitat is a subset of r is in the set $S(r,d)$. The set of existential spiders touching r in d is denoted by $ET(r,d)$. More formally,

$$ET(r,d) = \left\{s \in ES(d) : \eta(s) \cap r \neq \emptyset\right\}.$$

Moreover, in a shaded region there is an upper bound on the cardinality of the represented set. For example, d_1 in Fig. 4 tells us that there are at most two elements in $B - A$, because exactly two spiders touch $B - A$. The set of constant spiders touching a region, $CT(r, d)$, and the set of spiders touching a region, $T(r, d)$, are defined similarly. In d_1, Fig. 4,

$$\left|S\left(\left\{\left(\{B\}, \{A\}\right)\right\}, d_1\right)\right| = 1$$

and

$$\left|T\left(\left\{\left(\{B\}, \{A\}\right)\right\}, d_1\right)\right| = 2.$$

In d_2,

$$\left|S\left(\left\{\left(\{A\}, \{B\}\right)\right\}, d_2\right)\right| = \left|T\left(\left\{\left(\{A\}, \{B\}\right)\right\}, d_2\right)\right| = 3.$$

Unitary diagrams form the building blocks of *compound diagrams*, formed by using logical connectives.

Definition 2.2 An **abstract spider diagram with constants** is defined as follows.

1. Any unitary diagram with constants is a spider diagram with constants.
2. If D_1 and D_2 are spider diagrams with constants then $(D_1 \vee D_2)$ and $(D_1 \wedge D_2)$ are spider diagrams with constants.

Our convention will be to denote unitary diagrams by d and arbitrary diagrams by D. Some compound diagrams are not satisfiable (defined later). For convenience later, we introduce the symbol $\perp$, defined to be a unitary diagram that is not satisfiable.

3 Semantics

We now sketch, informally, the semantics of unitary spider diagrams. Regions represent sets. Missing zones represent the empty set. For example, in diagram d_1 in Fig. 3, the zones $(\{A, C\}, \{B\})$ and $(\{A\}, \{B, C\})$ are missing and so represent the empty set; from this we can deduce that sets represented by A and B are disjoint.

Now, for simplicity, suppose a unitary diagram d does not contain any ties. If region r is inhabited by n spiders in d then d expresses that the set represented by r contains at least n elements. If r is shaded and touched by m spiders in d then d expresses that the set represented by r contains at most m elements. Thus, if d has a shaded, untouched region, r, then d expresses that r represents the empty set. For example, in diagram d_1 in Fig. 3, the shaded region $\{(\{A\}, \{B, C\}), (\{A, C\}, \{B\})\}$ is untouched by any spider and therefore represents the empty set. In diagram d_2 in Fig. 3, the same region is shaded and touched by two spiders and so the set it represents contains at most two elements.

Each constant spider asserts that the individual it represents is in the set represented by its habitat. Moreover, the individuals represented by constant spiders are distinct from those represented by existential spiders. Therefore, if a region contains an existential spider and a constant spider, s, we can deduce that there are at least two elements in that region, including that represented by s. Within a unitary diagram, no two constant spiders represent the same individual unless they are joined by a tie. Constant spiders joined by ties represent the same individual if and only if there exists a zone, z, in their web and they both represent individuals in the set represented by z. So, the presence of a tie between two constant spiders has the effect of potentially reducing the upper and lower cardinality constraints placed on the set represented by the union of their habitats. In diagram d_3 in Fig. 3, the constant spiders s and t represent different individuals unless both the individuals they represent are in the set represented by the zone $(\{A\}, \{B, C\})$ or both are in the set represented by $(\{C\}, \{A, B\})$, in which case they must represent the same individual.

To formalize the semantics of spider diagrams with constants we shall map constant spider labels, contour labels, zones and regions to subsets of some universal set. We wish constant spider labels to act like constants in first-order predicate logic, so they will map to single element subsets of the universal set, unless the universal set is the empty set. We could, equivalently, choose to map constant spiders to elements of the universal set. However, the *semantics predicate* (defined below) is more elegant when we map constant spiders to sets, as are the details of some of the proofs below. Our formalization of the semantics extends that given for spider diagrams without constants in [12].

Definition 3.1 An **interpretation of constant spider labels, contour labels, zones and regions**, or simply an **interpretation**, is a pair (U, Ψ) where U is a set and $\Psi : \mathcal{CL} \cup \mathcal{Z} \cup \mathcal{R} \cup \mathcal{CS} \to \mathbb{P}U$ is a function mapping constant spider labels, contour labels, zones and regions to subsets of U such that the images of the zones and regions are completely determined by the images of the contour labels as follows:

1. for each zone (a, b), $\Psi(a, b) = \bigcap_{l \in a} \Psi(l) \cap \bigcap_{l \in b} \overline{\Psi(l)}$ where $\overline{\Psi(l)} = U - \Psi(l)$ and we define $\bigcap_{l \in \emptyset} \Psi(l) = U = \bigcap_{l \in \emptyset} \overline{\Psi(l)}$ and
2. for each region r, $\Psi(r) = \bigcup_{z \in r} \Psi(z)$ and we define $\Psi(\emptyset) = \bigcup_{z \in \emptyset} \Psi(z) = \emptyset$

and either the universal set is the empty set or the constant spiders map to singleton subsets of U. More formally

$$U = \emptyset \ \vee \ \forall s_i \in \mathcal{CS} \ \left|\Psi(s_i)\right| = 1.$$

We will write $\Psi : \mathcal{R} \cup \mathcal{CS} \to \mathbb{P}U$ when strictly speaking we mean $\Psi : \mathcal{CL} \cup \mathcal{Z} \cup \mathcal{R} \cup \mathcal{CS} \to \mathbb{P}U$.

We introduce a *semantics predicate* which identifies whether a diagram expresses a true statement, with respect to an interpretation.

Definition 3.2 Let D be a spider diagram with constants and let $m = (U, \Psi)$ be an interpretation. We define the **semantics predicate** of D, denoted $P_D(m)$. If $D = \perp$ then $P_D(m)$ is $\perp$. If D $(\neq \perp)$ is a unitary diagram then $P_D(m)$ is the conjunction of the following conditions.

1. **Plane Tiling Condition.** The union of the sets represented by the zones in D is the universal set: $\bigcup_{z \in Z(D)} \Psi(z) = U$.
2. There exists an extension of $\Psi : \mathcal{R} \cup \mathcal{CS} \to \mathbb{P}U$ to $\Psi : \mathcal{R} \cup \mathcal{CS} \cup ES(D) \to \mathbb{P}U$ such that the following conditions are satisfied.
   (a) **Spiders Condition.** Each spider represents the existence of an element (strictly, a single element set) in the set represented by its habitat and existential spiders do not represent the same elements as any constant spiders:

$$\forall s \in ES(D) \ \left(\left|\Psi(s)\right| = 1 \wedge \Psi(s) \subseteq \Psi\big(\eta(s)\big)\right)$$

and

$$\forall s \in CS(D)\ \big(\big|\Psi(s)\big| = 1 \wedge \Psi(s) \subseteq \Psi\big(\theta_D(s)\big)\big)$$

and

$$\forall e \in ES(D)\ \forall s_i \in CS(D)\ \Psi(e) \neq \Psi(s_i).$$

(b) **Existential Spiders Condition.** No two existential spiders represent the existence of the same element:

$$\forall e_1, e_2 \in ES(D)\ \big(\Psi(e_1) = \Psi(e_2) \Rightarrow e_1 = e_2\big).$$

That is, the function Ψ is injective when the domain is restricted to $ES(d)$.

(c) **Constant Spiders Condition.** Two constant spiders represent the same individual if and only if they both represent an individual in the set denoted by some zone in their web:

$$\forall s_i, s_j \in CS(D)\ \big(\Psi(s_i) = \Psi(s_j) \\ \Leftrightarrow\ \exists z \in \omega_D(s_i, s_j)\ \Psi(s_i) \cup \Psi(s_j) \subseteq \Psi(z)\big).$$

(d) **Shading Condition.** Each shaded zone, z, represents a subset of the set of elements represented by the spiders touching z:

$$\forall z \in Z^*(D)\ \Psi(z) \subseteq \bigcup_{s \in T(\{z\}, D)} \Psi(s).$$

If $\Psi : \mathcal{R} \cup ES(D) \rightarrow \mathbb{P}U$ ensures $P_D(m)$ is true then Ψ is a **valid extension to existential spiders** for D. If $D = D_1 \vee D_2$ then $P_D(m) = P_{D_1}(m) \vee P_{D_2}(m)$. If $D = D_1 \wedge D_2$ then $P_D(m) = P_{D_1}(m) \wedge P_{D_2}(m)$. We say m **satisfies** D, or m is a **model** for D, denoted $m \models D$, if and only if $P_D(m)$ is true. If all the models for D_1 are models for D_2, then D_1 **semantically entails** D_2, denoted $D_1 \vDash D_2$. If $D_1 \vDash D_2$ and $D_2 \vDash D_1$, then D_1 and D_2 are **semantically equivalent**, denoted $D_1 \equiv_{\vDash} D_2$.

As an example, the interpretation $m = (\{1, 2, 3, 4\}, \Psi)$ partially defined by $\Psi(s_1) = \{1\}$, $\Psi(s_2) = \{2\}$, $\Psi(L_1) = \{1, 2\}$ and $\Psi(L_2) = \{2, 3, 4\}$ is a model for d_1 in Fig. 4 but not for d_2.

Theorem 3.3 *Let d $(\neq \perp)$ be a unitary spider diagram with constants. Then d is satisfiable.*

The proof strategy is to construct an interpretation that we call a *standard model* for d, following a similar approach to that for spider diagrams without constants in [12]. Essentially, this contains only the elements that are forced to exist by the presence of spiders in the diagram: for each spider in the diagram we choose one the zones in its habitat and place an element there; in extending this construction to constants we just have to make sure that these elements are identified when ties require that to be so. It is straightforward to show that any standard model for d satisfies d. This standard model is also used in the proof of completeness. More formally, a standard model is defined as follows:

Definition 3.4 Let d be a unitary spider diagram with constants. Let $f : S(d) \to Z(d)$ be a function such that for each spider s, $f(s)$ is in the habitat of s. For each constant spider, s_i, we define

$$[s_i] = \{s_j \in CS(d) : f(s_j) = f(s_i) \wedge f(s_i) \subseteq \omega_d(s_i, s_j)\}$$

(these sets $[s_i]$ give rise to an equivalence relation and, hence, form a partition of $CS(d)$). Define

$$U = ES(d) \cup \{[s_i] : s_i \in CS(d)\}.$$

For each contour label, L, in d define

$$\begin{aligned}\Psi(L) = &\{e \in ES(d) : f(e) = (in, out) \wedge L \in in\} \\ &\cup \{[s_i] : s_i \in CS(d) \wedge f(s_i) = (in, out) \wedge L \in in\}\end{aligned}$$

and each constant spider, s_k, in d, maps to the set

$$\Psi(s_k) = \{[s_k]\}.$$

Then (U, Ψ) is a **standard model** for d.

4 Reasoning Rules

We will now develop a set of sound and complete reasoning rules for spider diagrams with constants. All of the reasoning rules given for spider diagrams without constants in [12] can be extended—sometimes in a non-trivial way—to spider diagrams with constants; we omit most of the formal definitions of the extended rules.

4.1 Unitary to Unitary Reasoning Rules

In this section we introduce a collection of reasoning rules that apply to, and result in, a unitary diagram.

Rule 1 (*Introduction of a shaded zone*) Let d_1 be a unitary diagram that has a missing zone. If d_2 is the same as d_1 except that d_2 contains a new, shaded and 'untouched' zone then d_1 is logically equivalent to d_2.

In Fig. 5, Rule 1 (introduction of a shaded zone) is applied to d_1 to give d_2. Applying the introduction of a shaded zone rule results in a semantically equivalent diagram. The next two rules are not information preserving.

Rule 2 (*Erasure of shading*) Let d_1 be a unitary diagram with a shaded region r. Let d_2 be identical to d_1 except that r is completely non-shaded in d_2. Then d_1 logically entails d_2.

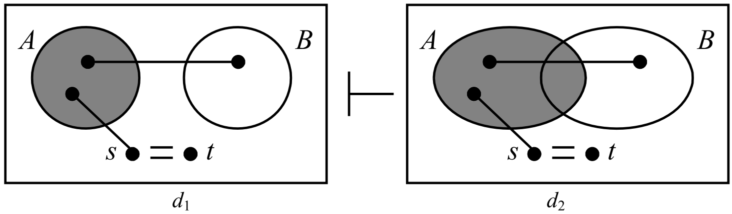


Fig. 5 An application of Rule 1 (introduction of a shaded zone)

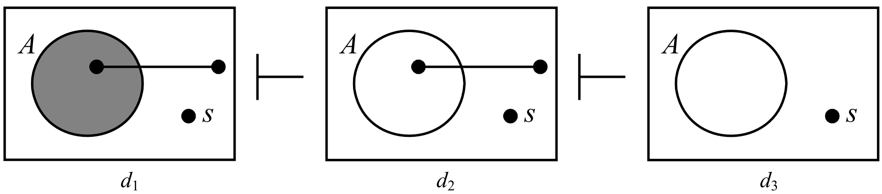


Fig. 6 Applications of Rule 2 (erasure of shading) and Rule 3 (erasure of a spider)

In Fig. 6, Rule 2 (erasure of shading) is applied to d_1 to give d_2.

Rule 3 (*Erasure of a spider*) Let d_1 be a unitary diagram containing a spider s with a completely non-shaded habitat. Let d_2 the same as d_1 except that d_2 does not contain s or any ties that were connected to s. Then d_1 logically entails d_2.

In Fig. 6, Rule 3 (erasure of a spider) is applied to d_2 to give d_3.

4.2 Unitary to Compound Reasoning Rules

We now specify five further rules, each of which is reversible, that allow a unitary diagram to be replaced by a compound diagram. The first of these rules allows us to introduce a contour. In the logic for spider diagrams without constants, the introduction of a contour rule applies to, and results in, a unitary diagram [12].

Before we formulate the introduction of a contour rule, we look at an example. In Fig. 7, we examine how to introduce the contour with label C to d_1, which contains constant spiders. When we do so, each zone must split into two new zones, thus ensuring that information is preserved. The habitats of the existential spiders are similarly altered. More care must be taken with the constant spiders, however, due to the presence of ties. Consider, for example, the constant spiders s and t. The individual represented by both s and t must be either in $C - (A \cup B)$ or in $U - (A \cup B \cup C)$. The constant spider u represents an individual that is either in $A - (B \cup C)$ or $(A \cap C) - B$. This gives rise to four possibilities, shown in d_2, d_3, d_4 and d_5. We call these four diagrams the C-*extensions* of d_1. The diagram d_1 is semantically equivalent to $d_2 \vee d_3 \vee d_4 \vee d_5$. We could replace d_1 with the disjunction of just two unitary diagrams, each with u having a two zone habitat: $(\{A\}, \{B, C\})$ and $(\{A, C\}, \{B\})$. However, it is not the case that the single unitary diagram d_6 in Fig. 8 is semantically equivalent to d_1. The constant spiders s and t must represent

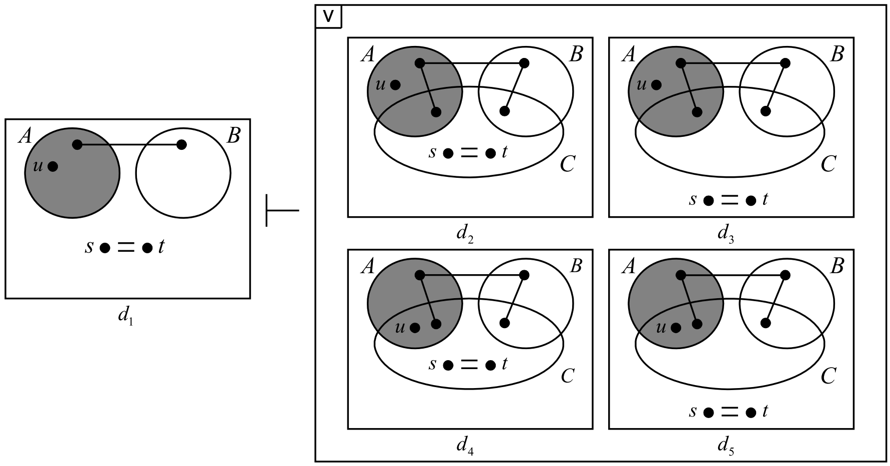
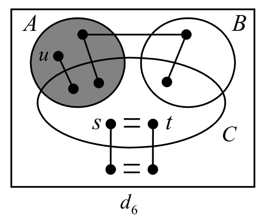


Fig. 7 A diagram with its C-extensions

Fig. 8 Introducing a contour: an incorrect application



the same individual in d_1 but this is not the case in d_6, since the semantics of ties are zone based.

To define this rule formally, we first define the component parts of the resulting disjunction. We call these component parts L_i-*extensions*, where L_i is the contour label introduced.

Definition 4.1 Let d_1 be a unitary diagram such that each constant spider in d_1 has a single zone habitat. Let L_i be a contour label that is not in d_1, that is $L_i \in \mathcal{CL} - L(d_1)$. Let d_2 be a unitary diagram such that each constant spider in d_2 has a single zone habitat. If the following conditions hold then d_2 is an $\boldsymbol{L_i}$**-extension** of d_1.

1. The contour labels of d_2 are those of d_1, together with L_i: $L(d_2) = L(d_1) \cup \{L_i\}$.
2. The constant spider labels match: $CS(d_1) = CS(d_2)$.
3. There exists a surjection, $h : Z(d_2) \to Z(d_1)$ defined by $h(a, b) = (a - \{L_i\}, b - \{L_i\})$ such that
   (a) each zone in d_1 is mapped to by two distinct zones in d_2,
   (b) each zone is shaded in d_2 if and only if it maps to a shaded zone,
   (c) the existential spiders match and their habitats are preserved under h: there exists a bijection, $\sigma : ES(d_1) \to ES(d_2)$ that satisfies

$$\forall e \in ES(d_1)\ \eta\big(\sigma(e)\big) = \big\{z \in Z(d_2) : h(z) \in \eta(e)\big\},$$

and

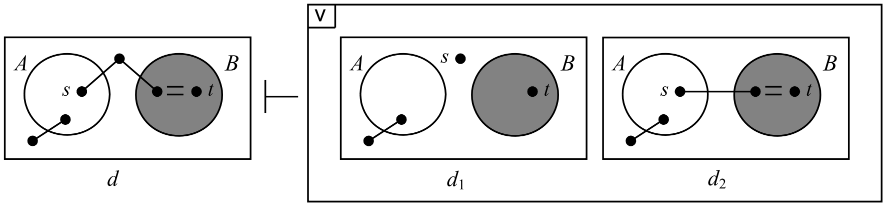


Fig. 9 An application of Rule 5, splitting spiders

(d) the habitat of each constant spider, c, in d_2 satisfies $h(\theta_{d_2}(c)) = \theta_{d_1}(c)$.

4. Spider webs are preserved. Since the constant spiders have a single zone habitat we may formalize this as follows:

$$\forall c_1, c_2 \in CS(d_2)\ \big(\omega_{d_1}(c_1, c_2) \neq \emptyset \Leftrightarrow \omega_{d_2}(c_1, c_2) \neq \emptyset\big).$$

We define $\mathcal{EXT}(L_i, d_1)$ to be the set of all L_i-extensions of d_1.

Rule 4 (*Introduction of a contour label*) Let d_1 ($\neq \bot$) be a unitary diagram such that each constant spider has a single zone habitat. Let $L_i \in \mathcal{CL} - L(d_1)$. Then d_1 is logically equivalent to the diagram

$$\bigvee_{d_2 \in \mathcal{EXT}(L_i, d_1)} d_2.$$

Rule 5 (*Splitting spiders*) Let d be a unitary diagram with a spider s touching every zone of two disjoint regions r_1 and r_2. Let d_1 and d_2 be unitary diagrams that are identical to d except that neither contains s, but instead each contains an extra spider, s_1 and s_2 respectively, whose habitats are regions r_1 in d_1 and r_2 in d_2. If s is a constant spider, then

1. s_1 and s_2 have the same label as s and
2. any ties joined to s in d are joined to the appropriate instance of s in d_1 and d_2.

Then d is logically equivalent to the diagram $d_1 \vee d_2$.

Figure 9 illustrates an application of the splitting spiders rule. The spider s in d splits into two spiders, one in d_1, the other in d_2. Intuitively, the individual represented by s is either in the set $U - (A \cup B)$ or the set $A \cup B$.

Rule 6 (*Excluded middle*) Let d be a unitary diagram with a completely non-shaded region r. Let d_1 and d_2 be unitary diagrams that are the same as d except that d_1 contains an extra existential spider whose habitat is r and in d_2 the region r is shaded. Then d is logically equivalent to the diagram $d_1 \vee d_2$.

For example, the diagram d in Fig. 10 can be replaced by $d_1 \vee d_2$ by applying the excluded middle rule.

Before we introduce the next rule, we look at an example, and then make a definition that is key to formulating the rule itself. Given a unitary diagram, d, that has only non-empty models (in which case d contains at least one spider), we can deduce that the

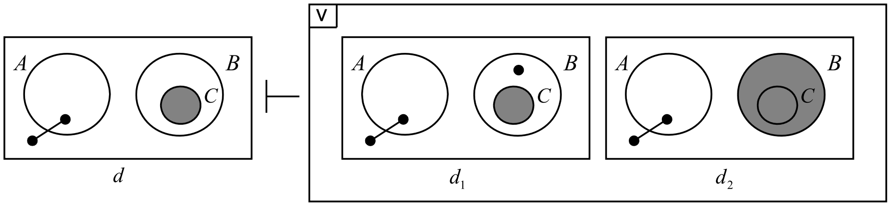


Fig. 10 An application of Rule 6, excluded middle

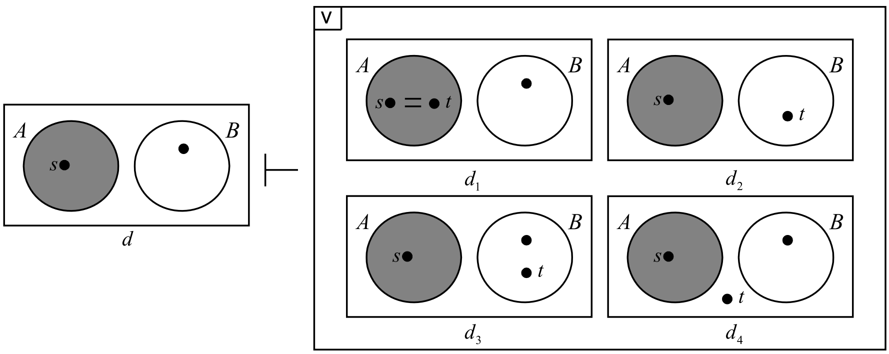


Fig. 11 A unitary diagram with its t-extensions

individual represented by a constant spider label, t, belongs to one of the sets denoted by the zones in d. Moreover, this individual must either be the same as, or different from, the elements already represented in d.

As an example, consider d in Fig. 11 which has only non-empty models. Thus, in any model for d the constant spider (label) t maps to some individual (technically, single element set). Then t is in $A - B$, $B - A$ or $U - (A \cup B)$. If t is in $A - B$ then it must equal s, since the region inside A is entirely shaded, shown in d_1. If t is in the set $B - A$ then it may be either equal to or different from the element represented by the existential spider in B in the diagram d; these cases are represented by d_2 and d_3 respectively. Finally, if t is not in $A - B$ or $B - A$ then, since $A \cap B = \emptyset$, t must be in $U - (A \cup B)$, represented by d_4. The diagrams d_1, d_2, d_3 and d_4 are called *t-extensions* of d. A diagram in which all spiders have a single zone habitat is called an $\boldsymbol{\alpha}$**-diagram**.

Definition 4.2 Let d_1 be a unitary α-diagram such that $S(d_1) \neq \emptyset$ and there exists $s_i \in \mathcal{CS} - CS(d_1)$. Let d_2 be a unitary α-diagram. If the following conditions are satisfied then d_2 is an $\boldsymbol{s_i}$**-extension** of d_1.

1. The zones match: $Z(d_1) = Z(d_2)$.
2. The shaded zones match: $Z^*(d_1) = Z^*(d_2)$.
3. The constant spiders match except that s_i is in d_2: $CS(d_1) \cup \{s_i\} = CS(d_2)$.
4. The habitats of the existing constant spiders are preserved: $\theta_{d_1} = \theta_{d_2}|_{CS(d_1)}$.
5. The existing webs are preserved: $\omega_{d_1} = \omega_{d_2}|_{CS(d_1) \times CS(d_1)}$.

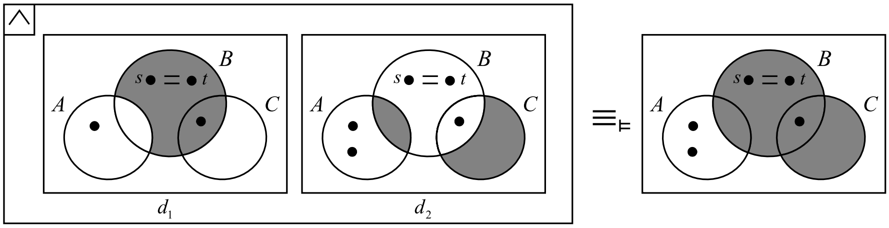


Fig. 12 Combining diagrams

6. If s_i has a shaded habitat, z, in d_2 then either the number of existential spiders inhabiting z is one less than the number in d_1 or s_i is joined to another (constant) spider by a tie: if $\theta_{d_2}(s_i) \subseteq Z^*(d_2)$ then
   (a) $\forall s_j \in CS(d_1)\ \omega_{d_2}(s_i, s_j) = \emptyset \wedge \exists e \in ES(\theta_{d_2}(s_i), d_1)\ ES(d_2) = ES(d_1) - \{e\}$ or
   (b) $\exists s_j \in CS(d_1)\ \omega_{d_2}(s_i, s_j) \neq \emptyset \wedge ES(d_1) = ES(d_2)$.
7. If s_i has a non-shaded habitat in d_2 then either the number of existential spiders inhabiting z is the same as, or one less than the number in d_1 or s_i is joined to another (constant) spider by a tie and the number of existential spiders is the same: if $\theta_{d_2}(s_i) \cap Z^*(d_2) = \emptyset$ then
   (a) $\forall s_j \in CS(d_1)\ \omega_{d_2}(s_i, s_j) = \emptyset \wedge (ES(d_1) = ES(d_2) \vee \exists e \in ES(\theta_{d_2}(s_i), d_1)$ $ES(d_2) = ES(d_1) - \{e\})$ or
   (b) $\exists s_j \in CS(d_1)\ \omega_{d_2}(s_i, s_j) \neq \emptyset \wedge ES(d_1) = ES(d_3)$.

We define $\mathcal{EXT}(s_i, d_1)$ to be the set of all s_i-extensions of d_1.

Rule 7 (*Introduction of a constant spider*) Let d_1 be a unitary α-diagram such that $S(d_1) \neq \emptyset$ and there exists $s_i \in \mathcal{CS} - CS(d_1)$. Then d_1 is logically equivalent to the diagram

$$\bigvee_{d_2 \in \mathcal{EXT}(s_i, d_1)} d_2.$$

Introducing the constant spider t to d in Fig. 11, results in $d_1 \vee d_2 \vee d_3 \vee d_4$.

The final rule in this section, called *combining*, replaces two unitary α-diagrams, with the same zone sets and constant spider label sets, taken in conjunction by a single unitary diagram, illustrated in Fig. 12. We combine $d_1 \wedge d_2$ to give d^*. Any shading in either d_1 or d_2 occurs in d^*. Moreover, the number of spiders in any zone in d^* is the same as the maximum number that occur in that zone in d_1 or d_2. The diagram $d_1 \wedge d_2$ is semantically equivalent to d^*.

We now give a further example in a build-up to the definition of the combining rule. In Fig. 13, d_1 and d_2 contain contradictory information. We observe the following.

1. The zone $z_1 = (\{A\}, \{B, C\})$ is shaded in d_1 and contains more spiders in d_2. Moreover, z_1 represents the empty set in any model for d_1. In any model for d_2, z_1 does not represent the empty set.
2. The constant spider u has different habitats in the two diagrams. In any model for d_1, u represents an individual that is not in the set $A \cup C$. In any model for d_2, u represents an individual in the set C.

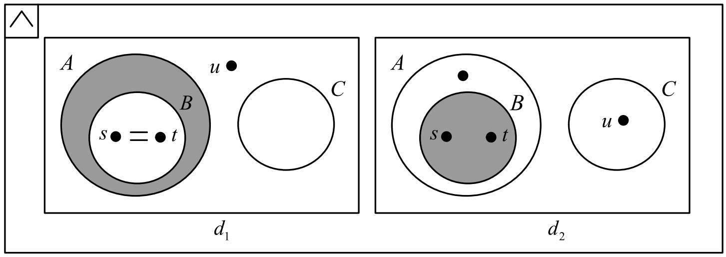


Fig. 13 An unsatisfiable diagram

3. The constant spiders s and t are joined by a tie in d_1 but not in d_2. In any model for d_1, s and t represent the same individual, but in any model for d_2 they represent distinct individuals.

From any one of these three observations we can deduce that $d_1 \wedge d_2$ is unsatisfiable.

Definition 4.3 Let d_0 and d_1 be unitary α-diagrams. Then d_0 and d_1 are **comparable** if one of the following three conditions holds.

1. $Z(d_0) = Z(d_1)$ and $CS(d_0) = CS(d_1)$.
2. $Z(d_0) = Z(d_1)$.
3. for one of the d_is where $i \in \{0, 1\}$, $Z^*(d_i) = Z(d_i)$ and $S(d_i) = \emptyset$.
4. $d_0 = \bot$ or $d_1 = \bot$.

Recall that $S(\{z\}, d) = \{s \in S(d) : \eta(s) = \{z\}\}$.

Definition 4.4 Let d_0 and d_1 be comparable unitary α-diagrams. Then d_0 and d_1 are in **contradiction** if one of the following four conditions holds.

(i) Either $d_0 = \bot$ or $d_1 = \bot$.
(ii) There is a zone that is shaded in one diagram and contains more spiders in the other. More formally, there exists $z \in Z(d_i)$ for some $i = 0, 1$ such that $z \in Z^*(d_j)$ and $|S(\{z\}, d_i)| > |S(\{z\}, d_j)|$ where $j = 1 - i$.
(iii) There is a constant spider with different habitats in d_0 and d_1. More formally, $\theta_{d_0} \neq \theta_{d_1}$.
(iv) There are two constant spiders that are joined by a tie in one diagram but not the other. More formally, $\omega_{d_0} \neq \omega_{d_1}$.

It may be helpful to note that if d_0 and d_1 are comparable and not in contradiction then $\omega(d_0) = \omega(d_1)$.

Lemma 4.5 *Let d_0 and d_1 be comparable unitary α-diagrams. Then d_0 and d_1 are in contradiction if and only if $d_0 \wedge d_1$ is unsatisfiable.*

Definition 4.6 Let d_0 and d_1 be comparable unitary α-diagrams. Then their **combination**, denoted $d^* = d_0 * d_1$, is a unitary α-diagram defined as follows.

1. If d_0 and d_1 are in contradiction then $d_0 * d_1 = \bot$.

2. Otherwise $d^* = d_0 * d_1$ is a unitary α-diagram such that the following hold.
   (a) The set of zones in the combined diagram is the same as the set of zones in the original diagrams: $Z(d^*) = Z(d_0)$.
   (b) The shaded zones in $d^* = d_0 * d_1$ are those that are shaded in at least one of the original diagrams: $Z^*(d^*) = Z^*(d_0) \cup Z^*(d_1)$.
   (c) The number of existential spiders in any zone in the combined diagram is the maximum number of existential spiders inhabiting that zone in the original diagrams:

$$\forall z \in Z\big(d^*\big)\ ES\big(\{z\}, d^*\big) = ES\big(\{z\}, d_0\big) \cup ES\big(\{z\}, d_1\big).$$

   Equivalently, $ES(d^*) = ES(d_0) \cup ES(d_1)$.
   (d) The constant spiders in the combined diagram are the same as those in the original diagrams: $CS(d^*) = CS(d_0)$.
   (e) The habitats of the constant spiders in the combined diagram are the same as those in the original diagrams: $\theta_{d^*} = \theta_{d_0}$.
   (f) The webs of the constant spiders in the combined diagram are the same as those in the original diagrams: $\omega(d^*) = \omega(d_0)$.

Rule 8 (*Combining*) Let d_0 and d_1 be comparable unitary α-diagrams. Then $d_0 \wedge d_1$ is logically equivalent to $d_0 * d_1$.

4.3 Logic Reasoning Rules

We now introduce a collection of rules, all of which have (obvious) analogies in symbolic logic. The next rule is analogous to $P \vdash P \vee Q$, for any propositions P, Q.

Rule 9 (*Connecting a diagram*) Let D_1 and D_2 be spider diagrams. Then D_1 logically entails $D_1 \vee D_2$.

Rule 10 (*Inconsistency*) The diagram $\perp$ logically entails any diagram.

Rule 11 ($\vee$-*Idempotency*) Any spider diagram D is logically equivalent to $D \vee D$.

Rule 12 ($\wedge$-*Idempotency*) Any spider diagram D is logically equivalent to $D \wedge D$.

Rule 13 ($\vee$-*Commutativity*) Let D_1 and D_2 be spider diagrams. Then $D_1 \vee D_2$ is logically equivalent to $D_2 \vee D_1$.

Rule 14 ($\wedge$-*Commutativity*) Let D_1 and D_2 be spider diagrams. Then $D_1 \wedge D_2$ is logically equivalent to $D_2 \wedge D_1$.

Rule 15 ($\vee$-*Associativity*) Let D_1, D_2 and D_3 be spider diagrams. Then $D_1 \vee (D_2 \vee D_3)$ is logically equivalent to $(D_1 \vee D_2) \vee D_3$.

Rule 16 ($\wedge$-*Associativity*) Let D_1, D_2 and D_3 be spider diagrams. Then $D_1 \wedge (D_2 \wedge D_3)$ is logically equivalent to $(D_1 \wedge D_2) \wedge D_3$.

Rule 17 ($\vee$-*Distributivity*) Let D_1, D_2 and D_3 be spider diagrams. Then $D_1 \vee (D_2 \wedge D_3)$ is logically equivalent to $(D_1 \vee D_2) \wedge (D_1 \vee D_3)$.

Rule 18 ($\wedge$-*Distributivity*) Let D_1, D_2 and D_3 be spider diagrams. Then $D_1 \wedge (D_2 \vee D_3)$ is logically equivalent to $(D_1 \wedge D_2) \vee (D_1 \wedge D_3)$.

Rule 19 ($\vee$-*Simplification*) Let D_1, D_2 and D_3 be spider diagrams. If diagram D_2 can be transformed into diagram D_3 by one of reasoning rules then $D_1 \vee D_2$ logically entails $D_1 \vee D_3$.

Rule 20 ($\wedge$-*Simplification*) Let D_1, D_2 and D_3 be spider diagrams. If diagram D_2 can be transformed into diagram D_3 by one of the reasoning rules then $D_1 \wedge D_2$ logically entails $D_1 \wedge D_3$.

4.4 Obtainability

To conclude this section on reasoning rules we define obtainability.

Definition 4.7 Let D_1 and D_2 be two spider diagrams with constants. Diagram D_2 is **obtainable** from D_1, denoted $D_1 \vdash D_2$, if and only if there is a sequence of diagrams $\langle D^1, D^2, \ldots, D^m \rangle$ such that $D^1 = D_1$, $D^m = D_2$ and D^{k+1} can be obtained from D^k (where $1 \leq k < m$) by applying a reasoning rule. If $D_1 \vdash D_2$ and $D_2 \vdash D_2$, we write $D_1 \equiv_{\vdash} D_2$.

5 Soundness

In this section we show the soundness of the logic of spider diagrams with constants introduced in Sect. 4.

To prove that the system is sound, the strategy is to start by showing that each of the reasoning rules is sound. We show that the introduction of a constant spider rule is sound as an illustration but omit the remaining proofs. The soundness theorem then follows by a simple induction argument.

Lemma 5.1 *Rule* 7 (*introduction of a constant spider*) *is sound. Let* d_1 *be unitary* α-*diagram such that* $S(d_1) \neq \emptyset$ *and there exists* $s_i \in \mathcal{CS} - CS(d_1)$. *Then*

$$d_1 \equiv_{\models} \bigvee_{d_2 \in \mathcal{EXT}(s_i, d_1)} d_2.$$

Proof Let $m = (U, \Psi)$ be an interpretation. Assume that $m \models d_1$. We will show that $m \models d_2$, for some $d_2 \in \mathcal{EXT}(s_i, d_1)$. Let $\Psi_1 : \mathcal{R} \cup \mathcal{CS} \cup ES(d_1) \rightarrow \mathbb{P}U$ be a valid extension to existential spiders for d_1. Using d_1 and Ψ_1, we define a diagram, d_2, as follows.

1. The zones match: $Z(d_1) = Z(d_2)$.

2. The shaded zones match: $Z^*(d_1) = Z^*(d_2)$.
3. The constant spiders in d_1 are in d_2 and, additionally, d_2 contains s_i: $CS(d_1) \cup \{s_i\} = CS(d_2)$.
4. The habitats of the constant spiders match and the habitat of s_i in d_2 is determined by Ψ_1:

$$\theta_{d_1} = \theta_{d_2}|_{CS(d_1)}$$

and

$$\theta_{d_2}(s_i) = \{z\}$$

where z is the unique zone in $Z(d_1)$ such that $\Psi(s_i) \subseteq \Psi(z)$. Such a zone exists because the plane tiling condition holds for d_1.
5. The existing webs in d_1 are preserved in d_2: $\omega_{d_1} = \omega_{d_2}|_{CS(d_1)\times CS(d_1)}$.
6. We now consider three cases in order to define the existential spiders (and their habitats) and the remaining webs of d_2.
   (a) *There is an existential spider, s, in d_1 such that $\Psi_1(s) = \Psi(s_i)$.* In this case, we choose $e_n(\{\eta(s)\})$, where $(n, \eta(s)) \in ESD(d_1)$, and we define $ES(d_2) = ES(d_1) - \{e_n(\eta(s))\}$. For the remaining webs, we define, for all $s_j \in CS(d_1)$, $\omega_{d_2}(s_i, s_j) = \emptyset$. We note, by the spiders condition for d_1, $\theta_{d_2}(s_i) = \eta(s)$.
   (b) *There is a constant spider, c, in d_1 such that $\Psi(c) = \Psi(s_i)$.* In this case, $ES(d_1) = ES(d_2)$, and, for the remaining webs, we start by defining $\omega_{d_2}(s_i, c) = \theta_{d_1}(c)$; since d_1 is an α-diagram, $\theta_{d_1}(c)$ is a single zone. It follows that s_i is also joined by a tie to all the constant spiders that are joined to c in d_1 and, by (5) above and the transitivity of ties, not joined by a tie to any other constant spiders. We note, by the spiders condition for d_1, $\theta_{d_2}(s_i) = \theta_{d_1}(c)$.
   (c) *No spider, s, in $S(d_1)$ satisfies $\Psi_1(s) = \Psi(s_i)$.* In this case, we have $ES(d_1) = ES(d_2)$ and for all $c \in CS(d_1)$, $\omega_{d_2}(s_i, c) = \emptyset$.

It is straightforward to verify that d_2 is an s_i-extension of d_1.

We now show that $m \models d_2$. Clearly, the plane tiling condition holds for d_2, since $Z(d_1) = Z(d_2)$. If case 6(a) holds then we suppose, without loss of generality, that $s = e_n(\eta(s))$. If either case 6(b) or 6(c) holds then no supposition is necessary. We define an extension of Ψ to the existential spiders in d_2 by $\Psi_2 = \Psi_1|_{\mathcal{R}\cup\mathcal{CS}\cup ES(d_2)}$. The function Ψ_2 is a valid extension of Ψ to existential spiders for d_2. Hence $m \models d_2$, and it follows that

$$d_1 \vDash \bigvee_{d_2 \in \mathcal{EXT}(s_i, d_1)} d_2.$$

For the converse, it can be shown that each $d_2 \in \mathcal{EXT}(s_i, d_1)$ satisfies $d_2 \vDash d_1$. Assuming that $m \models d_2$, the proof strategy is to take a valid extension of Ψ to existential spiders for d_2 and use this to construct a valid extension of Ψ to existential spiders for d_1. Thus,

$$\bigvee_{d_2 \in \mathcal{EXT}(s_i, d_1)} d_2 \vDash d_1.$$

Hence

$$d_1 \equiv_\models \bigvee_{d_2 \in \mathcal{EXT}(s_i, d_1)} d_2,$$

that is, Rule 7 (introduction of a constant spider) is sound. □

Theorem 5.2 (*Soundness*) *Let D_1 and D_2 be spider diagrams. If $D_1 \vdash D_2$ then $D_1 \models D_2$.*

Proof The proof is by induction on the length, n, of a sequence establishing $D_1 \vdash D_2$, since each individual step can be shown to be sound along the lines of the proof of Lemma 5.1 above. □

6 Completeness and Decidability

In this section we show the completeness and decidability of the logic of spider diagrams with constants introduced in Sect. 4. We begin with an informal overview, before giving details of the various stages of the proof.

6.1 Overview

The completeness proof strategy for spider diagrams without constants given in [12] extends to the more general case here. The extended strategy, outlined in Fig. 14, is as follows. Suppose that $D_1 \models D_2$. The aim is to transform D_1 and D_2 into disjunctions of unitary α-diagrams using reversible rules (i.e. those which are logical equivalences) where, roughly speaking, each unitary part has some specified contour label set and constant spider label set.

Firstly, we split the constant spiders in D_1 and D_2 until, in each unitary part, all the constant spiders have a single zone habitat, giving D_1^S and D_2^S respectively. This allows us to add contours to the unitary parts in both D_1^S and D_2^S using the reversible Rule 4 (introduction of a contour label), until each (non-false) unitary part has the same contour label set, L. This gives D_1^L and D_2^L respectively. For the next step, zones are introduced to each unitary part until all (non-false) unitary parts have the same zone set, Z. This is done using the reversible Rule 1 (introduction of a shaded zone) and yields D_1^Z and D_2^Z respectively. Now we obtain α-diagrams using the reversible Rule 5 (splitting spiders), yielding D_1^α and D_2^α respectively. The formalization of the diagrams D_i^L, D_i^Z and D_i^α readily generalize those given in [12] for spider diagrams without constants.

We wish to introduce constant spiders to each side until each unitary part has the same constant spider label set. However, we can only introduce constant spiders when our diagrams contain at least one spider (ensuring non-empty models). Thus the next step we take is to apply the excluded middle rule to both sides until all the (non-false) unitary parts are either entirely shaded or contain at least one spider. The reversible Rule 7 (introduction of a constant spider) is then applied, introducing constant spiders to all unitary parts that contain a spider, until all such unitary parts have some specified constant spider label set, C. This gives D_1^C and D_2^C respectively.

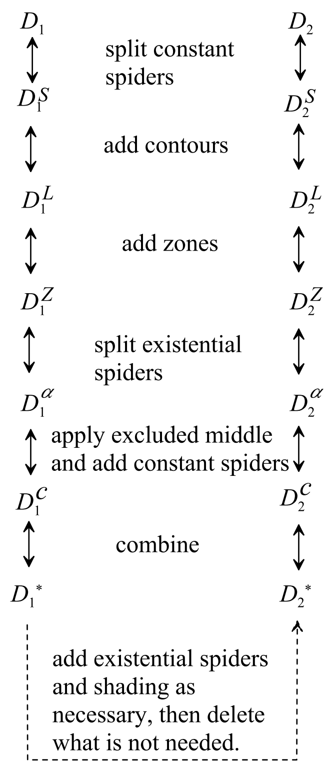


Fig. 14 The completeness proof strategy

We now apply Rule 8 (combining) to remove all the conjuncts, giving two disjunctions of unitary α-diagrams, D_1^* and D_2^*. We call D_1^* (D_2^*) the **disjunctified diagram associated with D_1 (D_2) given D_2 (D_1)**. All of the unitary parts of D_1^* and D_2^* are either

1. $\perp$,
2. have zone set Z and are entirely shaded and contain no spiders, or
3. have zone set Z and constant spider label set C.

Note that $D_1 \equiv_\vdash D_1^*$ and $D_2 \equiv_\vdash D_2^*$, since all the rules applied so far are reversible. The diagram D_i^* is a normal form that reflects the semantics of D_i clearly. We now apply the excluded middle rule to D_1^* until there are sufficiently many existential spiders and there is enough shading to ensure that each unitary part on the left hand side syntactically entails a unitary part of D_2^*.

The details of the proof are given in the following sections. The major differences between the completeness proof strategy here and that for spider diagrams without constants are the addition of the first step (splitting the constant spiders), with knock on changes to details of the other steps, and the insertion of an extra stage between splitting existential spiders and combining diagrams. In addition, we note that the details of the proofs are more complex.

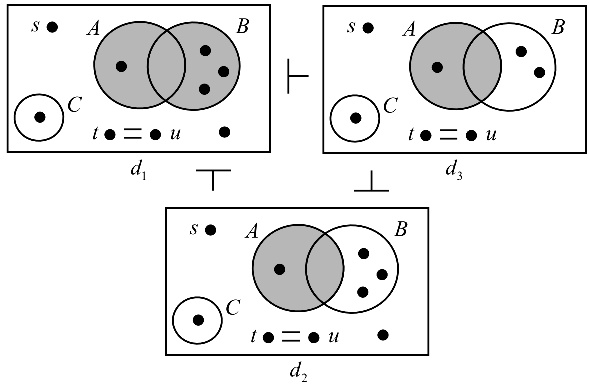


Fig. 15 Completeness for unitary α-diagrams

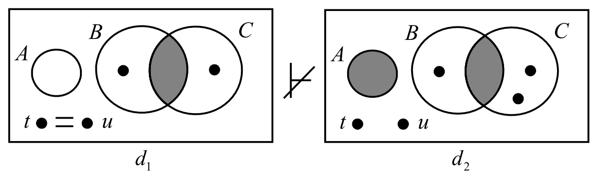


Fig. 16 Completeness for unitary α-diagrams

6.2 Completeness for Unitary α-Diagrams

We show that if $d_1 \vDash d_2$, where d_1 and d_2 are unitary α-diagrams with some fixed zone set and constant spider label set, then we can erase existential spiders and shading from d_1 to give d_2.

Example The diagrams d_1 and d_2 in Fig. 15 satisfy the following.

(a) Every shaded zone in d_2 is shaded in d_1 and contains the same number of existential spiders in both diagrams.
(b) Every zone in d_2 contains the same number or fewer existential spiders than in d_1.
(c) The constant spiders habitats match, as do their webs.

Under these conditions, the diagram d_2 can be obtained from d_1 by applying Rule 2 (erasure of shading), and Rule 3 (erasure of an existential spider) can then be used to give d_3. The properties (a), (b) and (c) above relate to properties 3(a), 3(b) and 3(c) in Theorem 6.1.

Example The diagram d_2 in Fig. 16 cannot be obtained from d_1 for three reasons.

(a) The zone $(\{A\}, \{B, C\})$ is shaded in d_2 but not shaded in d_1. There is a model for d_1 that will cause the shading condition for d_2 to fail whenever the spiders condition for d_2 holds.
(b) The zone $(\{C\}, \{A, B\})$ contains a two existential spiders in d_2 but only a single existential spider in d_1. Again we can deduce that there is a model for d_1 that does not satisfy d_2. For example, at least one model, $m = (U, \Psi)$ for d_1 ensures that $|\Psi(\{C\}, \{A, B\})| = 1$. In the interpretation m, it cannot be that case that both the spiders condition and the existential spiders condition hold for d_2.

(c) The constant spiders t and u have the same habitat in both diagrams, but different webs. In any model for d_1, t and u represent the same individual, but in any model for d_2 they represent distinct individuals.

From any one of the above observations we can deduce that $d_1 \not\models d_2$.

The following theorem gives syntactic conditions on unitary α-diagrams equivalent to semantic and syntactic entailment. The theorem forms the heart of the proof of completeness and is modified from the corresponding result in [12] to take account of the fact the our spider diagrams now include constant spiders.

Theorem 6.1 *Let d_1 $(\neq \bot)$ and d_2 $(\neq \bot)$ be two unitary α-diagrams. If $Z(d_1) = Z(d_2)$ and $CS(d_1) = CS(d_2)$ then the following three statements are equivalent*:

1. $d_1 \vdash d_2$.
2. $d_1 \models d_2$.
3. (a) *every zone that is shaded in d_2 is shaded in d_1 and contains the same number of existential spiders in both diagrams*:

$$Z^*(d_2) \subseteq Z^*(d_1) \ \wedge \ \forall z \in Z^*(d_2) \ ES\big(\{z\}, d_2\big) = ES\big(\{z\}, d_1\big),$$

(b) *every zone in d_2 contains at most the same number of existential spiders as in d_1*:

$$\forall z \in Z(d_2) \ ES\big(\{z\}, d_2\big) \subseteq ES\big(\{z\}, d_1\big),$$

and

(c) *the constant spiders have the same habitats and the same webs in both diagrams*: $\theta_{d_1} = \theta_{d_2}$ *and* $\omega_{d_1} = \omega_{d_2}$.

Proof By soundness, $d_1 \vdash d_2 \Rightarrow d_1 \models d_2$.

We now show that 2 (i.e., $d_1 \models d_2$) implies 3. Suppose that $d_1 \models d_2$ and let $m = (U, \Psi)$ be a standard model for d_1. We define, for each existential spider, e_1, in d_1, $\Psi_1(e) = \{e\}$ and the mapping Ψ_1 yields a valid extension to existential spiders for d_1. Since $d_1 \models d_2$, m is a model for d_2. Let $\Psi_2 : \mathcal{R} \cup \mathcal{CS} \cup ES(d_2) \to \mathbb{P}U$ be a valid extension to existential spiders for d_2. We will show that Ψ_2 induces an injective, habitat preserving map $\sigma : ES(d_2) \to ES(d_1)$. Now, Ψ_2 ensures that the spiders condition holds for d_2. Therefore, for each existential spider, e_2, in d_2, there exists an existential spider, e_1, in d_1 such that $\Psi_2(e_2) = \{e_1\}$ (each constant spider, s_i, in d_2 maps to $[s_i]$). Define σ by

$$\sigma(e_2) \in \Psi_2(e_2).$$

By the spiders condition for d_1,

$$\big\{\sigma(e_2)\big\} = \Psi_1\big(\sigma(e_2)\big) \subseteq \Psi\Big(\eta\big(\sigma(e_2)\big)\Big)$$

and, by the spiders condition for d_2,

$$\big\{\sigma(e_2)\big\} = \Psi_2(e_2) \subseteq \Psi\big(\eta(e_2)\big).$$

We deduce that, since distinct zones in d_1 represent disjoint sets,

$$\eta\big(\sigma(e_2)\big) = \eta(e_2).$$

Therefore σ is habitat preserving. We now show that σ is injective. Suppose that $\sigma(e_2) = \sigma(e_3)$ for some $e_3 \in ES(d_2)$. Then $\Psi_2(e_2) = \Psi_2(e_3)$, which implies, by the existential spiders condition for d_2, $e_2 = e_3$. Hence σ is injective. We deduce that 3(b) holds. It can also be shown that, for all $z \in Z^*(d_2)$,

$$ES\big(\{z\}, d_2\big) = ES\big(\{z\}, d_1\big).$$

Moreover, it is obvious that $d_1 \vDash d_2$ implies $Z^*(d_2) \subseteq Z^*(d_1)$. Thus 3(a) holds.

We now consider 3(c). The spiders condition for d_1 states, in part,

$$\forall s_i \in CS(d_1)\ \Psi(s_i) \subseteq \Psi\big(\theta_{d_1}(s_i)\big).$$

Since $CS(d_1) = CS(d_2)$, we deduce that

$$\forall s_i \in CS(d_2)\ \Psi(s_i) \subseteq \Psi\big(\theta_{d_1}(s_i)\big). \tag{1}$$

The spiders condition for d_2 states, in part,

$$\forall s_i \in CS(d_2)\ \Psi(s_i) \subseteq \Psi\big(\theta_{d_2}(s_i)\big). \tag{2}$$

Since distinct zones in d_1 represent disjoint sets, it follows from (1) and (2) that

$$\forall s_i \in CS(d_2)\ \theta_{d_1}(s_i) = \theta_{d_2}(s_i).$$

Hence $\theta_{d_1} = \theta_{d_2}$. Suppose that constant spiders s_i and s_j are joined by a tie in d_1. That is,

$$\omega_{d_1}(s_i, s_j) = \theta_{d_1}(s_i).$$

Then $\Psi(s_i) = \Psi(s_j)$, by the constant spiders condition for d_1. By the constant spiders condition for d_2,

$$\exists z \in \omega_{d_2}(s_i, s_j)\ \Psi(s_i) = \Psi(s_j).$$

Therefore, s_i and s_j are joined by a tie in d_2. That is,

$$\omega_{d_2}(s_i, s_j) = \theta_{d_2}(s_i) = \theta_{d_1}(s_i).$$

Alternatively, suppose that spiders s_i and s_j are not joined by a tie in d_1. That is,

$$\omega_{d_1}(s_i, s_j) = \emptyset.$$

Then $\Psi(s_i) \neq \Psi(s_j)$ so it cannot be that s_i and s_j are joined by a tie in d_2. That is,

$$\omega_{d_2}(s_i, s_j) = \emptyset.$$

Hence $\omega_{d_1} = \omega_{d_2}$. Thus 3(c) holds.

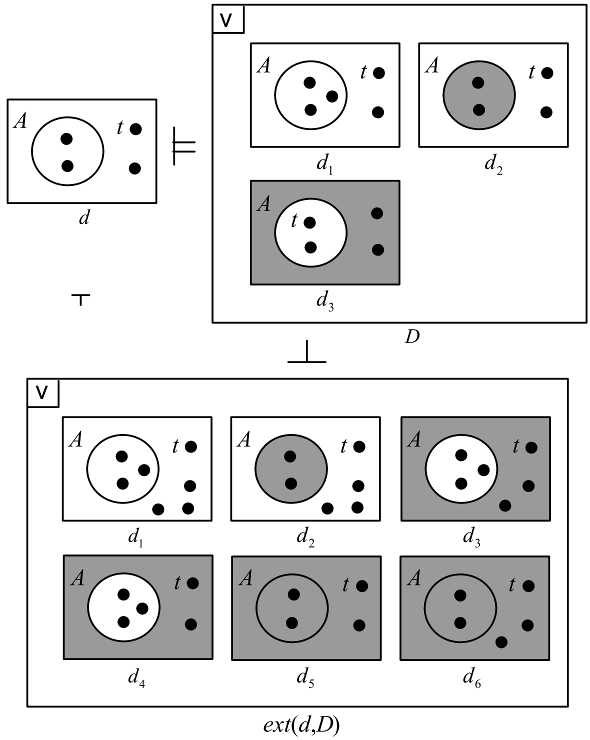


Fig. 17 An α-diagram and an extended diagram

Finally to show that 3 implies 1, it can be shown that shading and existential spiders can be deleted from d_1, using Rules 2 and 3 respectively, to give d_2. Hence all three statements are equivalent. □

6.3 *Extended Diagrams*

Example In Fig. 17, the diagram D is a semantic consequence of d but no unitary component of D is semantically entailed by d; that is $d \not\vDash d_1$, $d \not\vDash d_2$ and $d \not\vDash d_3$. The diagram $ext(d, D)$ can be obtained from d (and vice versa) by applying Rules 6 (excluded middle) and 19 ($\vee$-simplification). The spiders and shading introduced to d to obtain $ext(d, D)$ are determined by D. For example, consider the outside zone $(\emptyset, \{A\})$. In d_3, this zone is shaded and contains two existential spiders and no other unitary component of D contains more than two existential spiders in this zone. In $ext(d, D)$, this zone contains either one, two or three existential spiders in any unitary component. The process of constructing $ext(d, D)$ will be described in Definitions 6.2 and 6.3 below.

Note that we have

$$d_1' \vDash d_1, \quad d_2' \vDash d_2, \quad d_3' \vDash d_1, \quad d_4' \vDash d_1, \quad d_5' \vDash d_2, \quad \text{and} \quad d_6' \vDash d_2$$

so, for each unitary component d'_i of $ext(d, D)$, there exists a unitary component d_j of D such that $d'_i \vDash d_j$. In fact,

$$d'_1 \vee d'_3 \vee d'_4 \vdash d_1 \quad \text{and} \quad d'_2 \vee d'_5 \vee d'_6 \vdash d_2.$$

Therefore

$$ext(d, D) = d'_1 \vee d'_3 \vee d'_4 \vee d'_2 \vee d'_5 \vee d'_6 \vdash d_1 \vee d_2.$$

By Rule 9 (connecting a diagram) $d_1 \vee d_2 \vdash D$ and by transitivity $ext(d, D) \vdash D$. Therefore $d \vdash D$, since $d \equiv_{\vdash} ext(d, D)$.

In general, the diagram $ext(d, D)$ will be constructed by taking copies of d and adding shading and existential spiders, as specified below. The unitary components of $ext(d, D)$ are called *extended unitary components associated with* d, which we now define. Firstly, we define $comp(D)$ to be the set of all the unitary parts of D.

Definition 6.2 Let d ($\neq \bot$) be a unitary α-diagram and D be an α-diagram. Then, given D, a unitary α-diagram ${}^e d$ is an **extended unitary component associated with** d, denoted $d \sqsubseteq_e^D {}^e d$, if and only if the following seven conditions are satisfied.

1. The diagrams d and ${}^e d$ have the same zones: $Z(d) = Z({}^e d)$.
2. All shading in d occurs in ${}^e d$: $Z^*(d) \subseteq Z^*({}^e d)$.
3. All existential spiders in d occur in ${}^e d$: $ES(d) \subseteq ES({}^e d)$.
4. If zone z is shaded in d then the existential spiders match in d and ${}^e d$: $\forall z \in Z^*(d)$ $ES(\{z\}, d) = ES(\{z\}, {}^e d)$.
5. If zone z is not shaded in d but is shaded in some unitary component of D and the number, m say, of existential spiders that z contains in d is at most the number that z contains in any unitary component of D in which z is shaded then
   (a) if z is shaded in ${}^e d$ then z contains at most m spiders in ${}^e d$; and
   (b) if z is not shaded in ${}^e d$ then z contains $m + 1$ spiders in ${}^e d$.
   More formally:

$$\forall z \in Z(d) - Z^*(d)$$
$$\left(\left(z \in \bigcup_{d_i \in comp(D)} Z^*(d_i) \wedge ES\big(\{z\}, d\big) \subseteq \bigcup_{\substack{d_i \in comp(D) \\ z \in Z^*(d_i)}} ES\big(\{z\}, d_i\big)\right)\right.$$
$$\Rightarrow \left(\left(z \in Z^*\big({}^e d\big) \wedge ES\big(\{z\}, {}^e d\big) \subseteq \bigcup_{\substack{d_i \in comp(D) \\ z \in Z^*(d_i)}} ES\big(\{z\}, d_i\big)\right)\right.$$
$$\left.\left.\vee \left(z \in Z\big({}^e d\big) - Z^*\big({}^e d\big) \wedge \big|ES\big(\{z\}, {}^e d\big)\big| = \left|\bigcup_{\substack{d_i \in comp(D) \\ z \in Z^*(d_i)}} ES\big(\{z\}, d_i\big)\right| + 1\right)\right)\right).$$

6. If a non-shaded zone z in d is not shaded in any unitary component of D or z contains more spiders in d than any shaded occurrence of z in D then z is not shaded in ${}^e d$ and

z contains the same number of spiders in ${}^{e}d$ as in d. More formally:

$$\forall z \in Z(d) - Z^*(d)$$
$$\Big(z \notin \bigcup_{d_i \in comp(D)} Z^*(d_i) \vee ES(\{z\}, d) \supset \bigcup_{\substack{d_i \in comp(D) \\ z \in Z^*(d_i)}} ES(\{z\}, d_i)\Big)$$
$$\Rightarrow \big(z \in Z({}^{e}d) - Z^*({}^{e}d) \wedge ES(\{z\}, {}^{e}d) = S(\{z\}, d)\big).$$

7. The constant spiders and their webs match: $CS(d_1) = CS(d_2)$, $\theta_{d_1} = \theta_{d_2}$ and $\omega_{d_1} = \omega_{d_2}$.

If $d = \perp$ then the **extended unitary component associated with** d is $\perp$.

Definition 6.3 Let d be a unitary α-diagram and let D be a disjunction of unitary α-diagrams such that d is comparable to each $d_i \in comp(D)$. Given D, let $\mathcal{D}_e^d$ be the set of all extended unitary components associated with d

$$\mathcal{D}_e^d = \big\{d' \in \mathcal{D}_0 : d \sqsubseteq_e^D d'\big\}.$$

Then the diagram

$$ext(d, D) = \bigvee_{d' \in \mathcal{D}_e^d} d'$$

is the **extended diagram associated with** d **in the context of** D.

Example In Fig. 17, each d_i' $(i = 1, \ldots, 6)$ is an extended unitary component associated with d, given D. Indeed, all such extended components ${}^{e}d$ are present, so $ext(d, D)$ is the extended diagram associated with d in the context of D.

Theorem 6.4 *Let d be a unitary α-diagram and let D be a disjunction of unitary α-diagrams such that d is comparable to each $d_i \in comp(D)$. Then d is syntactically equivalent to $ext(d, D)$, the extended diagram associated with d in the context of D*:

$$d \equiv_{\vdash} ext(d, D).$$

Sketch of proof Follows by repeated application of Rules 6 (excluded middle) and 19 ($\vee$-simplification) to d in the case where $d \neq \perp$. When $d = \perp$ the result follows immediately. □

6.4 The Completeness Theorem

The next result is the final prerequisite to our proof of completeness.

Theorem 6.5 *Let d $(\neq \perp)$ be a unitary α-diagram such that $S(d) \neq \emptyset$. Let D be a disjunction of unitary α-diagrams such that d is comparable to each $d_i \in comp(D)$. Given D,* *let* ${}^e d \in \mathcal{D}_e^d$. *If* ${}^e d \vDash D$ *then there exists a unitary component of* D, *say* d_i, *such that* ${}^e d \vDash d_i$:

$${}^e d \vDash D \quad \Rightarrow \quad \exists d_i \in comp(D)\ {}^e d \vDash d_i.$$

Proof The proof is by contradiction. Assume ${}^e d \vDash D$ but there is no $d_i \in comp(D)$ for which ${}^e d \vDash d_i$. We will show that a standard model, $m = (U, \Psi)$, for ${}^e d$ does not satisfy D, giving the contradiction we seek. The interpretation m does not satisfy D if and only if m does not satisfy any unitary part, d_i, of D. There are three types of d_i to consider.

1. $d_i = \bot$. Clearly m does not satisfy $\bot$.
2. $Z(d) = Z(d_i)$ and $Z^*(d_i) = Z(d_i)$ and $S(d_i) = \emptyset$. Since d contains at least one spider, so too does ${}^e d$. Therefore $U \neq \emptyset$. But d_i has only one model: the empty model (that is, $U = \emptyset$). Therefore m does not satisfy d_i.
3. $Z(d) = Z(d_i)$ and $CS(d_i) = CS(d)$ and $S(d_i) \neq \emptyset$. Firstly, suppose that m satisfies d_i and we will reach a contradiction, thus completing the proof that m does not satisfy any unitary part of D. Since m satisfies d_i, it must be that $m \models {}^e d \wedge d_i$, so ${}^e d$ and d_i are not in contradiction. We immediately deduce, by Lemma 4.5, that the following conditions do not hold.

   (a$_1$) There is a zone that is shaded in one diagram and contains more spiders in the other diagram. More formally, either

   $$\exists z \in Z^*({}^e d)\ \big|S(\{z\}, d_i)\big| > \big|S(\{z\}, {}^e d)\big|$$

   or

   $$\exists z \in Z^*(d_i)\ \big|S(\{z\}, {}^e d)\big| > \big|S(\{z\}, d_i)\big|.$$

   (b$_1$) There are two constant spiders that are joined by a tie in one diagram but not the other. More formally, $\omega_{d_i} \neq \omega_{{}^e d}$.

   Since (b$_1$) does not hold, we deduce that

   $$\omega_{d_i} = \omega_{{}^e d}. \tag{3}$$

   Since (a$_1$) does not hold, we deduce that

   $$\forall z \in Z^*(d_i)\ \big|S(\{z\}, {}^e d)\big| \leq \big|S(\{z\}, d_i)\big|.$$

   Since $m \models d_i$, and the fact that d_i is an α-diagram, for all $z \in Z^*(d_i)$,

   $$\big|\Psi(z)\big| = \big|ES(\{z\}, d_i)\big| + \big|ConS(z, d_i)\big|. \tag{4}$$

   Moreover, if z is not shaded in ${}^e d$, then, by the construction of $ext(d, D)$, z contains more existential spiders in ${}^e d$ than in d_i:

   $$\big|ES(\{z\}, {}^e d)\big| > \big|ES(\{z\}, d_i)\big|.$$

   So,

   $$\begin{aligned}\big|\Psi(z)\big| &\geq \big|ES(\{z\}, {}^e d)\big| + \big|ConS(z, {}^e d)\big| \\ &= \big|ES(\{z\}, {}^e d)\big| + \big|ConS(z, d_i)\big| \quad \text{since } \omega_{d_1} = \omega_{{}^e d}\end{aligned}$$

$$> \left|ES\big(\{z\}, d_i\big)\right| + \left|ConS(z, d_i)\right|.$$

This contradicts (4). Therefore, it must be that z is shaded in ${}^e d$. Furthermore, it can be shown that $|ES(\{z\}, {}^e d)| = |ES(\{z\}, d_i)|$. Hence

$$\forall z \in Z^*(d_i)\; z \in Z^*\big({}^e d\big) \wedge ES\big(\{z\}, d_i\big) = ES\big(\{z\}, {}^e d\big). \tag{5}$$

Since ${}^e d \not\models d_i$, by Theorem 6.1 one of the following three conditions holds.

(a$_2$) $\exists z \in Z^*(d_i)\; z \notin Z^*({}^e d) \vee ES(\{z\}, d_i) \neq ES(\{z\}, {}^e d)$.
(b$_2$) $\exists z \in Z(d_i)\; ES(\{z\}, {}^e d) \subset ES(\{z\}, d_i)$.
(c$_2$) $\exists s_i, s_j \in CS(d_1)\; \omega_{d_1}(s_i, s_j) \neq \omega_{d_2}(s_i, s_j)$.

We now consider each of these three possibilities (a$_2$), (b$_2$) and (c$_2$) in turn. Firstly, (a$_2$) contradicts (5) above, so does not hold. Secondly, (c$_2$) contradicts (3) above, so does not hold. Finally we consider (b$_2$). In the model m for ${}^e d$ we have,

$$\left|\Psi(z)\right| = \left|ES\big(\{z\}, {}^e d\big)\right| + \left|ConS\big(z, {}^e d\big)\right|.$$

Now, because m is a model for d_i we have

$$\left|\Psi(z)\right| \geq \left|ES\big(\{z\}, d_i\big)\right| + \left|ConS(z, d_i)\right|$$

from which we deduce that

$$\left|ES\big(\{z\}, {}^e d\big)\right| + \left|ConS\big(z, {}^e d\big)\right| \geq \left|ES\big(\{z\}, d_i\big)\right| + \left|ConS(z, d_i)\right|.$$

Therefore, since $\omega_{d_i} = \omega_{{}^e d}$,

$$\left|ES\big(\{z\}, {}^e d\big)\right| \geq \left|ES\big(\{z\}, d_i\big)\right|.$$

Thus

$$\forall z \in Z(d_i)\; ES\big(\{z\}, d_i\big) \subseteq ES\big(\{z\}, {}^e d\big),$$

which contradicts (b$_2$). Thus in any of the three cases, m does not satisfy d_i.

It follows that the interpretation, m, does not satisfy any unitary part of D. Therefore m does not satisfy D giving a contradiction. Hence if ${}^e d \models D$ then there exists a unitary component of D, say d_i, such that ${}^e d \models d_i$:

$${}^e d \models D \quad \Rightarrow \quad \exists d_i \in comp(D)\; {}^e d \models d_i. \qquad \square$$

Theorem 6.6 (*Completeness*) *Let D_1 and D_2 be spider diagrams with constants. Then $D_1 \models D_2$ implies $D_1 \vdash D_2$.*

Proof Suppose that $D_1 \models D_2$. Let D_1^* be the disjunctified diagram associated with D_1 given D_2. Let D_2^* be the disjunctified diagram associated with D_2 given D_1. To recap, the diagrams D_1^* and D_2^* both have the following properties:

1. they are disjunctions of unitary α-diagrams, and
2. there exists a set of zones Z and a set of constant spider labels C such that each unitary part, d_i satisfies

(a) $d_i = \bot$,
(b) $Z(d_i) = Z$ and $Z^*(d_i) = Z(d_i)$ and $S(d_i) = \emptyset$, or
(c) $Z(d_i) = Z$ and $C(d_i) = C$ and $S(d_i) \neq \emptyset$.

For each unitary part, d_1 of D_1^* obtain the diagram $ext(d_1, D_2^*)$. Since $D_1 \equiv_{\models} D_1^*$, $D_2 \equiv_{\models} D_2^*$ and $D_1 \models D_2$ it follows that $d_1 \models D_2$. Therefore, $ext(d_1, D_2^*) \models D_2^*$. Thus, each unitary part, ${}^e d_1$ of $ext(d_1, D_2^*)$ satisfies ${}^e d_1 \models D_2^*$. By Theorem 6.5, ${}^e d_1 \models d_2$, for some $d_2 \in comp(D_2^*)$. We now consider three possibilities for d_1.

1. $d_1 = \bot$. In this case, $d_1 = {}^e d$ and it is trivial that $d_1 \vdash d_2$.
2. $Z(d_1) = Z$ and $Z^*(d_1) = Z(d_1)$ and $S(d_1) = \emptyset$. In this case, $d_1 = {}^e d$. Since ${}^e d \models D_2^*$, it must be the case that some unitary part, d_2 say, of D_2^* has an empty model. In which case, d_2 does not contain any spiders and so, by the construction of D_2^*, is entirely shaded. Thus $d_2 = {}^e d$ and it is trivial that ${}^e d \vdash d_2$.
3. $Z(d_1) = Z$ and $C(d_1) = C$ and $S(d_1) \neq \emptyset$. In this case, ${}^e d \vdash d_2$ by Theorem 6.1.

In each case, we have shown that ${}^e d \vdash d_2$ and we deduce that ${}^e d \vdash D_2^*$, by Rule 9 (connecting a diagram). It follows that $ext(d_1, D_2^*) \vdash D_2^*$. By transitivity, $d_1 \vdash D_2$. Using Rule 19 ($\vee$-simplification), $D_1^* \vdash D_2^*$. Thus $D_1^* \vdash D_2$. By transitivity, $D_1 \vdash D_2$. Hence the system is complete. □

6.5 Decidability

The proof of completeness provides an algorithmic method for constructing a proof that $D_1 \vdash D_2$ whenever $D_1 \models D_2$. It is simple to adapt this algorithm to determine, for any D_1 and D_2, whether $D_1 \vdash D_2$.

Theorem 6.7 (*Decidability*) *There exists an algorithm that determines whether, for any spider diagrams D_1 and D_2, $D_1 \vdash D_2$.*

7 Implementation

We have seen that equality between spider diagrams including constants is decidable, and so it is possible to build computer-based tools that will be able to check decidability, but also which can construct equality proofs when they exist, whether automatically or with user guidance. In this short section we discuss the state of the art in implementing tools for this and other purposes.

The development of tools to support diagrammatic reasoning is well underway, and recent advances provide a basis for automated support for spider diagrams with constants. Such tools require varied functionality and the research challenges can be viewed as more broad than for symbolic logics. There are at least two major differences: first, it is more difficult to parse a 2D diagram than a 1D symbolic sentence; more significantly, when automatically generating proofs, the diagrams must be laid out in order for the user to read the proof. In respect of the second difference, possibly the hardest aspect of spider diagram layout is in the initial generation of the underlying Euler diagram. There have been many recent efforts in this regard, including [1, 5, 15, 19, 26]. Spiders can be automatically added later, as demonstrated in [17].

In terms of automated reasoning, this has been investigated for unitary Euler diagrams [24] and, to some extent, for spider diagrams, for example [7]. The approaches used rely on a heuristic search, guided by a function that provides a lower bound on proof length. Roughly speaking, the better this lower bound, the more efficiently the theorem prover finds proofs. It has been possible to produce better proof search techniques for reasoning with unitary spider diagrams [7] than for compound diagrams [6]. As was demonstrated in [25], the translation of a unitary spider diagram with constants results in (except in trivial cases), a compound diagram. So, it is highly likely to be beneficial, from an automated reasoning perspective, to develop theorem provers for spider diagrams with constants using the rules presented in this paper rather than use translations and subsequently employ theorem provers for spider diagrams. An Euler diagram theorem prover, called EDITH, is freely available for download from http://www.cmis.brighton.ac.uk/research/vmg/autoreas.htm. We note that the main goals of automated reasoning in diagrammatic systems need not include outperforming symbolic theorem provers in terms of speed; of paramount importance is the production of proofs that are accessible to the reader and it may be that this readability constraint has a big impact on the time taken to find a proof.

8 Conclusion

We have provided formal syntax and semantics for the language of spider diagrams with constants and presented a set of reasoning rules for this language. We have shown that the resulting system is sound, complete and decidable. Although the inclusion of constant spiders does not increase expressive power, we believe that if one wishes to make statements about specific individuals then it is natural to do so using constants explicitly. Thus augmenting with constants, although it brings no expressiveness benefits, is likely to increase the usability of the notation. With the reasoning rules developed in this paper, users can reason with the language when constants are included. Such reasoning systems provide an essential basis for permitting diagrams to be used for mathematical formalization and reasoning.

In the future, we plan to investigate the use of constants in notations that extend spider diagrams. These include constraint diagrams [14] and their generalizations [22]. Recent research has begun to develop a variation of constraint diagrams that is suitable for specifying and reasoning about ontologies [13, 18].

Acknowledgement This work is supported by the UK EPSRC grant "Defining Regular Languages with Diagrams" [EP/H012311/1].

References

1. Chow, S., Ruskey, F.: Drawing area-proportional Venn and Euler diagrams. In: Proceedings of Graph Drawing 2003, Perugia, Italy. LNCS, vol. 2912, pp. 466–477. Springer, Berlin (2003)

2. Clark, R.: Failure mode modular de-composition using spider diagrams. In: Proceedings of Euler Diagrams 2004. Electronic Notes in Theor. Comput. Sci., vol. 134, pp. 19–31 (2005)
3. De Chiara, R., Erra, U., Scarano, V.: VennFS: a Venn diagram file manager. In: Proceedings of Information Visualisation, pp. 120–126. IEEE Comput. Soc., Los Alamitos (2003)
4. Euler, L.: Lettres a une princesse d'Allemagne sur divers sujets de physique et de philosophie. Opera Omnia **2**, 102–108 (1775)
5. Flower, J., Howse, J.: Generating Euler diagrams. In: Proceedings of 2nd International Conference on the Theory and Application of Diagrams, Georgia, USA, pp. 61–75. Springer, Callaway Gardens (2002)
6. Flower, J., Masthoff, J., Stapleton, G.: Generating proofs with spider diagrams using heuristics. In: Proceedings of Distributed Multimedia Systems, International Workshop on Visual Languages and Computing, pp. 279–285. Knowledge Systems Institute, San Francisco (2004)
7. Flower, J., Masthoff, J., Stapleton, G.: Generating readable proofs: a heuristic approach to theorem proving with spider diagrams. In: Proceedings of 3rd International Conference on the Theory and Application of Diagrams, Cambridge, UK. LNAI, vol. 2980, pp. 166–181. Springer, Berlin (2004)
8. Gurr, C.: Aligning syntax and semantics in formalisations of visual languages. In: Proceedings of IEEE Symposia on Human-Centric Computing Languages and Environments, pp. 60–61. IEEE Comput. Soc., Los Alamitos (2001)
9. Hayes, P., Eskridge, T., Saavedra, R., Reichherzer, T., Mehrotra, M., Bobrovnikoff, D.: Collaborative knowledge capture in ontologies. In: Proceedings of the 3rd International Conference on Knowledge Capture, pp. 99–106 (2005)
10. Howse, J., Molina, F., Shin, S.-J., Taylor, J.: Type-syntax and token-syntax in diagrammatic systems. In: Proceedings FOIS-2001: 2nd International Conference on Formal Ontology in Information Systems, Maine, USA, pp. 174–185. ACM, New York (2001)
11. Howse, J., Molina, F., Taylor, J., Kent, S., Gil, J.: Spider diagrams: a diagrammatic reasoning system. J. Vis. Lang. Comput. **12**(3), 299–324 (2001)
12. Howse, J., Stapleton, G., Taylor, J.: Spider diagrams. LMS J. Comput. Math. **8**, 145–194 (2005)
13. Howse, J., Stapleton, G., Taylor, K., Chapman, P.: Visualizing ontologies: a case study. In: International Semantic Web Conference 2011. Springer, Bonn (2011)
14. Kent, S.: Constraint diagrams: visualizing invariants in object oriented modelling. In: Proceedings of OOPSLA97, pp. 327–341. ACM, New York (1997)
15. Kestler, H., Muller, A., Kraus, J., Buchholz, M., Gress, T., Kane, D., Zeeberg, B., Weinstein, J.: VennMaster: area-proportional Euler diagrams for functional go analysis of microarrays. BMC Bioinformatics **9**(67) (2008)
16. Lovdahl, J.: Towards a visual editing environment for the languages of the semantic web. PhD thesis, Linkoping University (2002)
17. Mutton, P., Rodgers, P., Flower, J.: Drawing graphs in Euler diagrams. In: Proceedings of 3rd International Conference on the Theory and Application of Diagrams, Cambridge, UK. LNAI, vol. 2980, pp. 66–81. Springer, Berlin (2004)
18. Oliver, I., Howse, J., Stapleton, G., Nuutila, E., Törmä, S.: Visualising and specifying ontologies using diagrammatic logics. In: 5th Australasian Ontologies Workshop. Conf. Res. Pract. Inf. Technol., vol. 112. CRPIT, Melbourne (2009)
19. Rodgers, P., Zhang, L., Fish, A.: General Euler diagram generation. In: International Conference on Theory and Applications of Diagrams, pp. 13–27. Springer, Herrsching (2008)
20. Shimojima, A.: Inferential and expressive capacities of graphical representations: survey and some generalizations. In: Proceedings of 3rd International Conference on the Theory and Application of Diagrams, Cambridge, UK. LNAI, vol. 2980, pp. 18–21. Springer, Berlin (2004)
21. Stapleton, G.: Spider diagrams augmented with constants: a complete system. In: Visual Languages and Computing, pp. 292–299 (2008)
22. Stapleton, G., Delaney, A.: Evaluating and generalizing constraint diagrams. J. Vis. Lang. Comput. **19**(4), 499–521 (2008)
23. Stapleton, G., Thompson, S., Howse, J., Taylor, J.: The expressiveness of spider diagrams. J. Log. Comput. **14**(6), 857–880 (2004)
24. Stapleton, G., Masthoff, J., Flower, J., Fish, A., Southern, J.: Automated theorem proving in Euler diagrams systems. J. Autom. Reason. **39**, 431–470 (2007)

25. Stapleton, G., Taylor, J., Howse, J., Thompson, S.: The expressiveness of spider diagrams augmented with constants. J. Vis. Lang. Comput. **20**(1), 30–49 (2009)
26. Stapleton, G., Rodgers, P., Howse, J., Zhang, L.: Inductively generating Euler diagrams. IEEE Trans. Vis. Comput. Graph. **17**(1), 88–100 (2011)
27. Swoboda, N., Allwein, G.: Using DAG transformations to verify Euler/Venn homogeneous and Euler/Venn FOL heterogeneous rules of inference. J. Softw. Syst. Model. **3**(2), 136–149 (2004)
28. Zhao, Y., Lövdahl, J.: A reuse based method of developing the ontology for e-procurement. In: Proceedings of the Nordic Conference on Web Services, pp. 101–112 (2003)

G. Stapleton (✉) · J. Howse · J. Taylor · P. Chapman
Visual Modelling Group, University of Brighton, Brighton, UK
e-mail: g.e.stapleton@brighton.ac.uk

J. Howse
e-mail: john.howse@brighton.ac.uk

J. Taylor
e-mail: john.taylor@brighton.ac.uk

P. Chapman
e-mail: p.b.chapman@brighton.ac.uk

S. Thompson
School of Computing, University of Kent, Canterbury, UK
e-mail: s.j.thompson@kent.ac.uk

A Practice-Based Approach to Diagrams

Valeria Giardino

Abstract In this article, I propose an operational framework for diagrams. According to this framework, diagrams *do not work* like sentences, because we do not apply a set of explicit and linguistic rules in order to use them. Rather, we become able to manipulate diagrams in meaningful ways once we are familiar with some specific practice, and therefore we engage ourselves in a form of reasoning that is stable because it is shared. This reasoning constitutes at the same time discovery and justification for this discovery. I will make three claims, based on the consideration of diagrams in the practice of logic and mathematics. First, I will claim that diagrams are tools, following some of Peirce's suggestions. Secondly, I will give reasons to drop a sharp distinction between vision and language and consider by contrast how the two are integrated in a specific manipulation practice, by means of a kind of manipulative imagination. Thirdly, I will defend the idea that an inherent feature of diagrams, given by their nature as images, is their ambiguity: when diagrams are 'tamed' by the reference to some system of explicit rules that fix their meaning and make their message univocal, they end up in being less powerful.

Keywords Diagrammatic reasoning · Practice-based philosophy of mathematics · Peirce's diagrams · Manipulative imagination · Productive ambiguity

Mathematics Subject Classification (2010) Primary 00A66; Secondary 03A05 · 97C30

1 Introduction: Diagrammatic Reasoning in Logic and Mathematics

The claim that diagrammatic reasoning has a role in the work of a logician or a mathematician, both in education and research, is not controversial. Even a scholar who defends a conception of logic or mathematics as mere games of symbols would not deny that objects such as circles or figures are helpful as heuristic tools in directing reasoning. A common saying is that in problem solving, if we draw the 'right' diagram, then we are half of the way to finding its solution. Diagrams seem to guide thought: their heuristic power is evident, both in the classroom—where concepts and theorems are often taught by making reference to objects such as sketches on the blackboard—and in research—when logicians and mathematicians interact and communicate by holding a pen in their hands, drawing, changing and deleting columns of formulas or shapes.

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic,
DOI 10.1007/978-3-0348-0600-8_8,

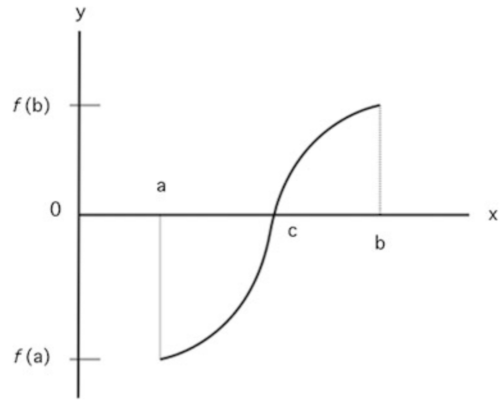


Fig. 1 A figure displaying the *Intermediate Zero Theorem*

Consider for example the case of the *intermediate zero theorem*.[1] According to this theorem, if a function f is continuous on the interval $[a, b]$ and f changes sign from negative to positive (or vice versa), then there is a c between a and b such that $f(c) = 0$. The figure displaying such a function is shown in Fig. 1.

By looking at the figure, we are inclined to believe that the theorem holds, because we can clearly 'see' it. The line of the argument would be the following: the horizontal axis divides the space on the right side of the vertical axis into two regions; in the region above the horizontal axis, the sign of the function is positive, while in the region below the horizontal axis, the sign of the function is negative. The function f would then be described as performing the action of *moving* from $f(a)$ to $f(b)$: if the function changes sign from negative to positive (or vice versa), then it *has to go* from below the horizontal axis to above it (or vice versa). The figure is straightforward: we have the impression that our reasoning corresponds to what we 'see'.

Nevertheless, a certain conception of logic and mathematics, by subscribing to the old distinction between a context of justification and a context of discovery, has relegated our natural disposition to use figures to display the content of our statements to the psychological context of discovery. This conception has discarded figures and diagrams as not being rigorous. I will call this view the *suspicious* view, both because it denies that diagrammatic reasoning plays a role in the context of justification, and because I want to put this approach into question. I mention here that I deliberately use the term 'suspicious' ambiguously; as I will show later in the article, ambiguity will become important in connection to diagrams.

According to the suspicious view, diagrams, despite their being good heuristic tools in discovery and explanation, are not sufficient when it comes to proof. The view does not directly deny the cognitive and computational advantages in using diagrams, but rather puts forward an image of logic and mathematics that is strongly dominated by proof. Proofs, in turn, are conceived as particular syntactic objects, namely derivations and thus verbal/symbolic entities. If formal proofs are at the core of mathematics, then diagrams cannot be part of them, unless they are 'tamed' by the definition of some syntactical rules that control their use. In such a framework, justification is defined by proof and rigor, and they alone lead to truth: mathematical knowledge happens only when we are *truly*

[1]The theorem is mentioned, among others, in [4].

justified in our belief, namely, when we have a formal proof, and not when we are merely justified in believing that a statement is true, as in the case of Fig. 1 for the intermediate zero theorem. As Fallis explains,

> even though there are many ways in which a mathematician might be justified in believing that a mathematical statement is true, there is only one way in which a mathematician feels that she is truly justified. Specifically, the mathematician has to know a proof of the mathematical statement ([10], p. 46).

Polya, who famously devoted most of his work to the investigation of problem solving processes, claims that what is lost in rigor is made up in understanding [22]. Insight, understanding, and explanation, which are part and parcel of the work of the logician or of the mathematician, are not all necessarily included in a rigorous proof.

If mathematics is dominated in such a way by formal 'non-visual' proofs that preserve truth, all so called 'visual' proofs are discarded as non-rigorous. Diagrams are not reliable: they can mislead us and do not provide evidence. According to the standard and logocentric definition, proofs are

> syntactic objects consisting only of sentences arranged in a finite and inspectable way, and therefore diagrams can only be heuristic tools to prompt certain trains of inference ([2], p. 3).[2]

Against this assumption, Barwise and Etchemendy tried to expand formal logic by freeing it from having one mode of representation only, i.e. language, and by pushing it closer to reasoning, which is in their view a *heterogeneous* enterprise. As Shin sums up,

> all of us engage in and make use of valid reasoning, and in the process of reasoning human beings obtain information through many different kinds of media, including diagrams, maps, smells, sounds, as well as written or spoken statements ([26], p. 92).

However, Barwise and Etchemendy, as well as Shin later on, did not renounce the idea that proofs are at the core of logic and mathematics and that proofs are derivations. Therefore, their strategy was to force and 'tame' diagrammatic reasoning by defining syntactic rules to carry out diagrammatic manipulations: in their view, only 'tamed' diagrams, namely diagrams which we have control over, can become 'rigorous' elements of formal proofs.

By contrast, let us suppose that an alternative picture of logic and mathematics is possible. This picture does not have at its core formal proofs, but rather the practices that are shared by a community of scholars who in their ordinary work consider informal proofs, such as proofs based on induction or visual tools, as sufficient for being justified in believing that some statement holds. As Brown suggests, we should assume

> a somewhat more humble attitude toward our understanding of verbal/symbolic reasoning. First-order logic may be well understood, but what passes for acceptable proof in mathematics includes much more than that ([4], p. 164).

Moreover, as discussed by Mancosu, there are connections in mathematics that count as *necessary* and *truth preserving*, and others that are not held as such but nevertheless can count as *reasons* [17]. Rigor might preserve truth, but insight, understanding, and explanation hint at reasons. Furthermore, there are proofs that explain—which might be referred to as *causal* proofs—and proofs that convince but do not explain—which can

---

[2]Barwise and Etchemendy quote this passage from an article by Tennant (see [28]). They believe it expresses the 'dogma' of logocentricity that they want to challenge.

be referred to as *non-causal* proofs [3]. If we look at the practice of mathematics, we realize that verification is proof, but verification might not provide reasons: mathematicians are not satisfied with proving conjectures, since what they want is *reasons* for these conjectures [23].

Views such as the suspicious view, focusing on rigorous formal proofs, move away from the consideration of actual mathematical practice, both in contemporary mathematics as well as in the history of mathematics. According to Corfield, they apply what he calls the 'foundational filter'; because of this filter the only interesting questions in philosophy of mathematics are about the possible reduction of mathematics to some foundational system [7]. Corfield believes that the foundational filter is an "unhappy idea", because not only does it fail to detect the pulse of contemporary mathematics but it also "screens off the past to us as not-yet-achieved" ([7], p. 8). Mathematics is a human activity and, consequently, it is situated in time. The questions about mathematical practice "are to be addressed by an understanding of mathematical knowledge as historically situated rather than timeless" ([7], p. 15). The suggestion is to get free of the appeal to a timeless logic or mathematics that has inspired the assumptions behind the suspicious view. An interdisciplinary investigation may be helpful

> in the process demonstrating that philosophers, historians and sociologists working on pre-1900 mathematics are contributing to our understanding of mathematical thought, rather than acting as chroniclers of proto-rigorous mathematics ([7], p. 8).

The suspicious view discards as 'psychological', and therefore philosophically uninteresting, not only diagrammatic reasoning but also many other activities of the community of working logicians and mathematicians.

In this article, I will not give any normative criteria for the definition of what a proof should be. More modestly, I will simply claim that the search for the reasons why diagrams are apparently so effective in explanation and discovery is a philosophical issue. I will propose an *operational* framework, within which I will show that diagrams *do not work* like sentences: in fact, we do not necessarily apply a set of explicit and linguistic rules in order to use them. Rather, once we are familiar with some specific practice, we manipulate diagrams in meaningful ways, engaging ourselves in a form of reasoning that is stable because it is shared by the community and thus constitutes at the same time discovery and justification for that discovery. If this kind of operational framework works for diagrams, then a further issue will be to ask whether the same operational framework can be applied to other kinds of activities, and thus be generalized to a practice-based approach to logic and mathematics in general. This is a matter for further research.

In the following sections, I will make three claims based on the consideration of diagrams in the practice of logic and mathematics. First, I will claim that diagrams are tools and I will define what I intend by 'diagram' and by 'tool', following some of Peirce's suggestions. Secondly, I will give reasons to drop the opposition between vision and language, and consider by contrast how the two are integrated in a specific manipulation practice by means of a kind of *manipulative* imagination. Thirdly, I will defend the idea that an inherent feature of diagrams, given by their nature as images, is their ambiguity. Moreover, ambiguity promotes a wider variety of interpretation and understanding: when diagrams are 'tamed' by way of referencing to some system of explicit rules that fix their meaning and make their message univocal, they end up in being less powerful.

2 Diagrams in the Practice

2.1 Diagrams Are Not Pictures of Abstract Objects

I propose an account of reasoning *in* and *from* diagrams based on the conception of diagrams as tools used within a specific practice. First, I will claim that if diagrams are considered as tools, then old problems that lie behind the suspicious view lose their strength. Secondly, I will clarify what I intend by 'diagram' and by 'tool'.

Traditionally, two problems have been put forward to conclude that figures are not sufficient for providing justification of mathematical statements. I will define these problems (i) the *generality* problem and (ii) the *appropriateness* problem. The problems state that:

(i) it is not possible to get to a general conclusion by looking at some particular diagram, for the reason that *that* particular diagram has specific properties and specific features that do not directly depend upon the statement that is to be proven: to check for the truth of the statement, a single diagram should be capable of representing *all* the possible specific ways in which the situation described by the statement can be true, which is impossible;

(ii) diagrams are most of the time inappropriate for proving the statement in question because they are never precise enough: these imprecisions can occur in our reasoning and thus bring us to false conclusions and misinterpretations.

For example, in the case of the intermediate zero theorem, (i) we cannot be sure that the function f would *in any case* behave as depicted in Fig. 1 and (ii) we cannot be sure that we have properly drawn the figure.

Nevertheless, the assumption behind the claim that diagrams are never sufficiently general and never sufficiently appropriate is that diagrams are depictions—though partial and imprecise—of abstract objects. I want to contend this claim and propose that diagrams are not pictures of abstract objects but tools for reasoning about abstract relationships. Before doing that, I will present Brown's view of diagrams as tools and explain why it is different from the view I propose.

Brown's strategy for refuting problems (i) and (ii) is to subscribe to a Platonist view of abstract objects. According to him, figures are not and cannot be representations in the sense of being pictures. His argument is that if they were such representations, then there would be a kind of structural similarity captured by the concept of *isomorphism* between them and the abstract objects they depict. Nonetheless, despite the fact that in a wide variety of cases a good diagram is isomorphic to the situation it represents, this is not always the case.

Take for example the following number theory result on the sum of natural numbers up to n:

$$1 + 2 + 3 + \cdots = \frac{n^2}{2} + \frac{n}{2}$$

Now consider one of the possible figures that can be used to represent this result (Fig. 2).

The area of the figure in Fig. 2 is composed by 1 square plus 2 squares plus 3 squares plus ... n squares. Thanks to the arrangement of the squares, the area is *also* equal to $n \times n$ squares divided by 2 (the area of the white isosceles triangle with base and height

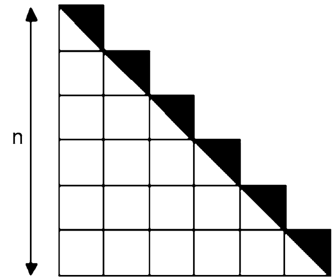


Fig. 2 A figure displaying the sum of n natural numbers

equal to n) plus half n (the area of the black triangles). Therefore, the figure shows that the equation holds: the same diagram that represents the right side of the equation also represents its left side.

However, Brown claims, the figure in Fig. 2 is strictly speaking just a 'picture', an illustration of the $n = 6$ case, and as a consequence

> we can claim that there is an isomorphism to some number structure with that cardinality. It is certainly not, however, isomorphic to all the numbers. True, it is homomorphic to the whole number structure. But note that a homomorphism to a larger structure is (at least in the case at hand) an isomorphism to a part ([4], p. 173).

Therefore, the diagram in Fig. 2 tells us something about the part of the structure it is isomorphic to, but nothing about the rest of the same structure. Brown's proposal then is to deny that figures such as the one in Fig. 2 are pictures at all, and assume that, in his words, "*some 'pictures' are not really pictures, but rather they are windows in Plato's heaven*": figures are tools in the same sense in which telescopes are tools for the unaided eye. For this reason, the problems (i) and (ii) do not apply, since diagrams are not depicting any possible situation and to some extent they are always inappropriate. In a nutshell, in Brown's view, it is possible to be a realist about abstract objects without being a realist about pictures [5].

Brown's solution is thus based on realism about abstract objects. My objection to his view is that it explains what is already mysterious—our common way of referring to figures in our reasoning—by means of even more mysterious entities: abstract objects in a Platonic heaven. Moreover, as Folina claims,

> telescopes are not *themselves* justificatory: it is not the telescope which is cited as the primary justification for an astronomical claim. Similarly for windows. . . . Rather, it is merely a tool which enables us to 'see' the evidence, or the abstract 'picture' ([11], p. 426).

According to Folina, Brown endorses Platonism in order to claim that figures can legitimately prove the truth of mathematical statements, but this is a non-starter. In fact, we can well accept the assumption that diagrams cannot be a part of formal proofs; the point is rather to claim that formal proofs are only a proper subset of a variety of justifications in mathematics. As in the criticism of Barwise and Etchemendy's program, the real challenge is not to make diagrams legitimate components of formal proofs, but rather to give an account of how they belong to other kinds of justifications in the variety. In this respect, Brown's view is not helpful.

I will side with Brown and agree with him that diagrams are not pictures of abstract objects but instruments, and therefore problems (i) and (ii) do not apply. Nevertheless, I will not claim that they are telescopes pointing at a Platonic heaven, because we are not sure of the existence of such a heaven. In the next section, I will discuss what I intend by saying that diagrams are tools.

2.2 Diagrams as Tools

For the purpose of this article, I use the label 'diagram' as a very general term to refer to two-dimensional displays.[3] I move within a Peircean perspective, and consider diagrams as instruments for thought along the following lines.

First, in order to be effective, diagrams must always be *interpreted* within a certain *context of use*. Secondly, the reference to them contributes to the very definition of this context and gives structure to the problem to solve. Thirdly, diagrams belong to the genus of 'representation' as "that character of a thing by virtue of which, for the predication of a certain mental effect, it may stand in place of another thing" ([13], vol. 1, par. 564; written in 1893). Nevertheless, this should not be taken literally, for diagrams do not 'directly' depict some abstract object whose existence is presupposed; rather, they embody a selection of *relevant* relations. Finally, diagrams are given *with an intention*, as all tools are, cognitive and epistemic tools being among them: they are conceived so as to achieve some particular aim, and the intention behind their creation must be acknowledged in order to appropriately interpret and use them.

Therefore, some of the physical features of a diagram refer to abstract elements that are not directly present in the diagram. By manipulating these physical features, the user—the 'interpreter'—learns or genuinely discovers something new about the relations the diagram embodies. As Peirce rightly pointed out, this *dynamic* aspect of diagrams triggers "a state of activity" in the interpreter that leads to experimentation.[4] Diagrammatic reasoning would then bring logic and mathematics closer to the natural sciences: logicians and mathematicians experiment with the very same representations that constitute their instruments. Peirce goes even further by saying "all necessary reasoning without exception is diagrammatic" ([13], vol. 5, par. 162; written in 1903). Once more, I will not take any stance in the debate on what counts as necessary reasoning. My more modest suggestion is that in order to claim that diagrams are stable enough to provide justification, we have to consider the practice shared by the community of actors who experiment on them. Diagrams are representations used with the intention of embodying relations; moreover, they promote inference because they can be interpreted and manipulated in various ways according to the shared practice.

An important advantage of this operational approach is that it discards the opposition between visual reasoning and linguistic knowledge. In fact, the dichotomy of vision vs. language, which has led to the antithesis between visuocentric and logocentric views, is pernicious.

---

[3] I am not denying here the possibility that there are three-dimensional diagrams. I only want to exclude this possibility for the moment, because I am inclined to think that it implies additional considerations.

[4] "It is not, however, the statical Diagram-icon that directly shows this; but the Diagram-icon having being constructed with an Intention [. . .]. Now, let us see how the Diagram entrains its consequence. The Diagram sufficiently partakes of the percussivity of a Percept to determine, as its Dynamic, or Middle, Interpretant, a state [of] activity in the Interpreter, mingled with curiosity. As usual, this mixture leads to Experimentation." In [21].

In the next sections, I will first discuss the risks of siding with or against vision or language without considering their continuous interaction; secondly, I will discuss two features of diagrammatic reasoning that emerge in this operational framework: the role of action and manipulative imagination, and the importance of ambiguity and multi-dimensionality of meaning.

3 Beyond Visuocentric and Logocentric Views

3.1 Not Only a Question of Visual Properties

Let us consider a first risk in endorsing the dichotomy of vision vs. language. Suppose that we want to refute logocentrism and defend a view diametrically opposed to it, a kind of 'visuocentrism' or 'optocentrism'. According to this approach, it suffices to look at a figure to get to its content and to the message it conveys; the meaning of a diagram, compared to the meaning of a sentence, would be easier to grasp because it is 'directly seen' and therefore extracted 'for free'.

Nevertheless, this view is ill posed, because it is not only visual perception that is at stake in diagrammatic reasoning. In fact, the diagram user—its interpreter—is not interested in many of the visual properties of the diagram; she only attends to a selection of them. For this reason, talking of visual thinking is misleading; the user rather considers possible spatial configurations and new constructions in the diagram.

Let me discuss one attempt to enhance the visual features of a diagram with the aim of making its meaning easier to grasp. Consider Byrne's edition of Euclid's *Elements* published in 1847 [6]. This edition covers Euclid's first 6 books and presents them in form of colored pictures, using as little text as possible, most of the times in the form of labels: Byrne's attempt was to make the *Elements* 'more visual' by adding colors to them. Nevertheless, this attempt is unsuccessful, if the aim is to make the Elements easier to understand. In fact, colors are introduced as a new code that must be learned in its symbolic use. As a consequence, the figures, instead of becoming more straightforward, end up becoming more complex: the student must learn how to interpret colors, and therefore she has to become familiar with a completely new language, the color code language, thus increasing the cognitive load of the task instead of reducing it (see Fig. 3). Furthermore, she is deprived of the instructions given in the text to construct the diagrams, and is therefore driven away from learning the practice of Euclidean geometry.

To make Euclid diagrams easier to understand, attention should be focused on constructions and manipulations rather than on possibly new visual features. This example shows that diagrams are not simply 'seen' but must be 'read', i.e., interpreted. Their appropriate interpretation leads to the definition of their constructions and manipulations. Therefore, the assumption of a sharp distinction between vision and language along with the claim that vision matters does not give a good description of what happens in diagrammatic reasoning.

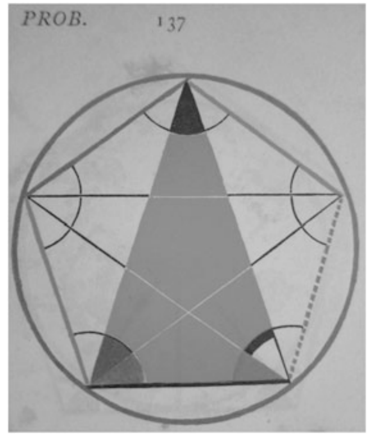


Fig. 3 A picture taken from Byrne's edition of Euclid's *Elements*

3.2 *Not Only a Question of Expressivity*

In the previous section, I claimed that diagrams are advantageous not simply because they are more 'visual' than sentences. In the same way, if the opposition between vision vs. language is assumed, not even a focus on the possible verbal/symbolic translations of the message conveyed by the diagram will yield a good description of what happens in diagrammatic reasoning.

Consider a case study from logic, which shows the contrast between the urge for rigor and the cognitive benefit of the diagrammatic representation.[5] In the 18th century, the mathematician Euler introduced his circles (D1) for the study of syllogisms. As Shin and Lemon explain, the representation in Euler's diagrams is governed by the convention according to which every object x in the domain is assigned a unique location in the plane, say $l(x)$, such that $l(x)$ is in the region R if and only if x is a member of the set the region represents [27]. Note that, despite its apparent naturalness, this move is already conventional since the choice of circles is arbitrary. Other logicians introduced systems that used points for objects and lines for sets [8, 15]. On the other hand, this convention exploits better than others the perceptual configuration that Lakoff and Nuñez have defined as the *Containment-schema*: circles and in general closed figures are more effective than lines in being interpreted as 'containing' the members of a set in the spatial region they identify [14].

Despite their straightforwardness, Euler circles have expressive limitations and retain some crucial ambiguities, for example when representing existential statements, or the empty set, or congruency among sets. In order to solve some of these ambiguities or limits, in 1881 Venn introduced his own system of diagrams (D2) based on 'primary diagrams'. Primary diagrams do not carry any particular information in themselves; in order to be meaningful, they need to be complemented by labels and shadings. Despite their greater generality, Venn diagrams present new expressive limitations, and for this reason Pierce, in the 20th century, modified them by introducing three new symbols (0, x, –), thus providing a new system (D3) by which existential statements, disjunctive information, probabilities, and relations could be represented (see Fig. 4).

[5] For a detailed discussion of this case-study, see [27].

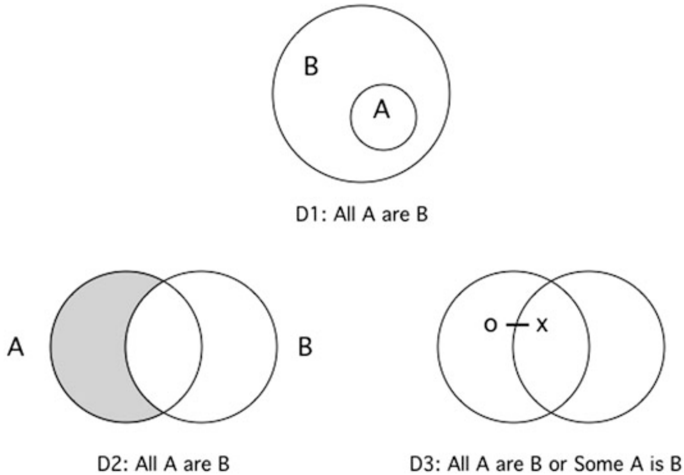


Fig. 4 Three examples of circles for syllogism: Euler (D1), Venn (D2) and Peirce (D3)

In going from D1 to D3, the new systems become thus more and more expressive. In the 90s, Shin continued on this path and proposed a formal system that improved Peirce's diagrams by making them even more expressively powerful [25].

Nevertheless, as Shin and Lemon discuss, Peirce's (as well as Shin's) introduction of new conventions increased the expressive power of the single diagrams at the expense of the visual clarity that Euler's original system enjoys. The new conventions are more arbitrary and the new representations more confusing. This is true, despite the fact that Peirce's choice of the symbol '0' for the empty set is not wholly arbitrary. According to the authors, when Peirce's revision was completed, most of Euler's original ideas about visualization were lost, except for the choice of a geometrical object, the circle, still used to represent (possibly empty) sets. I would slightly modify their claim and say that what was retained in D3 was the mere *appearance* of the circles of the original diagrammatic system, but not their *function* of exploiting the Containment-schema. As Euler himself claimed, "we may employ, then, spaces formed at pleasure to represent every general notion, and mark the subject of a statement by a space containing A, and the attribute by another which contains B."[6] These are the 'spaces' that mattered for him and his system, and not the particular figure they were of. Circles could have been changed into squares, for example, without changing their function of 'containment' and, therefore, without changing the expressive power of the system. By the introduction of primary diagrams, this function of 'containment' loses its centrality, and new elements such as shadows and labels are introduced in order for the system to be meaningful. The strategy of augmenting the expressivity of diagrams by fixing their meaning through the introduction of these arbitrary conventions provides an interesting 'diagrammatic' extension of sentential logic, but at the same time deprives spatial tools of their effectiveness and straightforwardness, and leaves behind our perceptual and imaginative capacities.

Consider now the fascinating collection of 'visual proofs' that Nelsen gives in his two books, entitled *Proofs Without Words* (*I*, *II*) [18, 19]. On closer inspection, these proofs are

[6] Letter 103, *Of Syllogism, and their different Forms, when the first Proposition is Universal*. See [9].

not simply 'without words', since the use of a diagram is not only a matter of vision. As in the case of the intermediate zero theorem, we have to know the mathematical statement in question in order to find it 'in' or 'represented by' the diagram in Fig. 1. Without knowing the statement, reasoning with a diagram would be equivalent to a riddle such as 'From this figure, find the statement in it' or 'Which statement is represented by the figure?'. We always need some linguistic explanation or justification for our use of some specific diagram.

More generally, in practice, it is very difficult to distinguish purely graphical systems from purely sentential ones [24]. In most cases, diagrams and text are so strongly interconnected that they cannot be considered in isolation one from the other; furthermore, thanks to this dialectic, appropriate interpretations and operational procedures are defined.

Let me point out to two examples in which the dialog between images and text have been challenged by the invention of printing. First, consider Diderot and D'Alambert's *Encyclopédie*, published in France between 1751 and 1772. The images and text in it were conceived as interconnected and continuously referring one another; nevertheless, because of the limitations of the printing techniques available at the time, it was impossible to show on the same page the labeled images and the corresponding text. In many cases, the text was very far from the images it referred to. For this reason, the reader had to move from the images to the text without having both together; this was very awkward and contributed to the sharp distinction between images and text, which does not reflect the way the two were conceived at the beginning as being strongly interconnected.

Something similar happened for the first printed editions of Euclid's *Elements*, where it was even more crucial to show images and text on the same page [1]. In Euclid's *Elements*, we never find a simple two-step process from the verbal to the visual moment that is a mere illustration of the former. By contrast, the text and its accompanying figure are engaged each time in a fruitful and rational intercourse: the text gives the instructions to outline the figure that stands by its side, and to draw the conclusion from it. The verbal instructions do not assume the complete figure at the start, but rather walks the readers through its construction, and therefore the readers continuously go back and forth between verbal and visual. In following these instructions, they *actively imagine* drawing lines.

Instead of accepting the unuseful opposition between visual perception and linguistic knowledge, I propose to focus on this kind of 'manipulative' imagination at play when reasoning with diagrams. This is further discussed in the next two sections.

4 An Operational Framework for Diagrams

4.1 Looking at How We 'Act' upon the Diagrams

Diagrams are always considered from within a specific practice and context. Imagine the diagram of a circle. It has been used in logic, for example, by Euler, in Euclidean geometry, and in Cartesian geometry. Was it 'the same' circle in all these cases? Did it play the same role? Of course it did not. In fact, there is nothing like a specific set of rules that could be fixed once for all for all circles or for some specific kind of diagram in general. My proposal is that in diagrammatic reasoning what counts is not the appearance of a diagram and a list of explicit rules that can be applied to it, but rather a set of *procedures*:

when one learns to use a certain diagrammatic system for performing some inferences, she learns a *manipulation practice*. The diagram becomes the mathematician's worksite, where operations, plans, and experiments are made in order to find solutions and reasons for these solutions. While syntactic rules are piecemeal, procedures are holistic.

From this perspective, when drawing a diagram, the user never reproduces it *mechanically*. To go back to the intermediate zero theorem, in order to draw the diagram in Fig. 1 not only does she need to 'see' it, but also to appropriately interpret it, understand what it means in *that* context, and finally learn its construction. This operational aspect of working with diagrams has been neglected because of the suspicious view, thus encouraging a straightening of the opposition between vision and linguistic processes; this opposition has left out the consideration of the continuous interaction between the two formats in reasoning.

In the framework I propose, it is the operational aspect of working with diagrams, namely what is *done* with them, that must be taken into account and not their visual features nor the sentential information they can possibly carry. Diagram-users share something like the *experience* of seeing in a diagram what they *have to*, focusing their vision on a selection of relevant features, which will bring them, as rational agents, to understand and *reproduce* the relevant features of diagrams in a non mechanical way and without 'damages'. The antidote to the suspicious view consists thus neither in assuming a visuo-centric view nor in fixing a set of syntactic rules, but in considering how diagrams are manipulated, in continuous interaction with language and within a specific practice, in order to infer some new conclusion.

My view moves from a purely syntactic approach to a semantic and, indeed, pragmatic approach to problem solving. As Grosholz claims in regard to the epistemology of mathematics, there are reasons for pursuing this approach and even considering the use of language in terms of its representational role in an historical context [12]. Diagrams and figures are inherently ambiguous: the operations on them are what fixes their meaning.

For example, consider someone who is inside the practice of Euclidean geometry. Without learning *in advance* any explicit rule, she perfectly knows that once she has constructed the figure, she is allowed to rotate or translate it, but she cannot for example stretch it differently from what can happen in other practices. This practice-based framework for diagrammatic reasoning presupposes the centrality of a form of *manipulative imagination* we refer to in our manipulation practices in logic and mathematics. This imagination is particularly effective because we are already familiar with it, since it derives from our perceptual experience and reproduces procedures that are to some extent similar to the manipulation of concrete objects. Yet, not all possible manipulations are allowed by the practice: among all the possible moves, only some of them are accepted.

Diagrams are cognitively advantageous because their use activates this manipulative form of imagination, as in the case of the Containment-schema for Euler circles, which is of course also informed by the context. The manipulations that are actively imagined on the diagram are controlled by interpretation and by the shared practice: diagrams do not offer a single message, but can be interpreted differently and, based on the interpretation, different actions can be performed on them in order to discover new relations. For this reason, one aspect of diagrams that becomes important is their inherent ambiguity: they can be 'read'—or interpreted—in different ways, and the practice of their manipulation and the procedures applied to them fix their meaning. Ambiguity is thus not a disadvantage in principle, but one of the strengths of diagrammatic reasoning: their multi-dimensionality

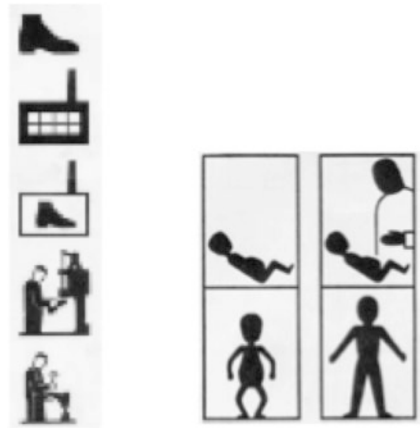


Fig. 5 Examples of pictures from Neurath's *Isotype* (taken from [20])

of meaning and non-unique interpretations can promote inference. In the next section, I will focus on this issue.

4.2 Diagrams Are (Hopefully) Ambiguous

I will first present an example to show that diagrams are ambiguous as are all images. Secondly, I will discuss how this ambiguity can be of help in mathematics by promoting inference.

Diagrams neither directly speak to the eyes nor convey a single message since, analogously to other images, they are inherently ambiguous. I want to mention here a particularly straightforward case that shows the ambiguous nature of images.

In the 1930s, the Austrian philosopher Neurath introduced *Isotype* (International System of Typographic Picture Education) to the aim of offering a tool that, in his view, could have solved the problems in communication caused by different levels of education among people, thus allowing free discussions of common problems and the dissemination of simple but important facts [20]. Isotype was meant to be a new way of conveying information that is at the same time easy to teach and learn, and is comprehensive and exact. It included a special dictionary and a special visual grammar, which, according to Neurath, created a new visual world analogous to the word world. As Neurath explains,

> the first step in Isotype is the development of easily understood and easily remembered symbols. The next step is to combine these symbolic elements.

In his view, icons such as the ones depicted in Fig. 5 are very easily recognized, no matter what the level of education is, and at the same time they express very complex situations.

Nevertheless, there are several problems in trying to provide a purely iconic system. First, a purely iconic system cannot easily convey changes of situations or convey that which contradicts some previous iconic representation. Let us suppose for example that at some point someone in our community decides that taking one's child for medical check-ups is not necessary for a child's health (in opposition to what the icons in Fig. 5 prescribe) and she shows this in Isotype by drawing the figure in Fig. 6.

Fig. 6 An example of a new (and ineffective) picture in *Isotype*, reworked from Fig. 5

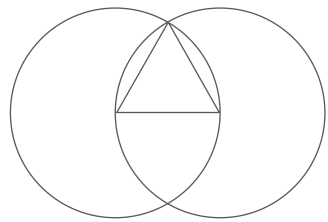


Fig. 7 The construction of an equilateral triangle starting from a given finite straight line

This time, the icons are evidently not as effective as in the previous case (i.e. as in Fig. 5). In fact, it is impossible to display *by them* the absence of the doctor or of anyone or anything else at all. This point expresses the difficulty in general of representing negation without using a specific symbol for it. Note that this impossibility is not too far from what we have seen in Euler circles' incapacity of representing empty sets.

Neurath's Isotype is based on the bad presupposition that there is in fact a distinction between a visual grammar on the one hand and a linguistic grammar on the other. By contrast, icons are not directly visual: each of them needs to be interpreted, to some extent by convention. To understand Isotype, the reader needs some background knowledge and, in the end, education. The same criticism can be formulated for any form of 'pictionary' that claims to be an universal dictionary.[7] Here we have a case similar to the 'colored' Euclidean geometry considered earlier: making it 'more visual' does not mean making it more straightforward. On the contrary, arbitrary conventions are introduced.

The above example shows that diagrams, analogously to other images, are inherently ambiguous since they do not convey one singular message that can be translated into language, but need interpretation in order to be understood. Nevertheless, in most cases, this is not a disadvantage but a strength.

Take the example of Euclid again, in Macbeth's reconstruction. Consider Book I, Proposition 1:

> *On a given finite straight line to construct an equilateral triangle.*

As Macbeth explains, the problem posed by this proposition is a *construction* problem. To get to the solution, it is necessary to reason *in* the diagram depicted in Fig. 7 [16].

[7] Still today, there are attempts to take this path, such as the pictionary 'Point it: Traveler's Language Kit', by Dieter Graf.

We start from the given finite line. Then, thanks to the construction of the two circles, we obtain a new figure. In order to get to the solution of the problem, the interpreter has to reason at different stages. At one stage, she must regard the lines in the diagram as the *radii* of the two circles: from this configuration, she concludes that all three lines are equal in length. At a second stage, she must regard the *same* lines as the *sides* of a triangle: from this new configuration, she concludes that the constructed figure is the desired equilateral triangle. Therefore, the diagram is in some way ambiguous, because it can be regarded as attending to different configurations. The interesting aspect is that all these possible configurations are compatible, and their compatibility is given by the fact that all configurations belong to the very same diagram.

As Macbeth explains, "one and the same lines are *now* regarded as parts of a circle and *later* as parts of a triangle" ([16], *italics mine*). In this case, there are three discernible levels of articulation: (i) primitive parts (points, lines, angles, and areas); (ii) geometrical figures (circles, triangles, and squares); and finally, (iii) the whole diagram. To find the solution of the problem stated by Euclid, one has to consider all these three levels together, precisely by exploiting the multidimensionality of meaning of the diagram. The generality of the diagram is given by the fact that it is constructed following the instructions in the text and is manipulated by imagination; the diagram is reliable because its use is considered inside a specific practice, Euclidean geometry, and its interpretation is intertwined with the rest of the shared system of knowledge and procedures that pertain to Euclidean geometry and serves as a guarantee for the correctness of the reasoning. The diagram is thus not a picture, but an instrument that promotes inference thanks to its possible manipulations. Moreover, the diagram does not need to be properly drawn, as long as the user is aware of the prescriptions contained in the instructions for its construction and is aware of its intended meaning.

We have seen in Fig. 2 something similar to the case described by Macbeth. In order to use the diagram for the sum of natural numbers up to n, we have to regard it *now* as a collection of squares going from 1 to n, corresponding to numbers from 1 to n, and *later* as a collage of triangles, the isosceles triangle with n as short sides, and the n half squares. The two configurations belonging to the *whole* diagram brought us to our conclusion. Moreover, consider Fig. 1 and the case of the intermediate zero theorem. In our linguistic description, we normally say that the function *goes* from below the horizontal axis to above it; we do that because we metaphorically reproduce its construction in our imagination.

My hypothesis is that, inside a specific practice, the space of the diagram perceived combines with the actions actually performed or imagined on it, in continuous interaction with linguistic knowledge. All these elements together contribute to mathematical meaning-making: manipulative imagination is at work to provide evidence in favor of some particular train of thought.

5 Conclusions

In this article, I tried to give arguments in favor of an operational framework for diagrammatic reasoning based on the practice of logic and mathematics. First, I presented what I defined as the suspicious view, according to which diagrams are not reliable enough to count as evidence for a conclusion. I claimed that this view is heavily based on a conception of proofs as syntactic objects and derivations, and I defended the idea that justification in practice is much more than proof only. Moreover, taking a Peircean perspective, I claimed that diagrams are a very special kind of representation, which is dynamic and needs an interpreter. According to the framework I propose, in order to give an account of diagrammatic reasoning it is necessary to focus on the practice shared by the community and on the actions performed on the diagram while considering two main aspects: (i) manipulative imagination; (ii) the role of ambiguity in triggering this imagination.

The dichotomy between visual thinking on the one hand and linguistic processes on the other has obscured the fact that what counts in diagrammatic reasoning is the manipulation practice, based on holistic procedures and not on the definition of explicit linguistic rules. According to this practice-based framework, it is the practice that fixes the meaning of the diagrams on each occasion, otherwise they are as ambiguous as other images are, and it is a kind of manipulative imagination that operates on them. Each practice is defined by procedures of manipulations and interconnected facts. All these elements taken together define in turn the system of knowledge shared by the community; this system encompasses diagrams, statements, particular notations and actions prescribed or allowed on them.

Acknowledgements I want to thank Mario Piazza, Achille Varzi, Roberto Casati, and two anonymous referees who gave me very useful suggestions in order to improve this article. Many thanks to Christopher Whalin for having proof-read the final version. Special thanks to the editors, Amirouche Moktefi and Sun-Joo Shin, for their careful work.

References

1. Baldasso, R.: Illustrating the book of nature in the Renaissance: drawing, painting, and printing geometric diagrams and scientific figures. PhD thesis (2007)
2. Barwise, J., Etchemendy, J.: Visual information and valid reasoning. In: Allwein, G., Barwise, J. (eds.) Logical Reasoning with Diagrams, pp. 3–25. Oxford University Press, London (1996)
3. Bouligand, G.: Premières leçons sur la théorie générale des groupes. Vuibert, Paris (1932)
4. Brown, J.R.: Proofs and pictures. Br. J. Philos. Sci. **48**, 161–180 (1997)
5. Brown, J.R.: Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures. Routledge, London (1999)
6. Byrne, O.: First Six Books of the Elements of Euclid, in Which Coloured Diagrams and Symbols Are Used Instead of Letters for the Greater Ease of Learners. William Pickering, London (1847)
7. Corfield, D.: Towards a Philosophy of Real Mathematics. Cambridge University Press, Cambridge (2003)
8. Englebretsen, G.: Linear diagrams for syllogisms (with relationals). Notre Dame J. Form. Log. **33**(1), 37–69 (1992)
9. Euler, L.: Letters of Euler to a German Princess: On Different Subjects in Physics and Philosophy (trans.: Hunter, H.). Thoemmes Continuum, London (1997)
10. Fallis, D.: Intentional gaps in mathematical proofs. Synthese **134**(1–2), 45–69 (2003)
11. Folina, J.: Pictures, proofs, and 'mathematical practice': reply to James Robert Brown. Br. J. Philos. Sci. **50**(3), 425–429 (1999)
12. Grosholz, E.: Representation and Productive Ambiguity in Mathematics and the Sciences. Oxford University Press, London (2007)
13. Hartshorne, C., Weiss, P. (eds.): Collected Papers of Charles Sanders Peirce, vols. 1–6. Harvard University Press, Cambridge (1931–1935)
14. Lakoff, G., Nuñez, R.: Where Mathematics Comes from: How the Embodied Mind Brings Mathematics into Being. Basic Books, New York (2001)

15. Lambert, J.H.: Neues Organon. Akademie Verlag, Berlin (1990)
16. Macbeth, D.: Diagrammatic reasoning in Euclid's Elements. In: Van Kerkhove, B., De Vuyst, J., Van Bendegem, J.P. (eds.) Philosophical Perspectives on Mathematical Practice, pp. 235–267. College Publications, London (2010)
17. Mancosu, P.: Mathematical explanation: problems and prospects. Topoi **20**, 97–117 (2001)
18. Nelsen, R.: Proofs Without Words: Exercises in Visual Thinking (Classroom Resource Materials). Math. Assoc. of America, Washington (1997)
19. Nelsen, R.: Proofs Without Words: More Exercises in Visual Thinking (Classroom Resource Materials). Math. Assoc. of America, Washington (2001)
20. Neurath, O.: Visual education: a new language. Surv. Graph. **26**(1), 25 (1937). http://newdeal.feri.org/survey/37025.htm
21. Peirce, C.S.: Prolegomena to an apology for pragmaticism. Monist **16**(4), 492–546 (1906)
22. Polya, G.: Mathematics and Plausible Reasoning. Princeton University Press, Princeton (1968)
23. Rota, G.C.: The phenomenology of mathematical proof. Synthese **111**(2), 183–196 (1997) (Special Issue on Proof and Progress in Mathematics edited by A. Kamamori)
24. Shimojima, A.: The graphic-linguistic distinction. Artif. Intell. Rev. **15**, 5–27 (2001)
25. Shin, S.-J.: The Logical Status of Diagrams. Cambridge University Press, Cambridge (1994)
26. Shin, S.-J.: Heterogeneous reasoning and its logic. Bull. Symb. Log. **10**(1), 86–106 (2004)
27. Shin, S.-J., Lemon, O.: Diagrams. Entry in the Stanford Encyclopedia of Philosophy (2008). http://plato.stanford.edu/entries/diagrams/
28. Tennant, N.: The withering away of formal semantics? Mind Lang. **1**(4), 302–318 (1986)

V. Giardino (✉)
Institut Jean Nicod (CNRS-EHESS-ENS), Pavillon Jardin, Ecole Normale Supérieure,
29, rue d'Ulm, 75005 Paris, France
e-mail: valeria.giardino@ens.fr

Figures, Formulae, and Functors

Zach Weber

Abstract This article suggests a novel way to advance a current debate in the philosophy of mathematics. The debate concerns the role of diagrams and visual reasoning in proofs—which I take to concern the criteria of legitimate *representation* of mathematical thought. Drawing on the so-called 'maverick' approach to philosophy of mathematics, I turn to mathematical practice itself to adjudicate in this debate, and in particular to category theory, because there (a) diagrams obviously play a major role, and (b) category theory itself addresses questions of representation and information preservation over mappings. We obtain a mathematical answer to a philosophical question: a good mathematical representation can be characterized as a category theoretic *natural transformation*. Assuming that this is not some reductio against the maverick approach to these issues, this in turn moots some of the disagreement in the philosophical debate and provides better questions with which to go on.

Keywords Picture proofs · Mathematical practice · Representation · Category theory

Mathematics Subject Classification 00A30 · 00A66 · 18A15

1 Towards a Mathematics of Real Philosophy?

A current debate in the philosophy of mathematics is over the role of visual reasoning [6, 14]. Another is over the importance of engaging with actual mathematics as it is *practiced* [8, 23]; this has sparked a so-called 'maverick' approach to philosophy of mathematics (see Sect. 2.4). The two issues are already related, by dint of emphasis on the empirical, but in bringing them together I mean to achieve something specific: to go inside of mathematical practice, looking at a number of real examples, in search of a subject-matter-appropriate answer to a philosophical question. In this way we can both test the viability of the maverick approach, and obtain further insight into the role of figures and diagrams in mathematical practice.

The philosophical question, on the criteria for legitimate representation, is *external*—something about or 'over' mathematics. The methodological reply is internal, an answer from or inside mathematics. A very plausible bridge between these two viewpoints is category theory, because it is mathematics (a) in which visual reasoning becomes, at least prima facie, indispensable, and (b) in part designed to address philosophical questions, but in a mathematician's, not a philosopher's, way. Category theory, that is, seems perfectly suited to bring these two problems together and deliver whatever sort of solution is appropriate.

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic,
DOI 10.1007/978-3-0348-0600-8_9,

In debating the status of pictures and visual reasoning—can a picture be a proof?—the notion of *representation* is crucial. There are two senses of representation to consider. One can talk about representing one mathematical object or structure with another, for example, giving a concrete representation of an abstract group, or using a formal power series to represent a function on a particular domain. One can also talk about using physical artifacts, like diagrams or expressions, to represent mathematical objects and assertions. Although these are prima facie two very different senses of the word 'representation', we will see how the two blur together, to the point of becoming indistinct. For example, which sort of representation is an equation, like $f(x) = x^2$, of a geometric object, like a sharp curve? I am open to, and even welcome, the possibility of a mathematical answer (representation of the first type) to philosophical questions (about representations of the second type).

Similarly, and even more basically, while there is an apparent difference between figures and formulae (pictorial versus syntactic imaging), that distinction too becomes complicated; examples are given in Sect. 3 and throughout. This paper sets up the problem of understanding representations of the second sort, diagrams and formulas, and then moves to consider representations of the first sort, representations internal to mathematics. I urge that within category theory, the first sort of representation, at least insofar as it is germane to mathematics, can be completely and satisfactorily subsumed by the second—insofar as we accept the maverick methodology and are looking for an answer that is authentically mathematical. The category-theoretic schematic at which we arrive in Sect. 6 is a mathematical explanation of diagrammatic representation in mathematics.[1]

2 Pictures, Proofs, and Mavericks

Across a great deal of mathematics and its philosophy, figures and formulae are conceived of as *oppositional*. In modern[2] Germany, Weierstrass led a distinct migration away from geometric visualization, toward the syntactic, with an enthusiasm for the arithmetization of nearly everything. Today, by contrast, there is a movement amongst a loose-knit group of 'mavericks' to return our attention to the pictures. This section provides a very surface level summary of this literature. Positions are presented with varying degrees of urgency, which we now sketch in turn.

[1]An initial worry about this whole approach may be put as follows (and thanks to a referee for doing so). If an answer is mathematical then that means that the question turns out to be mathematical as well—for wouldn't it be very surprising if questions in metaphysics or epistemology had mathematical answers? It is the burden of this paper and similarly motivated projects to show how *mathematical philosophy* can be fruitful, noting for now that the idea has been around at least since Descartes and Leibniz. As a research program, the mathematical approach would be interesting and informative even if it is ultimately unsuccessful.

[2]This is to use 'modern' in the technical sense of Jeremy Grey: "Modernism is defined as an autonomous body of ideas, having little or no outward reference, placing considerable emphasis on formal aspects of the work and maintaining a complicated—indeed, anxious—rather than a naive relationship with the day-to-day world, which is the de facto view of a coherent group of people, such as a professional or discipline based group that has a high sense of the seriousness and value of what it is trying to achieve" [17, p. 1].

2.1 Pictures Prohibited

We begin with old and familiar claims. Pictures should always be eliminated from valid reasoning. A pair of quotes can stand for all:

> A theorem is only proved when the proof is completely independent of the diagram. (Hilbert 1894)
> If for the grasp of a proof the corresponding figure is indispensable then the proof does not satisfy the requirements that we imposed on it. ... In any complete proof the figure is dispensable. (Pasch 1882)

The second claim is actually stronger, since it concerns our *grasp* of a proof. The justification is simply the old observation (to which we will return) that diagrams can hide assumptions (e.g. the Jordan curve property), elide proof steps, and even suggest outright falsities (e.g. that all functions are continuous). Familiar examples of the failure of spatial intuition are Cantor dust, or Peano's space filling curve—but also *non-Hausdorff spaces* containing points p, q such that every open neighborhood of p intersects some open neighborhood of q, but $p \neq q$. Because they lead us astray, pictures must always be eliminable from clear and valid reasoning.

2.2 Pictures in Proofs

Can a picture be part of valid reasoning? Driven by quite traditional proof-theoretic backgrounds, a positive answer is espoused by Barwise and Etchemendy, and separately by Shin. While a single image may not be enough to count as a proof, some steps in a proof may include unapologetically diagrammatic elements.

> We claim that visual forms of representation can be important, not just as heuristic and pedagogical tools, but as legitimate elements in mathematical proofs. ... Not all valid reasoning is (or can be cast) in the form of a sequence of sentences from some language [4, p. 3].

To make their point, Barwise and Etchemendy have developed formal reasoning systems using diagrams. Shin has developed a sophisticated account of diagrammatic reasoning [26]. Most recently, Avigad et al. have formalized Euclidean diagrammatic reasoning [1].

The strongest claim is that pictures can be proofs. The main proponent here is James Brown:

> ...the prevailing attitude is that pictures are really no more than heuristic devices; they are psychologically important—but they *prove* nothing. I want to oppose this view and make a case for pictures having a legitimate role to play as evidence and justification—a role well beyond the heuristic. In short, pictures can prove theorems [6, p. 26].

Brown's unusually bold claim is made in support for his Platonism, to solve Benacaraff's problem about our epistemic access to abstract objects. Brown asserts that pictures are an instrument, like a telescope, for knowing about mathematical truth. "Some pictures are not really representations, but are rather windows into Plato's heaven" [6, p. 44].

2.3 Pictures in Practice

A moderate thought between these two fairly polar 'right' and 'left' wing views (again, which I've only presented in cartoon form): Visualization is essential in some aspects of mathematical practice. The views from the right and the left concern *proofs* almost exclusively; the centrist view notices that proving is only a part of what mathematicians do. This balanced view, associated with Paolo Mancosu and Marcus Giaquinto, arises from the study of "the interaction between perception, visual imaging, concepts and belief formation" [24, p. 23]. This stance does not entail that visualizations count as proofs in the traditional sense. The focus here—especially for Giaquinto—tends into the psychological, on how we learn ("discovery") and come to form beliefs. Mancosu takes the main point to be

> Exclusive attention to the goal of justification is unacceptable. There are many other important epistemic goals, such as discovery, explanation, understanding, genesis of concepts, etc., that philosophy of mathematics should account for [24, p. 26].

This position does not necessarily dispute the critique of diagrammatic reasoning as dangerous. Rather, it provides a new justification for the critique, and so a new way forward:

> The reasons for why such tools are problematic is not necessarily on account of some intrinsic feature of the visual medium. It is rather that one must always check that the visual medium does not introduce constraints of its own on the representation of the target area [24, p. 26].

I think that Giaquinto is correct here, and that the important and difficult task is to make precise the idea of 'checking' the adequacy of visual media.

As I have already indicated, the key term to be investigated is representation (see Sect. 4.3). It seems like a very plausible thought that good mathematics represents its subject matter in a way that sheds more light, but without distortions. (Bad representations, by turns, substitute *obscurum per obscurius*.) We translate and abstract, without losing essential information. In examples to come, we will revisit the theme that good mathematics, and so good mathematical representation, captures *enough*, but *not too much*. We will find rigorous terms to characterize this.

2.4 Philosophers of 'Real' Mathematics

Recently, a strain of philosophy of mathematics has emerged and taken on the label of 'maverick'. Mancosu characterizes the maverick movement by three main tenets: Anti-foundationalism, anti-logicism, and attention to actual current mathematical practice [23, p. 5]. Under this banner we have [8, 11, 15, 18], and perhaps [3].

In many ways the mavericks are the intellectual descendant of philosophers like Lakatos, and also Pólya. The name seems to originate with Aspray and Kitcher in 1988, who use the epithet for those asking questions like:

> How does mathematical knowledge grow? What is mathematical progress? What makes some mathematical ideas (or theories) better than others? What is mathematical explanation? [8, at p. 18]

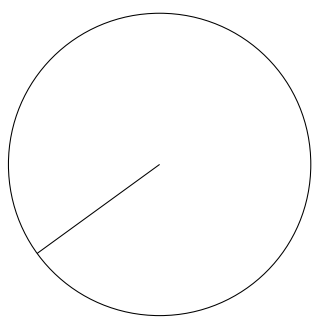
The maverick movement is as much about methodology as subject matter. But one point of detail is clear: it views the standard questions in philosophy of mathematics—Platonism versus formalism versus intuitionism versus structuralism, and the Benacerraf problems (what are numbers? how do we know?)—as old and exhausted research programs. "Mathematics has been and remains a superb resource for philosophers," writes Corfield. "Let's not waste it" [8, p. 270].

Implicit in the maverick approach is, I think, a hypothesis: Some philosophical questions can have mathematical answers. This is rather stronger than many quite full-blooded claims; more often the idea from Corfield et al. is that mathematics itself should generate the questions to which philosophers address themselves. So the idea is to do something approaching an anthropological study of the mathematics, where all the questions are guided by the given behaviors. But the questions Aspray and Kitcher attribute to the mavericks, and to which Corfield ascribes, are inherently *philosophical* questions. They concern epistemology and metaphysics (realism and anti-realism). It is not a stretch to follow up the maverick shift by posing philosophical questions directly to mathematical datum. We ask of mathematical practice: What are the roles of figures and formulae, their relations and representations? The validity of such a method is the working hypothesis of this paper.

Having given an overview of some philosophical debates, we turn to the putative distinction between visual and non-visual aspects of mathematical practice, by way of some examples.

3 Figures and Formulae

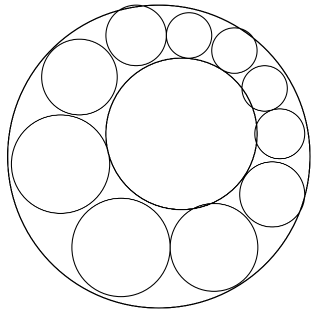
Here is a circle with radius r,



and here is a circle with radius r:

$$x^2 + y^2 = r^2$$

These are two ways of seeing a mathematical object, geometric and algebraic. To state an apparent platitude, appreciating the object in different ways may lead to different ideas. For instance, looking at the first circle, and perhaps playing around with a compass, may lead to *Steiner's porism*: Given two non-concentric circles, one inside the other, suppose circles are drawn successively touching them and one another. If the last circle in the chain touches the first, then the last will *always* touch the first, regardless of the position of the first circle in the chain.



Notice the 'if.' Not every pair of circles so inscribed allows for the generation of a Steiner chain. But the ones that do, always do [9, p. 87].

This seems like a straightforwardly *geometric* observation. Some thoughts about circles, on the other hand, have a less concrete character. Thinking more about r, some time and imagination might lead to the thought that *a straight line is the circumference of a circle with infinite radius*. This is a fundamentally abstract thought, although it still draws on a visualization thought-experiment. (Imagine the radius getting longer.[3])

Other properties emerge only at the level of complete divorce from the actual picture as embedded in space. From the picture alone, it hardly even seems worth pointing out that the circle divides the space in which it sits in two parts, i.e. that a closed curve divides a space into an interior and an exterior. This, the *Jordan curve theorem*, turns out not only to be a deep topological property, but one that is very hard to prove. Its statement only arises at the point where a closed curve is not a visual object in the space X at all any more, but rather an algebraic mapping $f : [0, 1] \longrightarrow X$.

Geometric insights lead to powerful algebraic expressions; algebra, in turn, leads to developments that beg for geometric interpretation.[4] Here are circles and triangles; here are quadratics and trigonometric functions. Here are solutions to equations of the form $x = \sqrt{-1}$; here is an interpretation of these numbers as rotations in the complex plane [28, p. 188]. The two modes of mathematical reasoning are mutually reinforcing, if not mutually dependent [17].

The two modes of practice do not simply harmonize; there is tension. Perhaps the discomfort is best expressed directly out of a mathematics text—Hartshorne's lovely *Geometry* book [19]. Let A, B, P, Q be distinct points in the Cartesian plane. Their *cross-ratio* is an element of the field defined as

$$(AB, PQ) = \frac{AP}{AQ} \cdot \frac{BQ}{BP}$$

Hartshorne then says:

> I can just hear someone asking, "What is the geometrical significance of the cross-ratio?" Although I first encountered the cross-ratio as a senior in high school, and have dealt with them many times since then, I must say frankly that I cannot visualize a cross-ratio geometrically. If you like, it is

[3]Gauss' definition of intrinsic curvature repeats this result, since the curvature $\kappa(t) = \frac{1}{r}$ of the osculating circle is identically 0 when r goes to infinity [21, p. 3].

[4]Stillwell (not historically unproblematically) observes that "the Greeks used curves to study algebra rather than the other way around" [28, p. 64].

> magic. Here is an algebraic quantity whose significance is impossible to understand, and yet turns out to do something very useful. It works. You might say it was a triumph of algebra to invent this quantity that turns out to be so valuable and could not be imagined geometrically. Or if you are a geometer at heart, you might say that it is an invention of the devil, and hate it all your life [19, p. 341].

The channels of communication, then, between algebra and geometry are not always completely clear. But the cooperation of geometry and algebra is fruitful more often than it is fraught. Problems that defy solution, indeed, concepts that defy expression, mechanically take care of themselves when couched in the language of formulae and rules. The algebra then presents new problems that can only be approached geometrically. For example, here is a difficult notion to come to grips with:

Physical reality has more than three dimensions.

And here is an easy notion to come to come to grips with.

$$\mathbb{R}^n = \left\{ \langle x_0, \ldots, x_n \rangle : x_0, \ldots, x_n \in \mathbb{R} \right\}$$

Grasping the geometric meaning of $\mathbb{R}^n$ appears hopeless. Grasping the syntax of n-tuples is easy. But then, it is only through the geometric language of topology and differentiable manifolds, tensor calculus (Sect. 4), and sketch after sketch of contorted annuli, that such spaces can be reasoned about and new directions suggested.[5] The inspiration and insight of geometry passes to the exactitude and algorithmic inevitability of algebra, and back again.[6]

More expansively, there are crucial visual elements to even pure formula manipulation. The apparently syntactic cancellation law

$$\begin{aligned} & x + y = x + z \\ \Rightarrow \quad & y = z \end{aligned}$$

has a striking visual aspect [14, p. 242]. It also has an explanation: the existence of inverse and unit elements. Supposing $x + y = x + z$, then

$$\begin{aligned} & -x + (x + y) = -x + (x + z) \\ \Rightarrow \quad & (-x + x) + y = (-x + x) + z \\ \Rightarrow \quad & 0 + y = 0 + z \\ \Rightarrow \quad & y = z \end{aligned}$$

This does not have the feel of learning, say, to conjugate verbs in natural language, or of declining nouns to fit moods. It does not appear to be a *linguistic* artifact. This is not a series of rules for expressing mathematics; this *is* mathematics.

[5] For methods of drawing topological objects, see [12]. For an abstract physics text that makes extensive use of diagrams, see [13].

[6] As long as Algebra and Geometry were separated, their progress was slow and their use limited; but once these sciences were united, they lent each other mutual support and advanced rapidly together towards perfection. (Lagrange 1795)

There are two observations to make in passing. First, the visual and syntactic tendencies drive mathematics towards one of its most general ends: *unification*. This means finding ever deeper connections between apparently disparate fields, fields opened by figures and formulae respectively, and showing how tools developed for use in one area can be put to very fecund work in another. Category theory is well-known to be directed toward unificationist ends [22]. Second, the means of reaching this unification is through *representation*. Understanding representations, I think, is the key to understanding the interaction between figures and formulae—because a representation can potentially hold between *anything*: between different types of formulae, between types of pictures, and between figures and formulae.

As a paradigm case of how representations play here, we cite the leitmotif of the magnificent *Rings of Continuous Functions* [16]. Let X be a topological space, and consider the set of all continuous functions from X into the reals $\mathbb{R}$. This set, $C(X)$, forms a ring under some obvious operations on the functions; so $C(X)$ has a clear algebraic structure. We then learn that, given two spaces X, Y, if X is homomorphic to Y, then their corresponding rings $C(X)$, $C(Y)$ are isomorphic. The structure of the underlying space completely determines the ring structure. That is striking enough. The bulk of *Rings*, though, addresses the question of what conditions on the algebraic structures will determine the topology of the spaces [16, p. 12]. Given that $C(X)$ is algebraically equivalent to $C(Y)$, what else do we need to know to conclude that X and Y are topologically equivalent? The answer constitutes a volume of fertile mathematics spanning topology, algebra, and set theory. By finding suitable representations of the space X in the ring $C(X)$, light is cast.

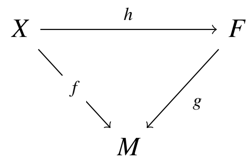
The question for us is what makes a representation more or less faithful, more or less useful. In part the answer to this question will depend on the particularities of any given case, on the sort of mathematics one is doing. There is still something to be gained, though, from asking the question at very high level of generality. We are looking for an abstract schema that expresses how a representation holds between structures. This is the relationship to examine now, through applied examples.

4 Examples

To get a basic sense of how the visual features in formal reasoning, let us consider a few formula-driven ideas. The first example is from abstract algebra; the second is the tensor calculus. Each reveals how syntax itself can, in a way, become a mathematical object. The third example, Galois theory, shows how representation internal to mathematics works.

4.1 Free Modules

A module is a generalization of a vector space. Let X be a subset of a module F. Say that F is *free* on X iff there is a mapping $h : X \longrightarrow F$ such that, for any $f : X \longrightarrow M$ from X to a module M there is a unique homomorphism $g : F \longrightarrow M$ such that $f = g \circ h$. Putting this notation into two dimensions makes a figure; and while the formal definition requires some time to absorb, the picture shows exactly and immediately what it means for F to be free [7, p. 38]: F is free on X when the diagram commutes:[7]



It turns out that any such X has a free module. This is extremely important, and is called the *universal mapping property* (UMP); we will return to it we discuss category theory in Sect. 5.

But now notice what the diagram does *not* tell us—which we would need to know to prove that for any set X there is an R-module free on X. The definition as a quantified sentence of mathematical English says this: A module F is *free* on X iff

$$\exists h\big(h : X \longrightarrow F \text{ and } \forall f\big(\text{if } f : X \longrightarrow M \text{ then } \exists g : F \longrightarrow M \text{ such that:}$$
$$g \text{ is a homomorphism, and } f = g \circ h,$$
$$\text{and } \forall g'\big(\text{if } g' \text{ is a homomorphism and } f = g' \circ h, \text{ then } g = g'\big)\big)\big)$$

The form of the sentence shows that the universal mapping property holds, or not, depending on the *existence* of an h with the required property. The diagram can carry neither the existential aspect, nor can it express the uniqueness of g. It does not capture logical relations of conditionality, nor that these are the salient properties. In the diagram, it looks like it is all already there.

In his recent textbook, Awodey explains that, indeed, it is the existence and uniqueness aspects that make the definition useful [2, p. 17]. His terminology will illuminate the entire discussion to come. To insure that there is "no junk" in the module, we want every member of the free module to be expressible as a product of elements of X, for some operation defined in terms of F. The uniqueness condition means that there is no junk. Further, there should be "no noise," meaning that any relations holding between elements are required to do so by the axioms. The existence part of the definition controls for noise. No trivial relations can be proved to hold between elements. These are the most important facts about free modules, which will recur often in the pages ahead. And a formula is required to know them.

4.2 Tensor Calculus

The tensor calculus is a notational system, devised by Ricci and his student Levi-Civita at the turn of the 20th century, to express ideas from mechanics and differential geometry. Tensor calculus streamlines the manipulation of a great deal of information, through syntactic rules and abbreviations. It was first called 'absolute differential geometry', to emphasize the calculus' freedom from coordinate systems. Einstein couched his theory

[7] A diagram like this one *commutes* when $g \circ h = f$, i.e. $g(h(x)) = f(x)$.

of relativity in the Ricci formalism.[8] As I illustrate, though, the system is not merely notation, but is itself a new piece of mathematics. The point of this section is that, among other things, formalism can be figurative. In the case of tensor calculus, the structure of its syntax is the syntax of the structure it describes.

A main theme in tensor calculus is to use indices judiciously. The calculus "guides the user through explicit computations—virtually a *machine-readable* calculus… The calculus *thinks for the user*" [20, pp. 31, 42]. This is achieved through a confluence of a background coordinate system with a foreground formalism.[9] Another is to compress information—both because there is so much information, and also because while coordinates may not be easy to display in dimensions higher than 2, their syntactic counterpart, indexed variables viz. $g_{\mu\nu\ldots\rho}$, is eminently flexible.

Here is a bite-sized example. The *Einstein summation convention* is that, when indices use the same variable, a summation is implied over that index, which means that $g^{\nu}_{\mu\nu}$ is to be understood to say $\sum_{\nu=0}^{n} g^{\nu}_{\mu\nu}$. Now define the Kronecker delta,

$$\delta^{\nu}_{\mu} = \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \neq \nu \end{cases}$$

Let us consider an N-dimensional space. Using Einstein's summation convention, we obtain

$$\delta^{\mu}_{\mu} = N$$

In other words, the dimension of a space can be recovered from the notation, but otherwise can remain an inessential part of the background.

The visual arrangement of syntactic marks on the page serves, in a structurally significant way, to express and characterize geometric information. The pages of Dirac's *General Relativity* [10] are themselves laid out in a spare and arid way, filled with empty white space, that itself contributes to the communication of ideas. Dirac shows that the curvature of space can be expressed with a concision and perspicacity that outstrips mere notational representation. The *Ricci tensor* is defined,[10] and Einstein's law of gravitation can be expressed in a heartbeat: In empty space,

$$R_{\mu\nu} = 0$$

[8]Einstein wrote to Levi-Civita, "I admire the elegance of your method of computation; it must be nice to ride through these fields upon the horse of true mathematics while the like of us have to make our way laboriously on foot." Quoted in Goodstein, J.R.: The Italian mathematicians of relativity. Centaurus **26**(3), 241–261 (2007).

[9]Dirac: "With a network of curvilinear coordinates the $g_{\mu\nu}$, given as functions of the coordinates, fix all the elements of distance; so they fix the metric. They determine both the coordinate system and the curvature of space" [10, p. 9].

[10]If the speed of light is taken to be unity, the metric on Minkowski space is

$$\left(dx^0\right)^2 - \left(dx^1\right)^2 - \left(dx^2\right)^2 - \left(dx^3\right)^2$$

Through some notational conventions for moving indices up and down, this reduces to $dx_{\mu}dx^{\mu}$, with Einstein's summation convention understood. A coordinate system with coefficients $g_{\mu\nu}$ is introduced. The invariant distance between a point x^{μ} and a nearby point $x^{\mu} + dx^{\mu}$ is

$$g_{\mu\nu}dx^{\mu}dx^{\nu}$$

All this pure syntax is not just bookkeeping. It is a method for capturing structure, in compressed and manipulable forms. A tensor is a geometric object, or even a physical object—a generalization of a 'directed quantity,' a vector. Yet a vector is just an inhabitant of a vector space; and a vector space in turn is any structure satisfying a list of algebraic axioms. (See Sect. 4.3.) The structure of the syntax, in other words, dictates any further syntax of the structure; cf. the cancellation law for groups (Sect. 3).

4.3 Galois Theory

Having considered how formalism can itself be figurative, we more overtly emphasize the pressing issue of representation. The example of Galois theory shows how mathematics unifies by connecting disparate branches, and how faithful representations between mathematical objects—in this case, algorithmic constructions and algebraic structures—make this unity possible.[11]

Galois found a way to answer questions about polynomials by checking for properties of groups. This is a particularly clear instance of a venerable mathematical strategy: given a problem about some intractable object X, find an object $[\![X]\!]$ that is tractable, and show that $[\![\cdot]\!]$ is a faithful means of representing X for all relevant purposes. In Galois' theory, he associated with each polynomial the set of its symmetries. These form a group. Galois showed that, if a polynomial is soluble, its Galois group will have a certain property; this is a device, then, for showing that certain polynomials are insoluble.

Galois' achievement is to find more and more sophisticated algebraic objects through which we can study target phenomena. Consider the question: Can a cube be duplicated using only a straightedge and compass? Let $\mathbb{F}$ be a field, and K, L be subfields of $\mathbb{F}$. A *field extension* is a map $i : K \longrightarrow L$ that is structure preserving:

$$\iota(x + y) = \iota(x) + \iota(y)$$

Now, if $L : K$ is a field extension, then, just because K and L are subfields and $K \subseteq L$, the operations

$$u + v : L \times L \longrightarrow L$$
$$\lambda u : K \times L \longrightarrow L$$

define on L the structure of a vector space over K. This is an instance of the universal mapping property from Sect. 4.1.

---

again with summations understood. If we define a *Christoffel symbol*

$$\Gamma_{\mu\nu\sigma} := 1/2(g_{\mu\nu,\sigma} + g_{\mu\sigma,\nu} - g_{\nu\sigma,\mu})$$

then we can define the Riemann-Christoffel curvature tensor

$$R^{\beta}_{\nu\rho\sigma} := \Gamma^{\beta}_{\nu\sigma,\rho} - \Gamma^{\beta}_{\nu\rho,\sigma} + \Gamma^{\alpha}_{\nu\sigma}\Gamma^{\beta}_{\alpha\rho} - \Gamma^{\alpha}_{\nu\rho}\Gamma^{\beta}_{\alpha\sigma}$$

Lowering suffixes, the tensor $R_{\mu\nu\rho\sigma}$ has 256 components. One can prove that space is flat if and only if $R_{\mu\nu\rho\sigma}$ vanishes. The *Ricci tensor* is defined: $R^{\mu}_{\nu\rho\mu} = R_{\nu\rho}$.

[11] For a rigorous but accessible presentation, see [27].

Define the *degree of a field extension* $L : K$ to be the number of its basis vectors of the vector space on L over K.[12] Just knowing about the degree of a field extension provides enough to answer solvability questions. Any point $\langle x, y \rangle$ in 2-space that is compass-and-straightedge constructible from a subfield K will have associated degrees that are powers of 2. But the duplication of a cube, or trisection of an angle, by compass and straightedge would generate a point (a, b) in the rational plane with degree 3. Unfortunately, 3 is not a power of 2.

The fundamental theorem of Galois theory establishes a one-to-one correspondence between a field extension and the automorphisms on it—the Galois group. Algebra captures the right kind of information about geometric constructions and equations, namely size (number of basis vectors) and shape (group structure). Ruler-and-compass constructions are algorithms for deriving one set of points from another; we learn about them through measurements of vector spaces. The representations are appropriate in an exact way. There are no pictures involved (although ruler-and-compass work is paradigmatic geometry) but we have a clear instance of high-fidelity transmission of information. This was the start of abstract algebra; let us now see if we can generalize on the processes we've seen here and with the universal mapping properties, by turning to abstract algebra.

4.4 The Argument So Far

Formalism can itself be figurative. From commutative diagrams to the visual aspect of syntax in the tensor calculus, we see how notation itself captures, compresses, and expresses mathematical information. A very simple example is the *ascending chain condition* on Noetherian rings [7, Chap. 3], that for some n of a chain,

$$C_0 \subseteq C_1 \subseteq \cdots \subseteq C_n = C_{n+1}$$

It is easy to *see* the condition in the symbolism, even if the reason is as simple as the similarity of the sign $\subseteq$ to a paper-clip chain. These examples, and Galois' theory, show how representation *within* mathematics—a geometric construction represented as a point in a field extension, for instance—are not importantly different from representations *of* mathematics by figures or diagrams. So the working distinctions of this paper—proofs versus pictures, and formulae versus figures—is becoming blurred. And this is as we expected; but we need now some accounting of how the vague boundary is mediated. We want a theory of how the 'same' mathematical information can be presented in such different ways, and to such different effects.

As a starting point, the relationship seemed to be expressed precisely by the existence and uniqueness of certain mappings, as seen in the UMP for free modules, and again when we defined the structure of a vector space over a field extension. The free module F on X must factor into members of X alone: this corresponds to the condition that, for any M such that X maps to M, the unique structure preserving map

$$g : F \longrightarrow M$$

[12] A vector space may have many different bases, but there are always the same number of basis vectors in each; a vector space is uniquely determined by its *dimension*, up to isomorphism.

exists. All relations between members of the module are determined by the algebraic structure: this corresponds to the condition that the map

$$h : X \longrightarrow F$$

exists. We turn, then, to a generalization of what we've already seen about free modules and commutative diagrams—to some 'abstract nonsense' that may yet make sense of all this.

5 Category Theory

Category theory concerns objects $X, Y, \ldots,$ and arrows between objects. The arrows form sets,

$$hom(XY) = \{f : f \text{ is an arrow from } X \text{ to } Y\}$$

Often there is an isomorphism ϕ between sets $hom(XY)$, $hom(UV)$, about which it is customary to abuse notation and write

$$\phi : hom(XY) \cong hom(UV)$$

The set $hom(X, Y)$ is a small picture of the mathematical life of the objects in a category, at an extremely general level. This is the level at which we can now do two things.

First, we have a mathematical discipline where figures and formulae become intermingled to the point of being almost indistinguishable. So we can validate our hypothesis that the simple yes/no distinction for pictures is unhelpful. Second, the discipline is itself a study of what mathematical information is preserved when figures and formulae become indistinguishable; and so we will try to appropriate some category theoretic notions to answer our original question, of what makes a good representation.

In Sect. 4.1 above we considered free modules and noted that F being free on X would be expressed by category theorists as a universal mapping property. Now we sketch one of the deeper instances of a UMP.

An *adjoint functor* is a category theoretic notion of tremendous scope. In logic, the existential quantifier is an adjoint. In topology, the image of a continuous function is an adjoint. There is no prima facie reason to see any connection between these two, but from a categorical point of view, they turn out to be essentially the same thing [2, Chap. 9]. In Galois theory, it is impressive that groups and solutions to equations are related, but it is not exactly *baffling*. Both are algebraic to begin with. In the case of adjoint functors, we are at the heights of structural abstraction. Our outline of adjunctions follows [5, Chap. XV].

Let F, G be functors between categories $\mathcal{C}, \mathcal{D}$. A *natural transformation* $\eta : F \longrightarrow G$ is a function which assigns each $X \in \mathcal{C}$ an arrow $\eta_X : F(X) \longrightarrow G(X)$ of $\mathcal{D}$ in such a way that every arrow $f : X \longrightarrow Y$ of $\mathcal{C}$ yields a commutative diagram

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ \downarrow{\scriptstyle \eta_X} & & \downarrow{\scriptstyle \eta_Y} \\ G(X) & \xrightarrow{G(f)} & G(Y) \end{array}$$

This elegant figure is the key to the whole story.

Let us have two categories $\mathcal{C}$, $\mathcal{D}$ and two functors F, G, such that

$$\mathcal{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathcal{D}$$

preserves map structure. Again stretching the use of notation, the relationship can be written

$$F : \mathcal{C} \longleftrightarrow \mathcal{D} : G$$

For $X \in \mathcal{C}$ and $Y \in \mathcal{D}$, we compare all the arrows $hom(F(X), Y)$ in $\mathcal{D}$ with all the arrows in $\mathcal{C}$ from X to $G(Y)$, the set $hom(X, G(Y))$. The *adjunction* of the functor F to the functor G is a natural bijection

$$\phi : hom_{\mathcal{D}}\big(F(X), Y\big) \cong hom_{\mathcal{C}}\big(X, G(Y)\big)$$

In this case we say that F, G are *adjoints*.[13]

Returning this terminology to Sect. 4.1, if we write the underlying set of the module M as $U(M)$ and the free module F as $F = V(X)$, the condition for an adjunction then is that every element of $hom(X, U(M))$ corresponds to exactly one element of $hom(V(X), M)$, and every monoid homomorphism restricts to a function on X. That is, for every set X and monoid M there is an isomorphism ϕ of sets

$$\phi : hom\big(V(X), M\big) \cong hom\big(X, U(M)\big)$$

So we see that the free module is adjoint to the underlying set. This result is appealed to in the next section.

By generalizing the basic faithfulness conditions of existence and uniqueness—'no junk, no noise'; enough, but not too much—to the austerity of adjoints, we have now formalized a notion ubiquitous in mathematics.[14] But more interestingly, I think, we have formalized a practice ubiquitous *over* mathematics. Perhaps this is the most distinctive feature of category theory, and what makes it striking from a philosophical point of view. Not only do images and inscriptions become rather indistinguishable in commutative diagrams; this very indistinguishably is exactly what the diagrams, and so category theory, is about. By comparison, on Klein's famous Erlangen view of geometry, a geometry is characterized by which properties are preserved under transformations. The Erlangen view of

[13]There is a related category theoretic notion of *representation*. A representation is a natural homomorphism (in the technical sense of natural, above) between S and a set theoretic object. One can prove that a functor $G : \mathcal{D} \longrightarrow \mathcal{C}$ has a left adjoint iff each $hom(X, G-)$ is representable, for every $X \in C$.

[14]An adjunction between two categories can also be an equivalence relation. An *equivalence* on functors $F : \mathcal{C} \longrightarrow \mathcal{D}$, $G : \mathcal{D} \longrightarrow \mathcal{C}$ is such that

$$hom(Z, X) \cong hom\big(F(Z), F(X)\big)$$

This is "isomorphism up to isomorphism" [2, p. 148]. Writing $Z = G(Y)$, it follows by using some natural isomorphisms that $hom(G(Y), X) \cong hom(F(G(Y)), F(X)) \cong hom(Y, F(X))$. Therefore any equivalence is an adjunction [2, p. 181].

category theory for this paper, then, would be to identify which properties of arrows are preserved under the transformation from formula to diagram and vice versa. The theory is about the kind of structure of maps that is respected in diagrammatic displays of information. And the most pleasing part is this: In the case of adjoints, *the structure preserved is information about what kind of structure is preserved*.

It remains to explicitly apply these concepts to the philosophical debates discussed at the outset.

6 A Mathematical Answer to a Philosophical Question

Giaquinto's program, the 'centrist' view about visual reasoning discussed in Sect. 2.3, is to find criteria for representation. From the conclusion of his recent book, Giaquinto recommends:

> The symbolic-diagrammatic distinction (or the algebraic-geometric distinction) is too coarse to advance our understanding of visual thinking in mathematics, let alone mathematics in general: we need taxa of a finer grain. Although there is a contrast to be drawn between symbolic thinking and diagrammatic thinking, these are vague opposites… no more apt than thinking that every human is wise or foolish [14, p. 253].

Through the various examples we considered, we begin to have some criteria for the finer grained taxonomy Giaquinto suggests. Having seen how representations work from within mathematics, we get an answer that is reasonably simple, but therefore that much harder to draw out: Representations, pictorial or otherwise, should be faithful, clean and effective. Notions appropriated from category theory will offer us an appropriate schematic form of this basic observation.

6.1 The Figures-and-Formulae Functor

The interplay of figures and formulae occurs in situations where figures X and Y stand in some kind of visual relationship if their representations as formulae stand in an analogous syntactic relationship. Schematically, given a suitable representation, $\llbracket\cdot\rrbracket$, and relation f, we could say that X and Y are congruent modulo $\llbracket\cdot\rrbracket$, and draw this picture:

$$\begin{array}{ccc} X & \xrightarrow{\quad f \quad} & Y \\ \Big\downarrow{\scriptstyle \llbracket\cdot\rrbracket} & & \Big\downarrow{\scriptstyle \llbracket\cdot\rrbracket} \\ \llbracket X\rrbracket & \xrightarrow{\quad \llbracket f\rrbracket \quad} & \llbracket Y\rrbracket \end{array}$$

This is a two-place, relational schema. For the study of single objects, X and its representation $\llbracket X\rrbracket$, the point of entry would be to ask which maps from X to $\llbracket X\rrbracket$ are illuminating—that is, about $hom(X, \llbracket X\rrbracket)$. The embedding $\llbracket\cdot\rrbracket$ is a *functor*, a generalized homomorphism with the following properties:

$$\big\llbracket (f : X \longrightarrow Y)\big\rrbracket = \llbracket f\rrbracket : \llbracket X\rrbracket \longrightarrow \llbracket Y\rrbracket$$

$$[\![f \circ g]\!] = [\![f]\!] \circ [\![g]\!], \quad \forall g : Y \longrightarrow X$$
$$[\![1_X]\!] = 1_{[\![X]\!]}$$

where 1_X is the unit element for X. A good representation is a kind of functor.[15]

The diagram captures a recurring theme in the examples considered, the specification of existence and uniqueness conditions. These conditions correspond with the criteria that an interesting mathematical theory must have enough structure, but not too much. Tensor calculus compresses information to the point that a (well trained) user can handle it easily, all the while with background facts like dimension recoverable from the notation if needed. Free modules, which we saw in application in Galois theory, and then powerfully generalized as adjunctions, capture these basic conditions of mathematical adequacy and practice.

Of course, at the current level of abstraction, we have no indication of what criteria the mappings must meet to takes figures to formulae or vice versa in any interesting way. The proposed functor is more like a label for what is wanted. All the real work will be done by the specifics of which modules, functions etc. exist and are unique; all the real work would have to be done rather piecemeal, looking at each case, and explaining what is meant by the word ‘interesting’. Although our examples above my give some ostensive indication of what counts as interesting, I do not pretend that sketches with arrows solve all the philosophical problems at hand. And it must be emphasized that there remain straightforward problems with pictures, in particular, the giving of existence and uniqueness conditions (as noted in Sect. 4.1). The figures we have considered go no distance in allaying any worries there. The philosophical payoff in pursuing the analogy, between something philosophers are concerned with and some work that mathematicians do, remains to be seen, perhaps through some examples of how the schematic can be applied in a philosophically illuminating way. So adjunctions and the like are not alone a panacea.

For present, though, it is progress to say that the language of adjunctions and natural transformations give us a schematic for describing successful instances of representations, and does so in a way drawn from mathematical practice itself.

6.2 Concluding Remarks

Figures and formulae, when used well, represent each other. For this relationship to hold, they cannot simply be the *same*; it is pointless to consider the representation of X by X itself. So the two are at a certain distance; but they can be so without conflict. The language and images of category theory have provided us with a way to express this abstract relationship. Seen rightly, the mathematics is providing us with a means to understand itself. This is not a new idea; Hilbert proposed metamathematics as a branch of mathematics,

[15]The figures-and-formulae functor is not the most general schema for the sort of representations we mean to capture, since it requires the existence of a unit element 1_X for any X. It seems clear that there can be interesting instances of X with no such element. So this schema can be taken as a partial result, that holds over domains with a sufficient amount of structure. Similarly, not every mathematical object is a group or a ring, but still group and ring theory *are* quite ubiquitous, and have a lot to tell us.

and that program has given us both more reliable and deeper insights into the nature of proof and truth than the two millennia did previously.

Now, though, a century after Hilbert, we want to know about other features, too, beyond proof and truth. How are problems expressed? How are they solved? Pursing such questions, some old philosophical debates are mooted; we find instead an answer *internal* to mathematics. Describing some of the earliest diagrammatic mathematics, Netz writes:

> The [mathematical] diagram is not a representation of something else; it is the thing itself. It is not the representation of a building; it is like a building, acted upon and constructed [25, p. 60].

We have hit upon a representation of representations, which is both a hermeneutic limitation and a very welcome insight: Further rigorous accounts of structure are always only more structure. And so the interplay of figures and formulae goes on.[16]

References

1. Avigad, J., Dean, E., Mumma, J.: A formal system for Euclid's elements. Rev. Symb. Log. **2**(4), 700–768 (2010)
2. Awodey, S.: Category Theory. Oxford Logic Guides (2006)
3. Azzouni, J.: Tracking Reason: Proof, Consequence, and Truth. Oxford University Press, New York (2006)
4. Barwise, J., Etchemendy, J.: Visual information and valid reasoning. In: Allwein, G., Barwise, J. (eds.) Logical Reasoning with Diagrams, pp. 3–25. Oxford University Press, New York (1996)
5. Birkhoff, G., Maclane, S.: Algebra. Am. Math. Soc., Providence (1999) (First edition 1967)
6. Brown, J.: Philosophy of Mathematics, 2nd edn. Routledge, London (2008)
7. Cohn, P.M.: An Introduction to Ring Theory. Springer, Berlin (2000)
8. Corfield, D.: Towards a Philosophy of Real Mathematics. Cambridge University Press, Cambridge (2003)
9. Coxeter, H.M.: Introduction to Geometry, 2nd edn. Wiley, New York (1969)
10. Dirac, P.M.: The General Theory of Relativity. Wiley, New York (1975)
11. Ferreirós, J., Grey, J.J. (eds.): The Architecture of Modern Mathematics. Oxford University Press, Oxford (2006)
12. Francis, G.K.: A Topological Picture Book. Springer, Berlin (2007)
13. Frankel, T.: The Geometry of Physics: An Introduction, 2nd edn. Cambridge University Press, Cambridge (2004)
14. Giaquinto, M.: Visual Thinking in Mathematics. Oxford University Press, Oxford (2007)
15. Gilles, D. (ed.): Revolutions in Mathematics. Clarendon, Oxford (1992)
16. Gillman, L., Jerison, M.: Rings of Continuous Functions. Van Nostrand, Princeton (1960)
17. Grey, J.: Plato's Ghost: The Modernist Transformation of Mathematics. Princeton University Press, Princeton (2008)
18. Grosholz, E.R.: Representation and Productive Ambiguity in Mathematics and the Sciences. Oxford University Press, Oxford (2007)
19. Hartshorne, R.: Geometry: Euclid and Beyond. Springer, Berlin (2000)
20. Jänich, K.: Vector Analysis (trans.: Kay, L.). Springer, New York (2001)
21. Lee, J.M.: Riemannian Manifolds: An Introduction to Curvature. Springer, New York (1997)
22. MacLane, S.: Mathematics, Form and Function. Springer, Berlin (1986)
23. Mancosu, P. (ed.): The Philosophy of Mathematical Practice. Oxford University Press, New York (2008)
24. Mancosu, P., Jorgensen, K.F., Pedersen, S.A. (eds.): Visualization, Explanation and Reasoning Styles in Mathematics. Synthese Library, vol. 327. Springer, Dordrecht (2005)

[16]Thanks to: Philip Catton, Cathy Legg, Clemency Montelle, Peter Smith, and Koji Tanaka.

25. Netz, R.: The Shaping of Deduction in Greek Mathematics. Cambridge University Press, Cambridge (1999)
26. Shin, S.-J.: The Logical Status of Diagrams. Cambridge University Press, Cambridge (1994)
27. Stewart, I.: Galois Theory, 3rd edn. Chapman & Hall, London (2004)
28. Stillwell, J.: Mathematics and Its History. Springer, New York (2002)

Z. Weber (✉)
University of Otago, Dunedin, New Zealand
e-mail: zach.weber@otago.ac.nz

Representation of Graphs in Diagrams of Graph Theory

Mitsuko Wate-Mizuno

Abstract The treatise of Dénes Kőnig *Theorie der endlichen und unendlichen Graphen* (1936) includes many diagrams. The style of the diagrams is typical of present-day texts of graph theory. In this treatise, many mathematical recreation problems are treated. Some of the problems were already treated by precedent mathematicians. In some of these early works, the diagrams were given, but the styles were not the same as that of Kőnig. Moreover, the way to use the diagrams in the early works is different from that of Kőnig.

Examining the diagrams, we will find that a certain type of diagrams became gradually influential. This historical aspect may be related to the formation of concepts of graph, but it will not be discussed here.

We will argue that Tarry's talk "Géométrie de situation: nombre de manières distinctes de parcourir en une seule course toutes les allées d'un labyrinthe rentrant, en ne passant qu'une seule fois par chacune des allées" (1886) played an important role in the way to use the diagrams.

Keywords Kőnig · Tarry · Recreations · Graph

Mathematics Subject Classification (2010) 00A08 · 01A55 · 01A60 · 01A70 · 97A20

1 Introduction

Dénes Kőnig (1884–1944), a Hungarian mathematician, is recognized as the "father of graph theory" with his treatise *Theorie der endlichen und unendlichen Graphen* written in 1936 [29]. In this treatise, some problems are selected from mathematical recreations, and they play an important part. This fact corresponds to the fact that Kőnig, when he was still a student, published two books on mathematical recreations. In fact, Kőnig's treatise of 1936 and one of his books on mathematical recreations are closely related to each other.[1]

I will analyze how, in the context of mathematical recreations, some features of the diagrams of graph theory took shape. Moreover, I will establish that Kőnig inherited the features of the diagrams from some publications on mathematical recreations.

In the treatise of 1936, many diagrams are used for representing graphs. They are mostly of a single kind of diagram, consisting of curved or straight lines which represent

[1] I examined this relationship in my thesis [51], Chap. 6.

A. Moktefi, S.-J. Shin (eds.), *Visual Reasoning with Diagrams*, Studies in Universal Logic, DOI 10.1007/978-3-0348-0600-8_10,

edges, and small circles which represent vertices. This representation of graphs in diagrams of graph theory continues to be used in the texts of graph theory today.

I focus on the question of how this representation of graphs in diagrams of graph theory took shape.

To address this issue, it appears that one needs to examine some problems from mathematical recreations which in the treatise of 1936, Kőnig treated on the basis of the way mathematicians before him had dealt with them. Four problems appear to have played a key part in this process: a problem of Königsberg bridges, a problem of describing polygons, a problem of configuring dominoes and a problem of circulating in mazes. This selection of problems will allow us to deploy a historical approach to diagrams, and to identify how some of the key features of the diagrams of graph theory took shape in different contexts.

In the earlier writings, different types of diagrams were used for different topics. Among these types of diagrams, a certain type became gradually influential.

This historical aspect may be related to the formation of concepts of graph. In Kőnig's treatise of 1936, the same concepts of graph are attached to the notion of graph in the different problems. In other terms, the concepts of graph allowed viewing problems that, when they appeared, looked unrelated, as depending on the same concepts attached to a single notion of graph, and thereby as related. In fact, in examining the changing solution given to these problems along the 19th century and early 20th century, one can identify a process of progressive integration of the problems throughout the various publications of mathematical recreations in which they were treated. In these publications, mathematicians introduced concepts that allowed to unify problems progressively, and that entered in the shaping of the concepts attached to the same notion, that is, the notion of graph in Kőnig's treatise of 1936. In the present article, however, we will not discuss further the relation between the concepts and the diagrams.

2 Kőnig's Way to Use the Diagrams

2.1 Why Did Kőnig Use the Diagrams in 1936?

The treatise *Theorie der endlichen und unendlichen Graphen* in 1936 [29] makes explicit the reason why Dénes Kőnig used diagrams in it. Since this gives us information as to how Kőnig viewed the diagrams he was using, let us first examine what the author said about them.

Kőnig treated various recreational problems and discussed them as graph theory in this treatise.

In the preface, he explained that graph theory can be understood from two standpoints: one standpoint views the topic as the first part of general topology, while the other understands it as a branch of *combinatorics* and abstract *set theory*.

If this treatise had embraced the first standpoint, that of general topology, we could expect that diagrams would be used for representing topological concepts.

However, Kőnig actually took the latter standpoint, that of *combinatorics* and abstract *set theory*. He explained the reason for this decision as follows:

> Kőnig: *Theorie der endlichen und unendlichen Graphen* in 1936 [29], Preface. (My translation from German.)
> In this book, we take this second standpoint, mainly because we attribute to the elements of graphs—points and edges—no geometrical content at all: the points (vertices) are arbitrary distinguishable elements, and an edge is nothing else but a unification of its two endpoints. This abstract point of view—which Sylvester (1873) emphasized already[2]—will be strictly kept in our representation, with the exception of some examples and applications.

In spite of Kőnig's decision to take the standpoint of *combinatorics* and abstract *set theory*, he used in his treatise a geometrical way to represent elements of a graph (points and edges). Moreover, he introduced diagrams that were not presupposing any geometrical point of view or any geometrical axioms.

We raise therefore naturally a very simple question: why did Kőnig, in this treatise, use a geometrical way of representing parts in a graph and diagrams, although no geometrical content is attributed to the elements of graphs?

Kőnig himself answered half of the question. He made clear that he used a geometrical way of representing elements of a graph because it gave him a very comfortable terminology.

The question thus remains: how were the diagrams to be read if they were not geometrical?

I suppose that Kőnig used diagrams for representing the geometrical notation used in this treatise, even though neither any "geometrical point of view" nor any "geometrical axiom" was presupposed.

To inquire further into this supposition, I will examine in Sect. 2.2 some of the diagrams shown in this treatise.

The book *Graph Theory, 1736–1936* (Norman L. Biggs, E. Keith Lloyd and Robin J. Wilson, 1976 [16]) contains most of the source material with which we shall deal later on. However, they interpret the early-day problems using the diagrams of modern graph theory. Therefore, they bypass the question of the emergence of these diagrams. This is the question with which I shall reconsider this source material and other documents.

2.2 Diagrams in Kőnig's Treatise of 1936

Let us consider here how Kőnig used diagrams in *Theorie der endlichen und unendlichen Graphen* in 1936. The point will be to compare these diagrams with those he himself published before 1936, and those in the publications by other mathematicians. As I wrote

[2]Kőnig made here a reference to the article by Sylvester entitled "On recent discoveries in mechanical conversion of motion" in 1873 [47]. This article treated a mode of producing motion in a straight line by a system of pure link-work without the aid of grooves or wheel-work, or any other means of constraint than that due to fixed centers, and joints for attaching or connecting rigid bars. Maybe here Kőnig had the following part of Sylvester's article in mind: "The theory of ramification is one of pure colligation, for it takes no account of magnitude or position; geometrical lines are used, but have no more real bearing on the matter than those employed in genealogical tables have in explaining the laws of procreation. [New paragraph] The sphere within which any theory of colligation works is not spatial but logical—such theory is concerned exclusively with the necessary laws of antecedence and consequence, or in one word of *connection* in the abstract, or in other words is a development of the doctrine of the compound parenthesis." (*The Collected Mathematical Papers of James Joseph Sylvester*, vol. 3, pp. 23–24.)

in Sect. 1, I will select only the diagrams used in the problems of bridges, polygons, dominoes and mazes, since these problems were also treated in many publications before his treatise of 1936, and we therefore have enough material to compare with Kőnig's treatise in 1936.

These problems are found in the books on mathematical recreations written by some mathematicians. In these books, certain mathematical problems were collected under the concept of "recreation", "pleasure", "delectation", "leisure", "amusement", "game" or "curiosity".[3]

Kőnig also published two books on mathematical recreations in 1902 and 1905 [27, 28].[4] The book of 1905 includes the problem of bridges which we will discuss in Sect. 2.3. This provides us with the publications to which he had access in his youth and allows us to analyze the evolution of his approach to some problems between 1905 and 1936.

The works of Édouard Lucas (1842–1891) on mathematical recreations were the topic of research by some historians. According to Anne-Marie Décaillot, Lucas was attracted by "geometry of situation", and, from the problems which had been considered as "geometry of situation", he drew recreations, but without any analysis ([19], p. 5). The "geometry of situation" was not yet well structured at that time (Pont: *La topologie algébrique des origines à Poincaré* [39], Epple: "Topology, matter, and space I: topological notions in 19th-century natural philosophy" [21]).[5]

Through the examination of diagrams used in the above mentioned four problems, we will clarify how a part of the "geometry of situation" was structured, and which diagrams were involved in this process.

---

[3]Some of the collections in this genre were cited by Kőnig in his books on mathematical recreations in 1902 and 1905 [27, 28], and some of them were cited even in his treatise of 1936 [29]: *Problèmes plaisans et délectables qui se font par les nombres* by Claude-Gaspar Bachet de Méziriac [5–8] in the early 17th century; *Récreations mathématiques et physiques* by Jacques Ozanam [36, 37] in the 17th century; *Récreations mathématiques* (4 volumes) and *L'Arithmétique amusante* by Édouard Lucas [31–35] and *Mathématiques et mathématiciens: pensées et curiosités* by Alphonse Rebière [40, 41] in the 19th century; *Mathematical Recreations and problems of past and present times* (*Mathematical Recreations and essays* for the 4th edn. and later) by Walter William Rouse Ball [9–15] (and many later editions), *Mathematische Mußestunden* by Hermann Schubert [43–45] and *Récréations arithmétiques* and *Curiosités géométriques* by Émile Fourrey [24, 25] around 1900; *Mathematische Unterhaltungen und Spiele* and *Mathematische Spiele* by Wilhelm Ahrens [1–4] in the early 20th century. Also in the period close to the publication of Kőnig's treatise of 1936, a book in this genre was published in Belgium: *La mathématique des jeux ou récréations mathématiques* by Maurice Kraitchik [30], which might suggest interest of mathematicians at that time in mathematical recreations.

Anne-Marie Décaillot made historical researches on Lucas' works [18–20].

David Singmaster made a precise examination of all the editions of Ball's book [46].

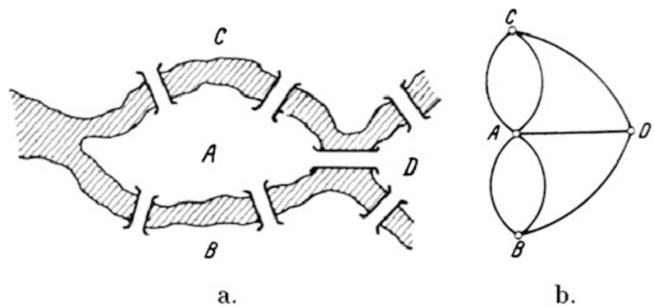
Albrecht Heeffer [26] worked on a movement toward creation of this genre in the 17th century.

In the UK and the USA in the 19th and the 20th century, amateur mathematicians or scientific writers—Samuel Loyd, Henry Ernst Dudeney, Martin Gardner and so on—also published in this domain. Some problems in their collections were taken from the collections listed above. This fact is interesting from the point of view of the popularization of mathematics, but, for the purpose of this article, we don't need to examine precisely these works.

[4]I translated these Hungarian books into English in my thesis [51].

[5]As Décaillot says, "among the mathematical games and recreations of Euler, the traces of strings on the chessboard of Vandermonde in the 18th century, the 'higher mathematics' of Riemann and the Analysis situs of Poincaré, the geometry of situation has a fluctuating content which is structured only slowly during the 19th century" ([19], p. 129, my translation from French).

Fig. 1 Taken from Sect. 2.2 of Kőnig's book of 1936 [29]



2.3 *Appearance of Graph-Like Diagram for the Problem of Seven Bridges of Königsberg*

In the chapter entitled "Eulersche und Hamiltonsche Linien (Eulerian and Hamiltonian lines)" of the book of 1936, Kőnig treated the problems of the so-called Eulerian circuits and the Hamiltonian circuits. An Eulerian circuit is a closed path which goes through each edge of a graph once and only once; a Hamiltonian circuit is a closed path which goes through each vertex of a graph once and only once.

Kőnig gave some theorems concerning Eulerian circuits in the first section of this chapter. One of the theorems is as follows:

> "One can go through all the edges of a graph in a closed path if and only if the graph is a connected Euler graph" (Theorem 2, my translation from German).

The terms appearing in this theorem—edge, graph, closed path, connect etc.—were defined in Chap. 1 using the concepts of set theory. It was proved by contradiction. In the proof, no diagram was used.

In this context, in the second section entitled "Das Brücken- und Dominoproblem (The problem of bridges and the problem of dominoes)", Kőnig mentioned the problem of seven bridges of Königsberg as an example of an application of the theorems.

This problem was first mathematically considered and published by Leonhard Euler (1707–1783). The problem is as follows: in Königsberg, there was a river flowing from east to west; across the river, there were seven bridges as shown in **a** of Fig. 1; is there a route to cross every bridge once and only once, or, more generally, is there such a route for any other forms of rivers and bridges?

Kőnig introduced this problem using a simplified map as **a** of Fig. 1, and then he gave a diagram as **b** of Fig. 1, which consists of small circles representing vertices corresponding to land areas, and straight or curved lines representing edges corresponding to bridges. In this way, the diagram shown in **b** of Fig. 1 represents geometrical elements which correspond to a graph. I will call such a diagram a "graph-like diagram".

In the diagram in Fig. 1.**b**, Kőnig displayed only the elements necessary *for solving* the problem. From this diagram, we see that each vertex is connected to an odd number of edges. Kőnig concluded, using the theorem of graph theory mentioned above, that there is not such a way to cross every bridge just once.

This diagram is thus useful *for solving* the problem using geometrical concepts representing a graph.

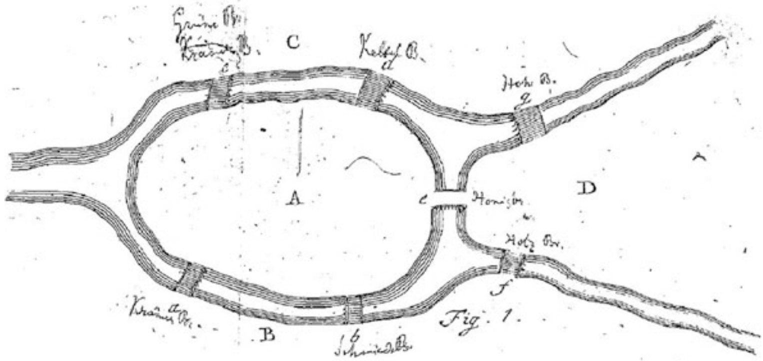


Fig. 2 Taken from Euler's article in 1736 [22]

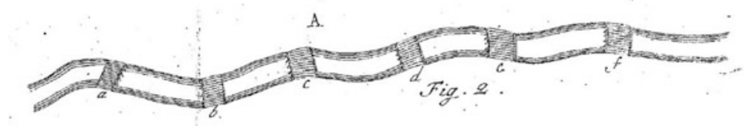


Fig. 3 Taken from Euler's article in 1736 [22]

Originally, how did Euler deal with the problem in his article "Solutio problematis ad geometriam situs pertinentis (solution of a problem relating to the geometry of situation)" in 1736 [22]?

The article of Euler consists of twenty-one sections as follows, each of which has no title.

§ 1: Aim of this article. This article gives a specimen of *Geometriam Situs*.

§ 2: Introduction of the problem. Seven bridges connect four land areas as shown in Fig. 2.
Is there a route to cross every bridge once and only once? More generally, is there such a route for any other forms of rivers and bridges?

§ 3: Choice of the way to solve the problem. One can solve the problem by checking over every possible course, but Euler chooses a simpler method, with which one will find if such a route exists or not.

§§ 4–5: Symbolization of the objects of the problem. Euler gives symbols A, B, C, D to the land areas and $a, b, \ldots, g$ to the bridges. Each route is described by a sequence of symbols of land areas in the order of passage. For example, ABD is a route depart from A via B to D, no matter which bridges are crossed. To describe a route to cross seven bridges, eight symbols are necessary.

§§ 6–9: Solution of the problem of seven bridges of Königsberg. Because there are two bridges between A and B, the sequence of symbols of the route demanded should include two sets of adjacent A and B. The same consideration is applied to the other bridges. Euler tries to find a law to judge if such a sequence of eight symbols exists or not. Suppose that a traveler crosses the bridge a of Fig. 3. In the sequence of the route, A appears once no matter if the traveler departs from A or arrives at A. Similarly, if A

has three bridges and the traveler crosses them, A appears twice in the sequence of the route. Generally, if A has any odd number of $2n + 1$ of bridges, A appears $n + 1$ times in the sequence of the route. In the case of Königsberg, A has five bridges, and each of B, C, D has three bridges. Therefore A appears three times, and each of B, C, D appears two times in the sequence of route. But such a sequence cannot be realized with only eight symbols. This means that there is no route to cross every bridge of Königsberg once and only once.

§§ 10–13: Generalization of the problem. Euler generalizes the problem to all the forms of bridges and land areas. Euler says that, if the sum of numbers of symbols to appear in the sequence is larger than "(the number of bridges) + 1", it is impossible to find such a route to cross every bridge once and only once. Euler says also that it is possible if the sum of numbers of symbols to appear in the sequence is equal to "(the number of bridges) + 1", but it is not proved.

In the case where A has an even number of bridges, we should consider if the traveler starts from A or not. If the traveler does not start from A which has $2n$ bridges, A appears n times in the sequence of symbols of the route. If the traveler starts from A which has $2n$ bridges, A appears $n + 1$ times in the sequence of symbols of the route. Put off considering the starting point, then a symbol of a land area appears $n + 1$ times if the land area has any odd number $2n + 1$ of bridges, and n times if the land area has any even number $2n$ of bridges in the sequence of symbols of the route. If the sum of the symbols to appear in the sequence is equal to "the number of bridges + 1", there is a route to cross every bridge once and only once, where the starting point cannot be any land area which has any even number of bridges. If the sum of the symbols to appear in the sequence is equal to "the number of bridges", there is a route to cross every bridge once and only once, where the starting point should be a land area which has any even number of bridges, so as to increase by 1 the number of symbols to appear in the sequence. But Euler gives no proof for the case where such a route exists.

§ 14: Invention of an algorithm depending on §§ 10–13. Euler gives an algorithm that he says can be used to know if one can cross every bridge once and only once in any form of rivers and bridges. But actually, it is rather an algorithm to know if such a route is impossible or not, because Euler gives no proof of the fact that where such a route exists.

1. Label the land area with symbols $A, B, C, \ldots$.
2. Write down "the number of bridges + 1".
3. Make a table with a column which consists of $A, B, C, \ldots$, and with the next column which consists of the number of bridges connected to each land area.
4. Asterisk the symbols of land areas which have any even number of bridges.
5. Make another column which consists of:
   - n if the land area has any even number $2n$ of bridges,
   - $n + 1$ if the land area has any odd number $2n + 1$ of bridges.
6. Sum up the numbers of the last column. If the sum is equal to the number written in the step 2, or if the sum is less by 1 than it, a route to cross every bridge once and only once is possible, where in the former case, the starting point should be one of the land areas without asterisk; in the latter case, the starting point should be one of the land areas with asterisk.

Euler makes such a table for the problem of Königsberg, and concludes that such a route is impossible.

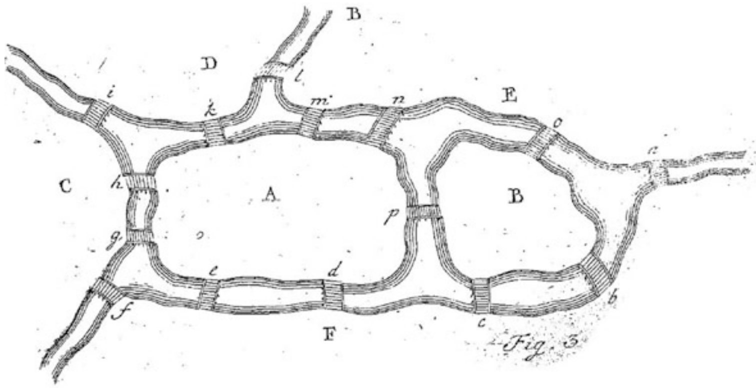


Fig. 4 Taken from Euler's article in 1736 [22]

§ 15: Example different from seven bridges of Königsberg. Euler gives an example where a route to cross every bridge once and only once exists (Fig. 4), applies the above-mentioned algorithm to it, and gives such a route. But we should still note that he gives no proof for the case where such a route exists.

§§ 16–17: Proof of the handshaking lemma. To obtain a simpler way to judge if a route to cross every bridge once and only once exists or not, Euler proves that the number of land areas which have any odd number of bridges is an even number. This is called the handshaking lemma in our times.

Proof: Count bridges which each land area has. The sum of these numbers is just twice as many as the number of all the bridges, because every bridge connects just two land areas, and it is counted double. Therefore the sum of the numbers of bridges which each land area has is an even number. If the number of land areas which have any odd number of bridges is an odd number, the sum of the numbers of bridges which each land area has cannot be an even number. So the number of land areas which have any odd number of bridges is an even number.

§§ 18–19: Simplification of the algorithm of § 14. Because the sum of the numbers of bridges which each land area has is twice as many as the number of all the bridges,

$$\frac{\text{(The sum of the numbers of bridges which each land area has)} + 2}{2}$$

is just the number written in the step 2 of § 14. If every land area has an even number of bridges, the sum in the step 6 of § 14 is less by 1 than the number written in the step 2, with which Euler means that a route to cross every bridge once and only once is possible. Because every land area has any even number of bridges, any of them can be the starting point.

If only 2 land areas have any odd number of bridges and the other areas have any even number of bridges, the sum in the step 6 of § 14 is just as same as the number written in the step 2, with which Euler means that a route to cross every bridge once and only once is possible. In this case, one of the land areas which have any even number of bridges should be the starting point.

But we should still note that he gives no proof for it.

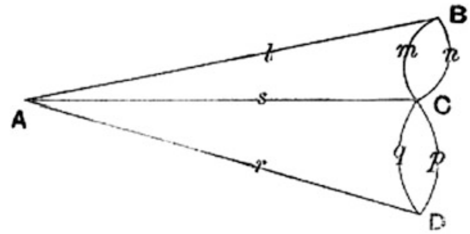


Fig. 5 Taken from Chap. 6 of Ball's 2nd edn. in 1892 [9]

If the land areas which have an odd number of bridges are 4, 6, 8 or more, the sum in the step 6 of § 14 is larger by 1, 2, 3 or more than the number written in the step 2. Then there is no route to cross every bridge once and only once.

§ 20: Summary of §§ 18–19. If more than two land areas have any odd number of bridges, it is impossible to cross every bridge once and only once.

If just two land areas have any odd number of bridges, and if the traveler choose one of such land areas as the starting point, it is possible to cross every bridge once and only once.

If every land area has any even number of bridges, no matter which land area is chosen as the starting point, it is possible to cross every bridge once and only once.

But we should still note that Euler gives no proof for the two latter propositions.

§ 21: Method to simplify the way to find the route. Remove pairs of bridges which connect two common land areas, and it will be easier to find a route to cross every bridge once and only once. After finding the route, put back the removed bridges as they were, and it will be easy to modify the route so as to include them.

These are all the sections Euler gave.

Euler introduced a map illustrating the situation. On the map, Euler displayed symbols for his proof. Euler's proof needed only the symbols of the land areas and the number of bridges connected to each land area. For him, the names of bridges were not necessary. Yet we see that Euler kept this information on the map (Fig. 2). Moreover, the proof does not make any reference to a diagram. In fact, more precisely, Euler did not give any graph-like diagram for this problem.

The information that was not necessary for the solution to the problem was to be removed from the diagrams included in the texts of subsequent mathematicians who addressed the problem. Indeed, the problem will be taken up in several publications devoted to mathematical recreations.

Let us consider them since this analysis will put in a situation to determine who first introduced a graph-like diagram in this context and how it influenced Kőnig for this feature of the diagrams.

In 1851, Émile Coupy translated this article of Euler into French [17]; in 1882, Édouard Lucas translated it again into French in the chapter about the problem of bridges in vol. 1 of his series on mathematical recreations [31]. However, neither Coupy nor Lucas made any significant modification to Euler's figures.[6]

In 1892, Walter William Rouse Ball (1850–1925) dealt with the problem within a more general context, since he mentioned it in the chapter about "unicursal problems" in

[6] I examined the difference of the translations of Coupy and Lucas in Chap. 3 of my thesis [51].

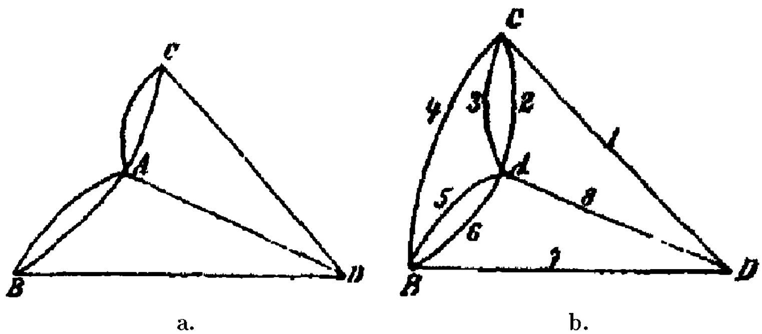


Fig. 6 Taken from Sect. 17.1 of Ahrens' book of 1901 [1]

his book devoted to mathematical recreations [9]. In this new context, Ball gave a new kind of diagram (Fig. 5) for the problem. In this diagram, Ball represented the bridges with lines—some lines curved and others straight—indicated with lowercase letters, and the land areas with points indicated with uppercase letters. Ball mentioned the correspondence of the notation introduced for bridges to Euler's map, but he did not use it in his consideration of the problem, and indeed it is not necessary for solving the problem. However, it is remarkable that lines in this diagram represent geometrical elements which correspond to a graph. As a diagram given to the problem of seven bridges of Königsberg, this is perhaps the first graph-like diagram.

In 1901, Wilhelm Ahrens (1872–1927) also, in his book on mathematical recreations [1], treated the problem of bridges. However, for him, the problem appeared in the chapter about "Brücken und Labyrinthe (bridges and labyrinths)", that is, within a classification of problems made on the basis of their topic. Still, he gave here diagrams quite similar to Ball's graph-like diagram (Fig. 6).

The diagram **a** of Fig. 6 has no mark for lines representing the bridges, but only associates letters to points representing the land areas. Thus we find here only the information necessary for solving the problem. We see that the diagram is drawn for the problem, and not as a representation of a general mathematical object. The diagram **b** of Fig. 6 represents the case with eight bridges, where one can pass over all the bridges once and only once. The diagram **b** has digits attached to lines, but their meaning is different from the lowercase letters shown in Ball's diagram: Ahrens let these numbers represent the order of passing over the bridges, therefore these numbers are necessary information for representing a solution to the problem.

In 1905, Kőnig himself also, in one of the books that in his youth he devoted to mathematical recreations in 1905—long before the publication of his treatise of 1936—treated the problem of bridges in the chapter about "A Königsbergi hidak (the bridges of Königsberg)" [28]. In relation to this problem, the book of 1905 quotes Euler [22], Lucas [31], Ball [11], Schubert [45], Ahrens [1]. However, as for the diagram given to this problem, he took it from Ahrens' book (Fig. 7). Fig. 7 is almost the same as Ahrens' graph-like diagram.[7] I suppose that Kőnig's diagram for the problem of bridges in 1905 was taken from Ahrens' diagram in 1901.

[7]A part of the line between A and B of Kőnig's diagram is not connected, but it is only an error of printing. It should be continuous for consistency of the text.

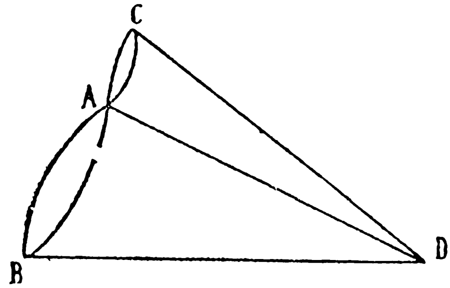


Fig. 7 Taken from Chap. 4 of Kőnig's book of 1905 [28]

It is interesting that this diagram of Kőnig in 1905 was still different from his diagram in his treatise in 1936 where, instead of the points indicated with uppercase letters, small circles were used, which represented the vertices of a graph. We will see that the representation of the vertices of a graph with small circles and the full notation of the elements of a graph in 1936 betray an influence different from the problem of bridges.

2.4 Polygons, Dominoes and the Introduction of the Flexible Strings

In the same section entitled "Das Brücken- und Dominoproblem (The problem of bridges and the problem of dominoes)" in Kőnig's treatise of 1936, he treated two other problems coming from mathematical recreations—one bearing on polygons and another one on dominoes. He dealt with them as other examples for his theorem on Eulerian circuits. Examining the history of the treatment of these problems and of their relation to each other will show how another feature of the diagrams for graphs took shape within this context.

The problem on polygons is formulated as follows: a polygon being given, can we go along every edge and every diagonal just once with only one stroke?

As for the problem on dominoes, it is formulated as follows: one set of dominoes consists of twenty-eight pieces; on each piece, a pair of integers from 0 to 6 is shown; we put aside here the double numbered pieces with $(0, 0)$, $(1, 1)$ etc. because they play no part in the question considered; the question is to arrange all the remaining twenty one pieces so that adjacent numbers are equal to each other.

Kőnig related the three types of problems to each other, which can be done when one reformulates them in terms of problems related to graphs. Previously, they were not precisely discussed together in the same context. As we will see soon, Kőnig was not the first one to have perceived their link, but he was the first mathematician to treat them explicitly as related problems, and he did this on the basis of diagrams of graph theory. Moreover, in the context in which the problems were understood as being connected with each other, another feature of the diagram took shape: the nature of the lines representing edges to be "flexible strings". Let us explain what we mean by these statements.

For the case of a heptagon, Kőnig represented the problem of polygons by a diagram shown in Fig. 8. We can trace every edge and every diagonal of this diagram just once with only one stroke.

Here, the diagram was used *for solving* the problem.

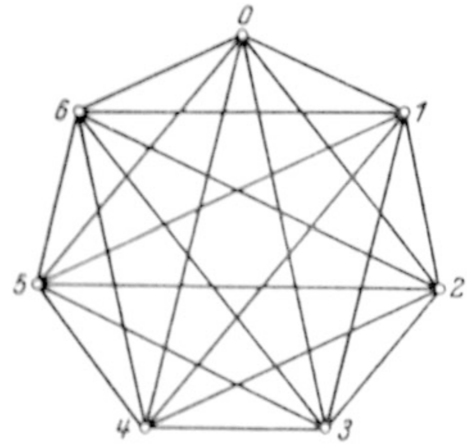


Fig. 8 Taken from Sect. 2.2 of Kőnig's book of 1936 [29]

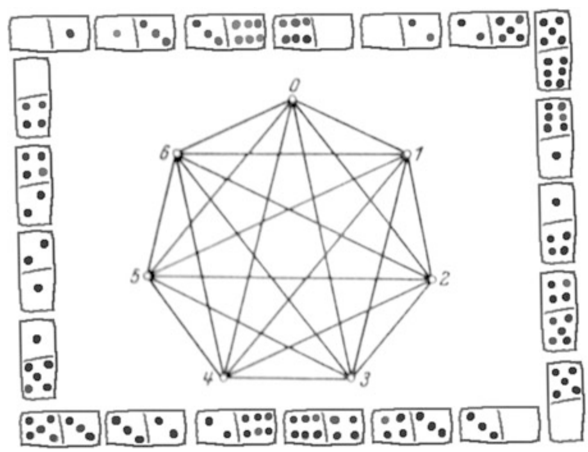


Fig. 9 Domino pieces which I arranged so as to correspond to Kőnig's diagram shown in Fig. 8

Moreover, it was also by using the same diagram of a heptagon that Kőnig solved the problem of dominoes, thereby displaying the link between the two. He let each vertex of a heptagon correspond to a number on a domino piece, and each edge of it to one domino piece. By means of this representation, the solution to the problem of dominoes corresponded exactly to the solution to the problem of a heptagon. Kőnig did not give any specific diagram for dominoes, but we understand easily the relation of the solutions to these two different problems in Fig. 9.

What is important is that the problem of polygons and the problem of dominoes were not treated as being the same in any of the previous mathematical texts in which they both appeared.

In 1809, Louis Poinsot (1777–1859) treated the problem of polygons in his lecture about "*les polygones et les polyèdres* (the polygons and the polyhedrons)". This lecture was published as a memoir in 1810 [38]. The problem we consider was described in Sect. 18 of this memoir.

Poinsot described the problem as follows:

> Poinsot (1809/1810) [38], Sect. 18, pp. 28–29, my translation from French.
> [...] The problem is, between points placed in the space as you like, to lead a same flexible string [*fil flexible*] which unites the points two by two in all the possible ways, so that the two ends of the string come to be rejoined at the end, and that the total length[8] should be equal to the sum of all the mutual distances.

Poinsot explained why the solution is possible only for an odd number of points. He did so, using the concept of a "flexible string": when the points are even in number, one can still lead a string which connects the points two by two in all the possible ways, but this string should pass twice from any of the points to any other, before the two ends be rejoined and, the string being closed, the total length will be equal to twice all the mutual distances of the proposed points.

Poinsot treated this problem with points in a space, which means the points and the flexible string do not necessarily form a polygon. However, in the succeeding sections in his memoir, he discussed this problem in the case that the points are projected onto a plane. By projection of the points onto a plane, we consider this problem as that of polygons. In fact, in Sect. 23, Poinsot related this problem in the case of four points to a quadrilateral with two diagonals; and finally in Sects. 24 and 25, he applied this problem to arbitrary polygons.

In his publication, Poinsot used no diagram to discuss this problem. It is nevertheless remarkable that he used the concept of a "flexible string" *to describe* the path and *to solve* the problem.

In fact, the idea of flexible strings can be traced back to an article by Alexandre-Théophile Vandermonde (1735–1796), which Poinsot mentioned at the beginning of the memoir. He wrote:

> Poinsot (1809/1810) [38], pp. 16–17, my translation from French.
> [...] *Vandermonde* gave, in the Memoirs of the Academy of science for 1771, a simpler solution,[9] which was deduced from a particular notation which he invented for this sort of problem, and which he applied also to the representation of a textile or net formed with the successive knots of several strings. [...]

We suppose therefore that Poinsot got the concept of flexible strings from Vandermonde's article "Remarques sur des problèmes de situation (remarks on the problems of situation)" in 1771 [50].

However, Vandermonde did not treat polygons in his article of 1771. Therefore Poinsot took his idea of flexible strings, and used it in a completely different context. Moreover, as we saw in the memoir of Poinsot in 1810 (see p. 183), Poinsot did not only use what Vandermonde used, but he added measures of length, that is, "geometry of situation" was lost in Poinsot's memoir.

On the other hand, Vandermonde's concern was the notation to be used by the workers who make a *braid*, a *net*, or *knots*. These workers do not conceive these spatial situations

---

8 Despite the fact that Poinsot speaks of a flexible string, he uses the length. This means that one looses the "geometry of situation".

9 Poinsot mentioned Vandermonde [50] in the context of the problem of the knight's move on a chess-board, as one of the problems concerning the "geometry of situation". He meant that Vandermonde's solution was simpler than Euler's [23].

in terms of size, but in terms related to the situation of strings with respect to each other. What the workers see is the order in which the strings are interlaced.

Vandermonde attempted to create a system of notation more to conform to the process of the worker's mind. This notation was the basis on which he would work out a solution for the problem. For this purpose, he needed a notation which would represent only the idea formed from his work, which could be sufficient for again making a similar thing at any time.

The object of Vandermonde's article of 1771 was only to give a hint of the possibility of this kind of notation, and its usage in questions related to textiles composed of strings.

For this purpose, Vandermonde described each point on a string with its spatial position. To represent the spatial position, one splits a 3-dimensional space into parallelepipeds. Each parallelepiped is indicated with a triple of numbers—Vandermonde called it "trois nombres assemblés, ainsi c_a^b (three numbers gathered, so that c_a^b)"—each term of which corresponds to a position of the parallelepiped on each axis of the space. By putting the triples in the order of the parallelepipeds where a string passed through, one gets a sequence of triples, which denotes a form of the string.

From such a sequence of triples, One can reproduce the textile or knots by making a string go through the parallelepipeds indicated by the triples in order.

Vandermonde applied this notation to a 2-dimensional space for solving the problem of the knight's move on a chessboard:

> Vandermonde (1771) [50], p. 568, my translation from French.
> Let the *knight* go all over the squares of a chessboard without visiting twice the same square, as a result determine a certain trace of the *knight* on the chessboard; or else, supposing a pin fixed at the center of each square, determine the course of a string passed one time around each pin, according to a law from which we will search the expression.

Vandermonde let the trace of a knight correspond to a string, each square to a pin.

To trace a knight on a chessboard, the above mentioned notation is applied. Because the trace is on a plane, the sequence corresponding to the trace is a sequence of pairs of numbers, each number of which consists of any of the numbers 1, 2, 3, 4, 5, 6, 7 and 8.

A knight's move on a chessboard is denoted as $\begin{smallmatrix} b & b\pm1 \\ a & a\pm2 \end{smallmatrix}$ or $\begin{smallmatrix} b & b\pm2 \\ a & a\pm1 \end{smallmatrix}$.

To simplify the solution, we use the symmetry of the knight's trace: if we create a sequence of a knight's move, and interchange the numbers of the pairs: 8 to 1, 7 to 2, 6 to 3, 5 to 4 and vice versa. Then we will get a new sequence denoting a trace of knight symmetry to the original one.

Therefore, to obtain a knight's trace visiting all the squares on a chessboard once and only once, we first need to create only a trace within the squares denoted with a sequence of pairs with numbers 1, 2, 3 and 4, and then we get another sequence by exchanging the numbers of one axis, still another sequence by exchanging the numbers of the other axis, and the other sequence by exchanging the numbers of both axis.

We thus obtain four sequences denoting four separate traces which, as a whole, visiting all the squares on a chessboard once and only once.

These four sequences can be connected, without breaking the knight's move, by joining two sequences, or by inserting one sequence between two pairs of another sequence.

We should pay attention to the fact that, *for solving* the problem, Vandermonde did not use the concepts of "strings" and "pins", but sequences of numbers which denote strings.

In spite of not using the formulation in terms of a string and of pins *for solving* the problem, Vandermonde gave a diagram corresponding to these concepts *to represent his result*

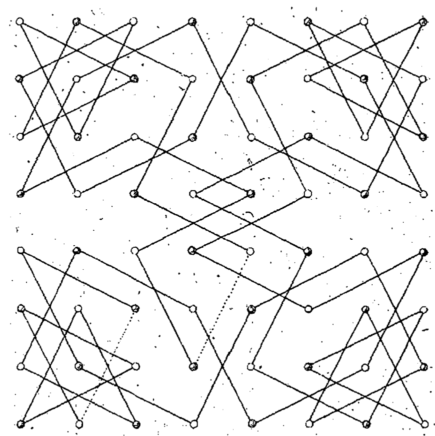


Fig. 10 Taken from Vandermonde's article in 1771 [50].

regarding "la forme de la trace du cavalier sur l'échiquier, déterminée par cette suite (the form of the trace of the knight on the chessboard, determined by this sequence)". Compare Fig. 10. We see in this diagram that Vandermonde used small circles representing pins fixed on a chessboard. These features of his diagram evoke the form of diagram used in Kőnig's treatise in 1936, as well as in the texts on graph theory of the present day.

I suppose that Poinsot, inspired by Vandermonde's article of 1771, had in mind not only the concept of "flexible string" but also the concept of "pins". In fact, Poinsot lead a flexible string between the "points placed in the space as you like", just like Vandermonde let a string go through each "pin fixed at the center of each square" on a chessboard. The way to use the points of Poinsot is the same as the way to use the pins of Vandermonde. In other words, Poinsot's "points" played the role of "pins" in Vandermonde's geometrical representation.

In terms used in Kőnig's treatise of 1936, Poinsot applied these concepts to a problem related to Eulerian circuits, whereas Vandermonde applied them to a problem related to Hamiltonian circuits. In fact, in 1936, Kőnig considered both problems as examples of more general problems: the problem of polygons was treated as a problem of Eulerian circuits, while the problem of a knight's moves was treated as a problem of Hamiltonian circuits.

However, Poinsot, conscious or not, related these two problems with the concepts of a flexible string and pins.

2.5 Polygons and Dominoes Again: A Single Diagram and a Single Way of Using It for Two Distinct Problems

Later on, Kőnig went further: he not only described the recreational problems on the basis of the common concepts of graph theory, but he also formulated the general problems bearing on the general object of graphs for which they were particular cases.

Poinsot's lecture of 1809 mentioned no relation between the problem of polygons and the problem of dominoes. In 1849 [49], Orly Terquem (1782–1862) published the commentary on Poinsot's works on polygons. In this commentary, he alluded the relation between the problem of polygons and the problem of dominoes by mentioning the question

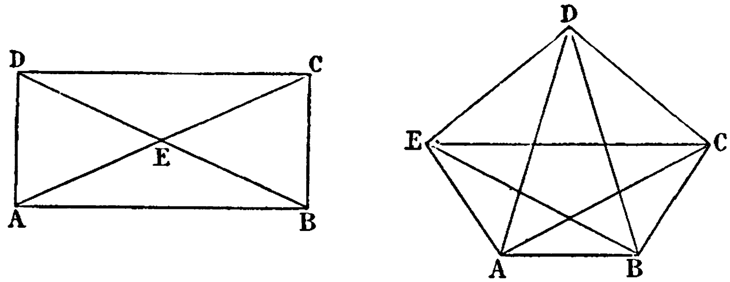


Fig. 11 Taken from Chap. 2 of Lucas' vol. 1 in 1882 [31]

of dominoes just after the description of Poinsot's problem of polygons that we examined in Sect. 2.4. At the end of the commentary, Terquem described as follows:

> Terquem (1849) [49], p. 74, my translation from French.
> [...] The determination of the *number* of possible solutions for an odd number of points is a problem of which the solution is desired. I proposed it to some distinguished geometers, but I got nothing. The domino game presents a question of this type: in how many ways can one place all the dominoes on only one line obeying the law of the game? One can put aside the *double*-numbered pieces.

Clearly, Terquem gave neither any precise description of the relation, nor any diagram representing this idea. However, this seems to be the first mention of a relation between the problem of polygons and the problem of dominoes.

Similarly, regarding the relation between the problem of polygons and the problem of dominoes, already in 1883 Lucas was aware of the relation between the problem of a heptagon and the problem of dominoes, because he mentioned it in his note to the chapter about "Le jeu de dominos (the domino games)" put at the end of vol. 2 of his series of mathematical recreations [32]. But he did not give any precise description about this relation at that time.

Finally in 1894, in the chapter about "La Géométrie des réseaux et le problème des dominos (the geometry of nets and the problem of dominoes)" in vol. 4 of his series of mathematical recreations [34], Lucas declared that the idea of relating a heptagon to the problem of dominoes mentioned in vol. 2 was given by Laisant.[10]

If we now go back to the history of the relationship between the problem of bridges and that of polygons, we note that in 1882 Lucas treated the problem of polygons with the diagrams shown in Fig. 11. Further, he included the problem in vol. 1 of his series on mathematical recreations, in the chapter devoted to the problem of bridges [31].

In this case, it is not by making use of common terms to formulate different problems that Lucas indicated something common between them. Lucas did so by classifying them in the same chapter of his book. This fact indicates that Lucas recognized that both problems could be treated in the same way.

Moreover, Lucas described in this chapter relationships between a wider set of problems, since he stressed the relations between four different topics of mathematical recreations—bridges, mazes, polygons and dominoes.

[10] Charles-Ange Laisant (1841–1920) was a mathematician, and was a director of some reviews of mathematics.

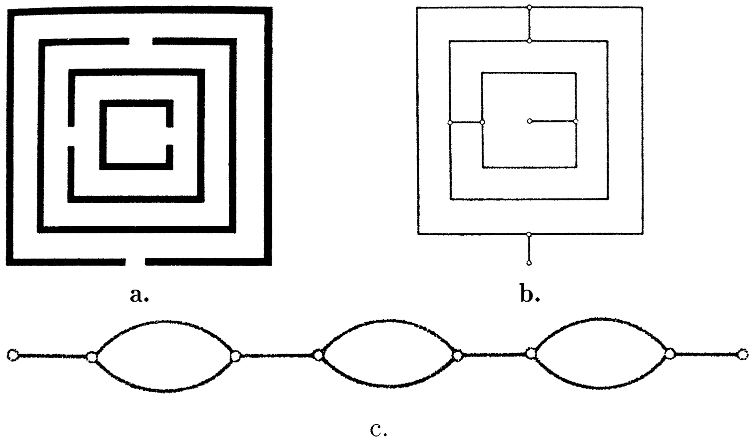


Fig. 12 Taken from Sect. 3.1 of Kőnig's book of 1936 [29]

However, Lucas did not discuss these four topics explicitly on the same basis, while Kőnig did on the basis of diagrams of graph theory.

The links that Lucas could establish between these topics depended on ideas that Tarry had presented at a conference in 1886 [48], which we will discuss more completely in Sect. 3.

Let us first consider the history of the treatment of the fourth type of problem that Lucas linked to the first three considered above.

2.6 Mazes of Which the Junctions Became Important

Kőnig's treatise from 1936 also included a chapter entitled "Das Labyrinthenproblem (The problem of mazes)". In it, he treated the following problem: how can I arrive at a certain place—a branching point or a loop—in walking in a maze without a map? He gave three different diagrams for the same example: in the diagram **a** in Fig. 12, the lines are the walls of a maze, and a traveler of the maze goes through the space between the lines; in the diagram **b** in Fig. 12, the lines and the small circles are paths and junctions of a maze, which is represented by means of the edges and vertices of a graph; the diagram **c** in Fig. 12 is a transformation of the diagram **b**. With diagram **c**, Kőnig showed that the absolute position of vertices and edges are ignored in graph theory, and that only the relation between the vertices and the edges is important for solving the problem.

A solution to this problem had been first published in 1882 by Lucas in the chapter about "Labyrinthes (mazes)" in vol. 1 of his series on mathematical recreations [31]. Lucas says about it that this solution had been given by Trémaux, a telegrapher and a former student of the polytechnic school. Lucas included a proof of the correctness of the solution. However, in Kőnig's treatise of 1936, he noted that "Dieser Lucassche Beweis ist nicht vollständig (this Lucas' proof is not complete)".

Let us concentrate on the diagrams used by Lucas. In the proof to the solution, Lucas used diagrams representing a part of a maze. One of Lucas' diagrams is shown in Fig. 13.

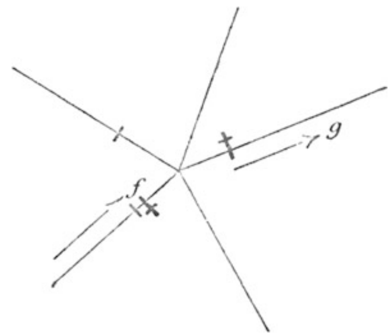


Fig. 13 Taken from Chap. 3 of Lucas' vol. 1 in 1882 [31].

In Lucas' diagrams, the lines represent "chemins (paths)", and points where lines intersect represent "carrefours (junctions)".

We see in this diagram other characteristics: the arrows indicate the directions of the walk; the marks crossing the lines indicate the paths through which one already walked. In other words, we have here a graph-like diagram with further marks.

Lucas used lines representing paths of a maze in 1882, just like Kőnig did in his treatise of 1936, using for this problem lines which represented edges of a graph. Kőnig thus used a general type of diagram to represent the maze problem whereas Lucas drew a diagram specific to the problem considered. However, both diagrams look alike. Moreover, Lucas did not use small circles in his diagrams for the problem of mazes. This is connected to the fact that Lucas did not consider the vertices of the graph as relevant elements for the solution. On the other hand in Kőnig's diagrams in 1936, small circles represented vertices, corresponding to junctions of a maze, and this representation was used also for all the other problems of bridges, polygons and dominoes. We will see in Sect. 3 how this common form of diagram became used in all the problems of graph theory, and how this detail bears witness to the historical process by means of which Kőnig adopted these representations.

To sum up our conclusions so far, we saw in this section that Kőnig, in his treatise of 1936, discussed different problems of mathematical recreations—bridges, polygons, dominoes and mazes—using the same general concepts attached to graph with the same type of diagram; on the other hand, the earlier works did not treat these topics in the same way.

We will see in the next section that Tarry's talk played an important role in the transformation of the way to use the diagrams.

3 The Significance of Tarry's Talk

I examined texts written by Kőnig and other mathematicians before 1936, and found that Tarry's talk in a conference in 1886 played important role in relation to my question—that is, how the representation of graphs in diagrams of graph theory took shape.

I shall now establish that the role played by Tarry's talk relates to his way of using diagrams.

Let us first say a few words about the person. Gaston Tarry (1843–1893) was a public servant working for the French financial administration in Algiers, and an amateur mathematician. He gave a talk in the 15th session of the *Association française pour l'avancement des sciences* in Nancy in 1886 [48]. Its title was: "Géométrie de situation: nombre de manières distinctes de parcourir en une seule course toutes les allées d'un labyrinthe rentrant, en ne passant qu'une seule fois par chacune des allées (Geometry of situation: number of distinct ways of walking in only one course along all the alleys of a recurring maze, in passing through each of the alleys only once)." Here, by "labyrinthe rentrant (recurring maze)" he meant a maze for which the number of alleys leading to each junction is always an even number.

This problem is different from our problem of mazes. The subject of this problem is, in modern terms, the number of all the Eulerian circuits of the maze read as a graph. That is, Tarry dealt with something related to the problem of bridges using the concepts attached to mazes, for example "walk", "alleys", "junctions" etc.

The proceedings of this session consist of two volumes: Volume 1 presents the abstracts of talks prepared by the secretariat of the Association, and vol. 2 contains the articles written by the speakers. The diagrams corresponding to the articles are placed at the end of vol. 2.

The abstract of Tarry's talk was written by someone else,[11] and it reads as follows:

Editor: secrétariat de l'Association (1887) [48], p. 81 of Part 1, my translation from French.
Mr. TARRY, in Algiers.
On a problem of the geometry of situation. — Mr. Tarry proves two theorems on the figures[12] which one can draw with only one continuous stroke, without interruption nor repetition. These two theorems allow one to find the number of solutions[13] in a very large number of cases; he applies his procedure to the problem of Reiss,[14] on the game of dominoes, and obtained again the results of Doctor Reiss in two pages, while the much longer solution of Reiss occupies 60 pages in No. 4 of the *Annali di Matematica*.
Discussion. — The president of the section[15] emphasizes the extreme elegance and the great simplicity of this new method.

Although dominoes are mentioned in this abstract, there is no mention of dominoes in Tarry's text itself. The details of the problem of dominoes perhaps were given only to the audience of his talk.

The question we need to tackle then is to understand the means by which the concepts and diagrams became possible.

---

[11] It is unclear who the authors of the abstracts were, but someone in the bureau of the section which contained Tarry's talk may have been the author: Président d'honneur: M. le Géneral FROLOW, major général du génie russe (Russian general major of engineering); Président: M. Ed. LUCAS, Prof. de math. spéciale au Lycée Saint-Louis (Professor of higher mathematics at the Saint-Louis High school); Vice-Président: M. Laisant, Député de la Seine, Anc. Él. de l'Éc. Polyt. (Deputy of the Seine, Alumnus of the Polytechnic School); Secrétaire: M. HEITZ, Él. de l'Éc. centr. des Arts et Manufact. (Student of the central school of Arts and Manufacture).

[12] Tarry talked about figures of mazes according to his article in vol. 2 of the proceedings.

[13] The problem treated by Tarry was therefore different from the problem treated by Poinsot, who described the possibility of tracing all the edges and diagonals of a polygon once and only once with one stroke.

[14] Michel Reiss (1805–1869) was a mathematician from Frankfurt who worked mainly on the theory of determinants. He published an article about dominoes "Evaluation du nombre de combinaisons desquelles les 28 dés d'un jeu du domino sont susceptibles d'après la règle de ce jeu" (1871) of 58 pages [42].

[15] That is, Édouard Lucas.

In the text of the proceedings, Tarry proved two theorems and a corollary. He first proved the "Théorème des impasses (Theorem of the dead-ends)". An alley both ends of which lead to an identical junction is called an "impasse (dead-end)". In the terminology of graph theory of today, the "impasse (dead-end)" corresponds to a loop on a vertex. The theorem is as follows:

Theorem of the dead-ends *In a recurring maze, if a dead-end is deleted, then the number of distinct courses of the maze is reduced, and then the number of distinct courses of the reduced maze multiplied by the number of the alleys leading to the junction situated on the deleted dead-end is equal to the number of distinct courses of the primitive maze. Each of the other dead-ends on the junction are also counted as two alleys.*

Let N be the number of distinct courses of the reduced maze. Let $2n$ be the number of its alleys leading to the junction that was situated on the deleted dead-ends. The theorem is written with N and $2n$: the number of distinct course of the primitive maze is equal to $N \times 2n$. The proof is as follows: consider any of the N distinct courses of the reduced maze; in this course, you will pass n times through the junction situated on the deleted dead-end; to walk in the primitive maze, in any of these n passages, you interrupt the walk when you arrive at the junction of the dead-end, walk entirely this dead-end, which is to be done in two different directions, and, after coming back to the junction, complete your walk in the maze; as a result, each of the N distinct courses of the reduced maze will supply $2n$ distinct courses of the primitive maze; the N distinct courses of the reduced maze will supply therefore $N \times 2n$ courses of the primitive maze; evidently, these $N \times 2n$ courses of the primitive maze are all distinct, and there is no other way to walk the primitive maze in only one course; the theorem is thus proved.

And then Tarry gave the following corollary:

Corollary *If* $2(n + k)$ *alleys lead to a junction, and* $2k$ *of them belong to* k *dead-ends, then the number of distinct courses of the given maze is equal to the product of* $n(n+1)(n+2)\ldots(n+k-1)2^k$ *and the number of distinct courses of the reduced maze gotten after deletion of* k *dead-ends of the given maze.*

In fact, if you add successively each of these k dead-ends to the reduced maze, then this procedure gives the numbers of distinct courses successively multiplied by $2n$, $2(n+1), 2(n+2), \ldots 2(n+k-1)$.

For calculating the number of distinct courses, using the theorem of the dead-ends, we eliminate the dead-ends of the given maze, and simplify the calculation to the case of the maze with no dead-ends.

Tarry then proved the following theorem:

Theorem (to reduce junctions) *A recurring maze consisting of* k *junctions without dead-ends is given;* N *is one of the junctions of the recurring maze; let* $2n$ *be the number of alleys leading to* N*; the number of distinct courses of the given maze is equal to the sum of the numbers of distinct courses of* $1 \times 3 \times 5 \times 7 \ldots (2n-1)$ *recurring mazes consisting of not more than* $k-1$ *junctions. These* $1 \times 3 \times 5 \times 7 \ldots (2n-1)$ *new mazes are obtained by the following procedure:*

1. *group the* $2n$ *alleys leading to the junction* N *into* n *pairs of alleys in all the possible ways*;
2. *and then*, *in each of the groups*,
   (a) *replace each pair of alleys with a new alley joining the* 2 *junctions to which the endpoints of the pair of alleys leads*, *or*,
   (b) *in the case that the* 2 *alleys of the pair lead to an identical junction*, *replace the pair of alleys with a dead-end at this junction*.

Tarry proved this theorem as follows: group the $2n$ alleys leading to the junction N into pairs of alleys in all the possible ways; we will get $(2n-1)(2n-3)\ldots\times 5\times 3\times 1$ different groups; to each of these groups, we relate all the courses of the given maze; in these courses, in each passage through the junction N, the alley leading to it and the alley away from it belong to a pair of the group considered; we see easily that the number of courses to be found will be equal to the sum of the numbers of distinct courses corresponding to each group, in the way shown above; consider the courses of one of these groups, and examine the n pairs of alleys that comprise the group; in each of these n pairs of alleys, the two alleys, which are considered as ways out of the junction N, lead to two different junctions A, B or to one identical junction C; in the former case, we replace the 2 alleys NA, NB with a new alley AB that joins the junctions A and B without changing the number of courses, because this change means replacing the track ANB or BNA with the equivalent tracks AB or BA; in the latter case, the two alleys joining the junctions N and C are replaced with a dead-end passing through the junction C; the theorem is thus proved.

After proving these theorems, Tarry gave a procedure to calculate the number of distinct courses of any recurring maze:

1. Apply the "theorem of dead-ends" and "the theorem to reduce junctions" to a given recurring maze. According to the theorems, the number of junctions of the maze will be reduced, and we will get an equation between the number of distinct courses of the reduced maze and that of the maze before it was reduced.
2. Repeat process 1 so that we finally get mazes containing only two junctions without dead-ends.
3. Count the number of mazes containing only two junctions without dead-ends: two junctions of such a maze are connected with $2n$ alleys, therefore the number of distinct courses is equal to $2(2n-1)(2n-2)\ldots 4\times 3\times 2\times 1$ if each direction of the walk is counted.
4. Substitute for the variable of the last equation the number of distinct courses of the last maze, that is, the maze containing only two junctions without dead-ends. We will thus get the value of the variable of the preceding equation.
5. Repeat the substitutions, and we will get finally the number of distinct courses of the primitive maze.

To give an application of this procedure, Tarry selected a recurring maze where the alleys form the edges and the diagonals of a heptagon. We recognize here that this application gives the number of possible solutions to Poinsot's problem of polygons.

Tarry used a new kind of diagram as shown in Figs. 14, 15, 16 and 17.

In these diagrams, he made use of several signs such as circles, and equilateral triangles joined to a circle.

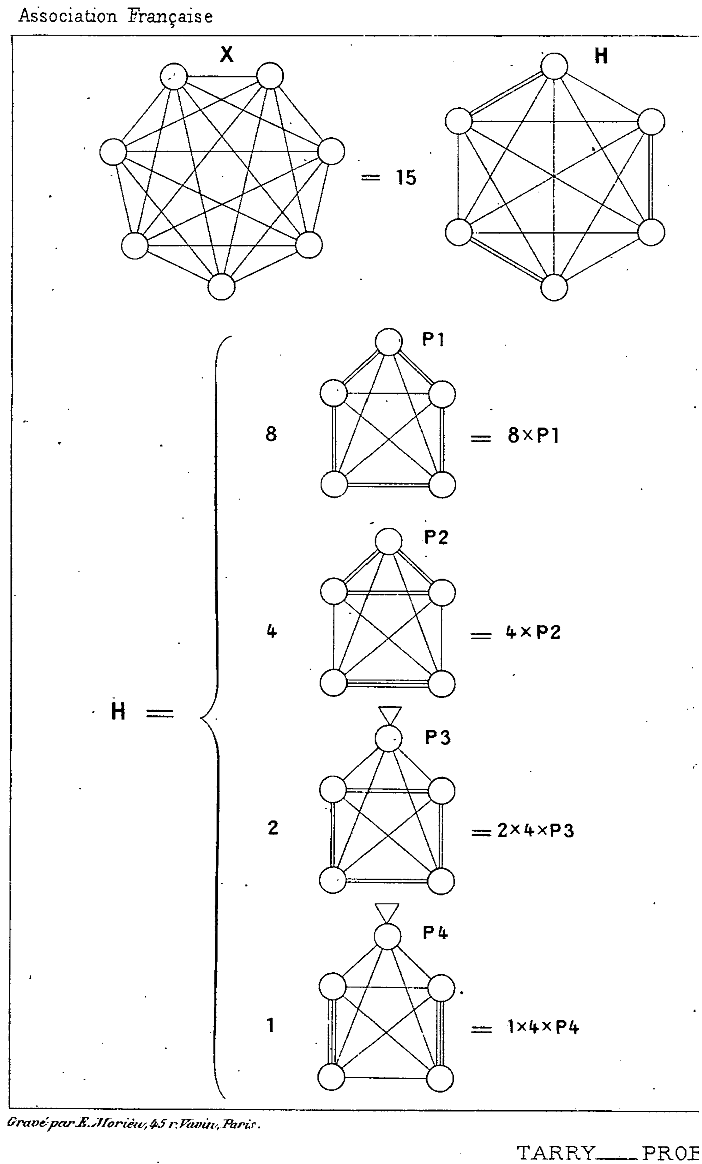


Fig. 14 Tarry's diagrams, the left half of the first sheet [48]

Tarry gave explanatory notes to read his diagrams. The sentences between "[" and "]" bellow are my comments.

Circle Junction of the maze.
[We recognize here that the elements that were to become the vertices of the graph are explicitly noted.]

Straight line connecting two circles Alley of the maze connecting two junctions.

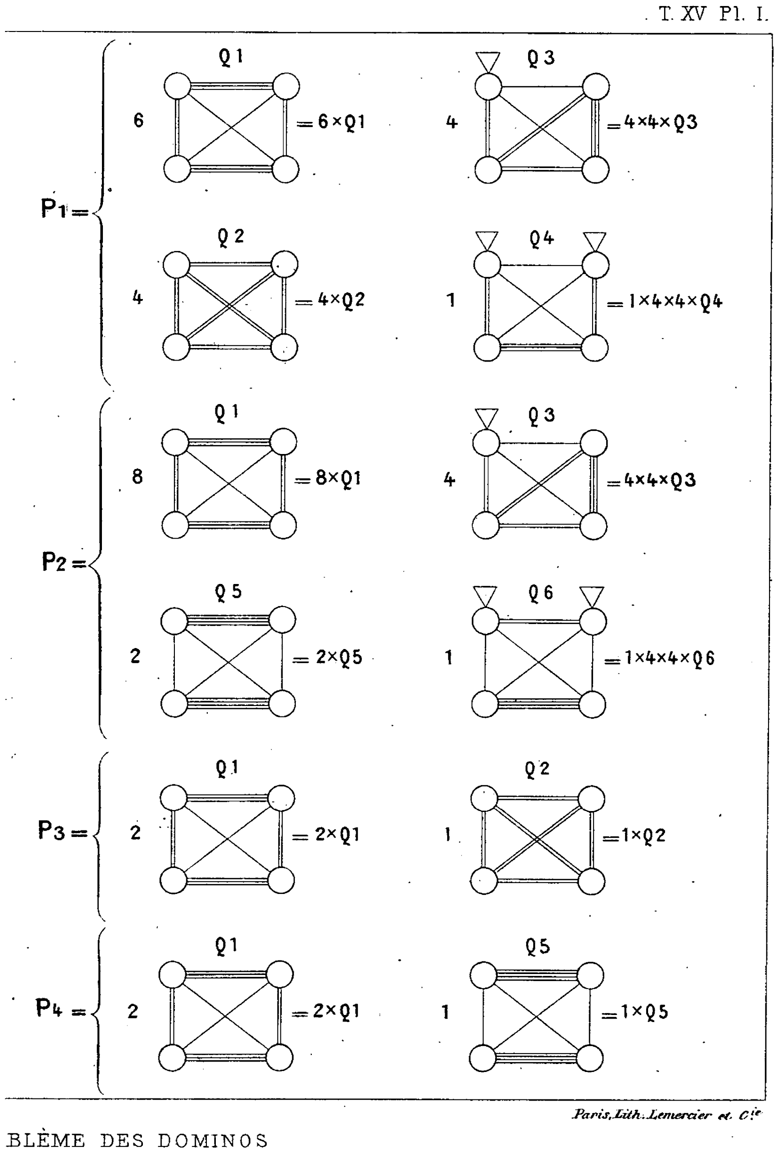


Fig. 15 Tarry's diagrams, the right half of the first sheet [48]

[A straight line represents any kind of alley and is used from the viewpoint that it is described, as was the case for the polygons above.]

Equilateral triangle having one corner on a circle Dead-end (alley, both ends of which lead to the same junction corresponding to the circle).

Letter beside each figure The letter indicates each figure, and at the same time in the equation, represents the number of courses corresponding to this figure. [See the detail bellow.]

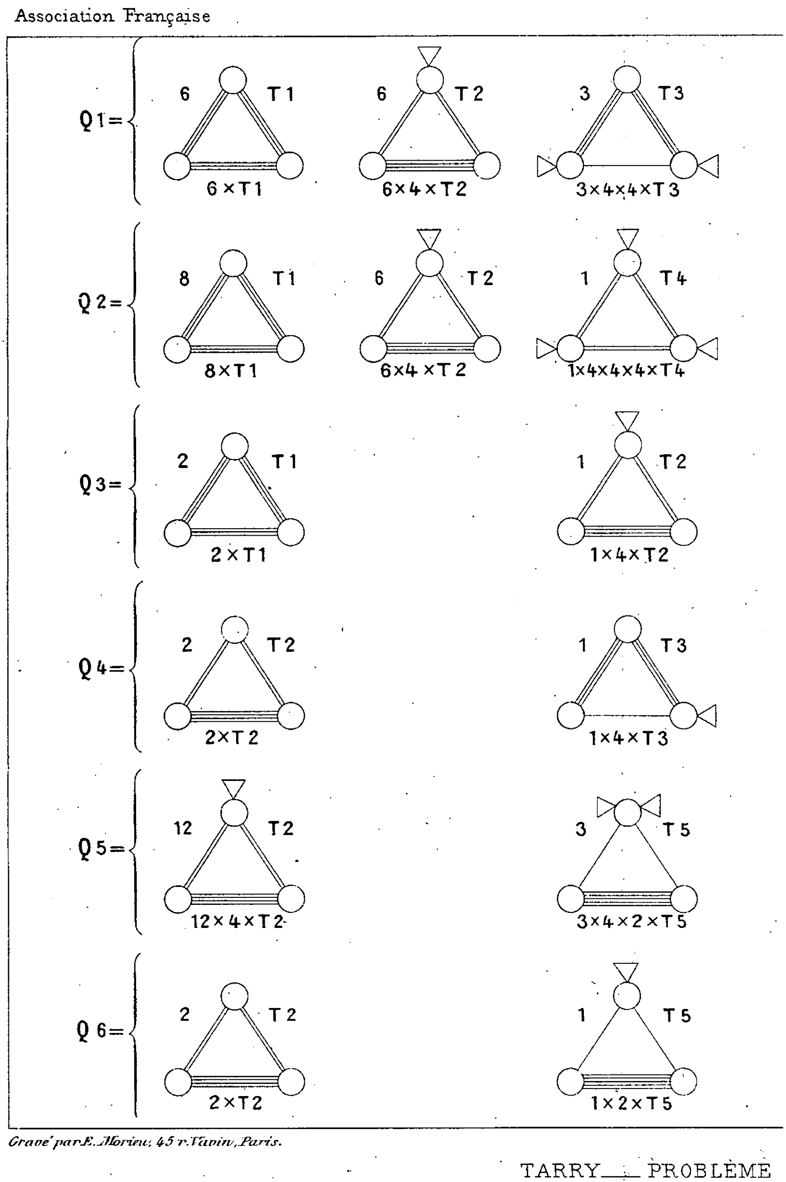


Fig. 16 Tarry's diagrams, the left half of the second sheet [48]

In the first diagrams of Fig. 14, we see the equation $X = 15H$. X represents the number of distinct courses of the heptagonal maze, which we want to obtain. H represents the number of distinct courses of the hexagonal maze, which is a reduced maze of the heptagonal maze. We obtain the number "15" from the number of alleys leading to one of the junctions of the heptagonal maze "6": when we group the six alleys leading to the junction into pairs of alleys in all the possible ways, we get $5 \times 3 \times 1$ different groups, that is, 15. Applying the theorem to reduce junctions, we get the equation $X = 15H$. When

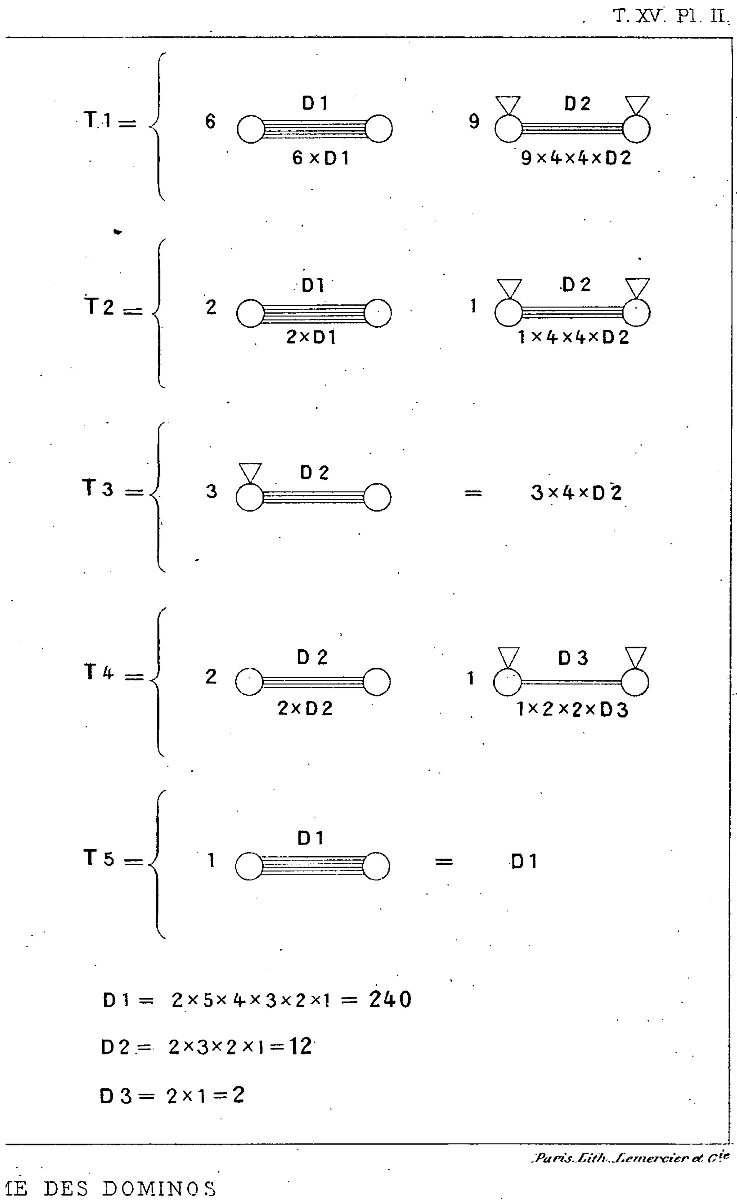


Fig. 17 Tarry's diagrams, the right half of the second sheet [48]

we delete the junction of the heptagonal maze X, the six alleys are reduced to three alleys, which form double lines in the figure of the hexagonal maze H.

In the next figures, the hexagonal maze H is reduced to the pentagonal mazes P_1, P_2, P_3, P_4. These four pentagonal mazes are drawn differently because multiple lines are differently connected depending on the ways of grouping of the six alleys leading to one of the junctions of the hexagonal maze into three pairs of alleys: for P_1, we count the groups bringing five double alleys; for P_2, we count the groups bringing three double alleys and one triple alley; for P_3, we count the groups bringing four double alleys; for P_4, we count the groups bringing two triple alleys. Moreover, each of P_3 and P_4 has a dead-end, we therefore apply to them the corollary with $k = 1, n = 2$. We thus multiply the count of groups by $(2 + 1 - 1)2^1 = 4$ to obtain the number of distinct courses. Then we get the equation $H = 8P_1 + 4P_2 + 4 \times 2P_3 + 4 \times P_4$.

Continuing the procedure similarly, we get the following equations successively:

$$X = 15H$$
$$H = 8P_1 + 4P_2 + 8P_3 + 4P_4$$

$$P_1 = 6Q_1 + 4Q_2 + 16Q_3 + 16Q_4$$
$$P_2 = 8Q_1 + 16Q_3 + 2Q_5 + 16Q_6$$
$$P_3 = 2Q_1 + Q_2$$
$$P_4 = 2Q_1 + Q_5$$

$$Q_1 = 6T_1 + 24T_2 + 48T_3$$
$$Q_2 = 8T_1 + 24T_2 + 64T_4$$
$$Q_3 = 2T_1 + 4T_2$$
$$Q_4 = 2T_2 + 4T_3$$
$$Q_5 = 48T_2 + 24T_5$$
$$Q_6 = 2T_2 + 2T_5$$

$$T_1 = 6D_1 + 144D_2$$
$$T_2 = 2D_1 + 16D_2$$
$$T_3 = 12D_2$$
$$T_4 = 2D_2 + 4D_3$$
$$T_5 = D_1$$

$$D_1 = 240$$
$$D_2 = 12$$
$$D_3 = 2$$

From these equations, we get finally the number of distinct courses of the heptagonal maze $X = 129976320$.

Tarry did not mention the problem of dominoes in his text of the proceedings, but we see on the diagram sheets the caption "TARRY—PROBLÈME DES DOMINOS (Tarry—problem of dominoes)". This caption supports the description of the abstract that Tarry applied his theorems to the problem of dominoes. Moreover, we recognize that the diagrams given for the calculation of the number of distinct courses of a heptagonal maze was, in his talk, used for the problem of dominoes. We recognize, therefore, that Tarry related the problem of polygons with the problem of dominoes.

In Tarry's diagrams, We find a clear representation of junctions of a maze, which corresponds to vertices of a graph in modern terms, while such a representation was not found in Lucas' diagram used to solve the maze problem (Fig. 13), nor in that of polygons

(Fig. 11). Tarry's representation suggests the importance of junctions, which correspond to the vertices of a graph, which is still important in Kőnig's treatise of 1936.

It is remarkable that Tarry related the problem of bridges to the concepts of mazes, though in this talk, he did not treat the same problem of mazes as we discussed in Sect. 2.6. Moreover, he applied his result to the problem of dominoes using diagrams of polygons.

4 Conclusion

We discussed how the representation of graphs by means of the diagrams of graph theory emerged.

We examined some problems Kőnig treated in his treatise of 1936, and analyzed the diagrams attached to them in the texts before 1936. Earlier, these diagrams did not always have the form of the present day, and the forms were not uniform. The features of diagrams in early days were different from each other, depending on the problem in which the diagram was used, while the features of diagrams in the present day are uniform in different topics.

Tarry's talk in 1886 [48] played an important role in the way of using the diagrams.

With respect to diagrams, the significance of Tarry's talk was that he used a uniform type of diagram in different topics, and that *for solving* the problem of dominoes in his talk, he connected this problem to other problems for which similar diagrams had been introduced and which had been reformulated as problems related to this kind of diagrams.

In fact, we identified another graphical representation and another conception of the object under study used at the beginning of the 19th century to connect a smaller set of distinct problems now understood as bearing on graphs: a form of diagram with lines and small circles which already appeared in an article of Vandermonde in 1771 [50]. However, its status was different at that time: Vandermonde used the diagram not for stating and then solving a problem, but *for representing his solution* to the problem. The remarkable contribution of Vandermonde (1771) to graph theory is that he used the concepts of "épingle (pin)" and "fil (string)" for a problem of the knight's move on a chessboard, thereby introducing in particular the line independently from its shape and distinguishing only some points to represent a situation. Moreover, he gave a diagram representing his result using these concepts. In Kőnig's treatise of 1936, the problem of the knight's move on a chessboard was considered, in the context of graph theory, to be a problem of a Hamiltonian circuit. But Kőnig's treatise of 1936 is not the first text in which the same concepts were used as a basis to define problems referring to questions related to circuits, the circuits being represented by means of the same elements. Poinsot, in his lecture of 1809 [38], applied the concepts of Vandermonde to the problem of polygons. In other words, Poinsot recognized that the same concepts can be used to formulate two apparently different problems, one on a Hamiltonian circuit and another one, which in Kőnig's treatise in 1936 was mentioned as a problem of Eulerian circuits. It is also remarkable that Poinsot used these concepts *for solving* the problem, not only representing the solution, though no diagram representing the concepts is found. It was Kőnig who, for the first time, explicitly related these two problems in the same chapter, and considered them using the same graph theoretic basis.

Tarry also related the problem of bridges to the concept of mazes. Moreover, he applied his result to the problem of dominoes using diagrams of polygons.

However, Tarry did not integrate explicitly all the four problems of bridges, polygons, dominoes and mazes, while Kőnig did in his treatise of 1936.

References

1. Ahrens, W.: Mathematische Unterhaltungen und Spiele. Teubner, Leipzig (1901)
2. Ahrens, W.: Mathematische Spiele. In: Meyer, W.F. (ed.) Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen, Band 1, pp. 1080–1093. Teubner, Leipzig (1902)
3. Ahrens, W.: Mathematische Unterhaltungen und Spiele, vol. 2. Teubner, Leipzig (1918) (Zweite, vermehrte und verbesserte Auflage)
4. Ahrens, W.: Mathematische Unterhaltungen und Spiele, vol. 1. Teubner, Leipzig (1921) (Dritte, verbesserte, anastatisch, gedruckte Auflage)
5. Bachet, C.-G.: Problèmes plaisans et délectables qui se font par les nombres. Pierre Rigaud, Lyon (1612)
6. Bachet, C.-G.: Problèmes plaisans et délectables qui se font par les nombres. Pierre Rigaud, Lyon (1624) (Seconde édition revue, corrigée et augmentée de plusieurs propositions, et de plusieurs problèmes par le même auteur)
7. Bachet, C.-G.: Problèmes plaisants et délectables qui se font par les nombres. Gauthier-Villars, Paris (1874) (Troisième édition, revue, simplifiée et augmentée par A. Labosne)
8. Bachet, C.-G.: Problèmes plaisants et délectables qui se font par les nombres. Gauthier-Villars, Paris (1879) (Quatrième édition, revue, simplifiée et augmentée par A. Labosne)
9. Ball, W.W.R.: Mathematical Recreations and Problems of Past and Present Times. Macmillan, London (1892) (First edition in February; second edition in May with no material changes)
10. Ball, W.W.R.: Mathematical Recreations and Problems of Past and Present Times, 3rd edn. Macmillan, London (1896)
11. Ball, W.W.R.: Récréations et problèmes mathématiques des temps anciens et modernes. Hermann, Paris (1898) (Troisème édition revue et augmentée par l'auteur, traduite par J. Fitz-Patrick)
12. Ball, W.W.R.: Mathematical Recreations and Essays, 4th edn. Macmillan, London (1905)
13. Ball, W.W.R.: Récréations et problèmes mathématiques des temps anciens et modernes, vol. 1. Hermann, Paris (1907) (Deuxième édition française traduite d'après la quatrième édition anglaise et enrichie de nombreuses additions par J. Fitz-Patrick)
14. Ball, W.W.R.: Récréations et problèmes mathématiques des temps anciens et modernes, vol. 2. Hermann, Paris (1908) (Deuxième édition française traduite d'après la quatrième édition anglaise et enrichie de nombreuses additions par J. Fitz-Patrick)
15. Ball, W.W.R.: Mathematical Recreations and Essays, 5th edn. Macmillan, London (1911)
16. Biggs, N.L., Lloyd, E.K., Wilson, R.J.: Graph Theory, 1736–1936. Oxford University Press, New York (1976) (First published 1976; reprinted with corrections 1997; first published in paperback 1986; reprinted with corrections 1998)
17. Coupy, É.: Solution d'un problème appartenant à la géométrie de situation, par Euler. Nouv. Ann. Math. **10**, 106–119 (1851)
18. Décaillot, A.-M.: L'arithméticien Édouard Lucas (1842–1891): théorie et instrumentation. Rev. Hist. Math. **4**(2), 191–236 (1998)
19. Décaillot, A.-M.: Édouard Lucas (1842–1891): le parcours original d'un scientifique dans la deuxième moitié du XIXe siècle. Ph.D. thesis, L'université René Descartes-Paris V (1999)
20. Décaillot, A.-M.: Géométrie des tissus. Mosaïques. Échiquiers. Mathématiques curieuses et utiles. Rev. Hist. Math. **8**(2), 145–206 (2003)
21. Epple, M.: Topology, matter, and space I: topological notions in 19th-century natural philosophy. Arch. Hist. Exact Sci. **52**, 297–392 (1998)
22. Euler, L.: Solutio problematis ad geometriam situs pertinentis. Comment. Acad. Sci. Imp. Petropol. **8**, 128–140 (1736) (Based on a talk presented to the Academy on 26 August 1735; printed in 1741; English translation in the book [16])
23. Euler, L.: Solution d'une question curieuse qui ne paroit soumise à aucune analyse. Mém. Acad. R. Sci. B.-Lett. **15**, 310–337 (1766)

24. Fourrey, É.: Récréations arithmétiques. Nony, Paris (1899) (1er edn. 1899; 2e edn. 1901; 3e edn. 1904; 4e edn. 1907; 8e edn. 1947; nouvelle edn. augmentée d'une étude de Jean-Louis Nicolas 1994/2001)
25. Fourrey, É.: Curiosités géométriques. Vuibert et Nony, Paris (1907)
26. Heeffer, A.: *Récréations mathématiques* (1624), a study on its authorship, sources and influence. Gibecière **1**, 77–167 (2006)
27. Kőnig, D.: Mathematikai mulatságok, vol. 1. Lampel Róbert, Budapest (1902) (Reprinted in 1991)
28. Kőnig, D.: Mathematikai mulatságok, vol. 2. Lampel Róbert, Budapest (1905) (Reprinted in 1992)
29. König, D.: Theorie der endlichen und unendlichen Graphen. Akad. Verlagsgesellschaft, Leipzig (1936) (Republished in 1950 by Chelsea Publishing Company in New York; republished in 1986, with "Leonhard Eulers Abhandlung über das Königsberger Brückenproblem (1736)", with preface by Paul Erdös, and with a contribution from Tibor Gallai: "Denes König—Ein biographischer Abriß" edited with an appendix by Horst Sachs, Teubner-Arch. Math., Band 6, S. 3)
30. Kraitchik, M.: La mathématique des jeux ou récréations mathématiques. Stevens Frères, Bruxelles (1930)
31. Lucas, É.: Récréations mathématiques, vol. I. Gauthier-Villars, Paris (1882) (2nd edn. 1891)
32. Lucas, É.: Récréations mathématiques, vol. II. Gauthier-Villars, Paris (1883) (2nd edn. 1896)
33. Lucas, É.: Récréations mathématiques, vol. III. Gauthier-Villars, Paris (1893)
34. Lucas, É.: Récréations mathématiques, vol. IV. Gauthier-Villars, Paris (1894)
35. Lucas, É.: L'Arithmétique amusante. Gauthier-Villars, Paris (1895)
36. Ozanam, J.: Récreations mathématiques et physiques, vols. 1–2. Jombert, Paris (1694) (Le Traité des horloges élémentaires est traduit de l'italien de Domenico Martinelli)
37. Ozanam, J.: Récreations mathématiques et physiques, vols. 1–4. Virguin, Paris (1778) (Nouvelle édition, totalement refondue & considérablement augmentée par M. de C.G.F.)
38. Poinsot, L.: Mémoire sur les polygones et les polyèdres. J. Éc. Polytech. **4**(10), 16–48 (1810) (Lu à la première classe de l'institut le 24 Juillet 1809)
39. Pont, J.-C.: La topologie algébrique des origines à Poincaré. Presses Universitaires de France, Paris (1974)
40. Rebière, A.: Mathématiques et mathématiciens: pensées et curiosités. Nony, Paris (1889) (282 pages)
41. Rebière, A.: Mathématiques et mathématiciens: pensées et curiosités, 3rd edn. Nony, Paris (1898) (566 pages)
42. Reiss, M.: Evaluation du nombre de combinaisons desquelles les 28 dés d'un jeu du domino sont susceptibles d'après la règle de ce jeu. Ann. Mat. Pura Appl. **5**, 63–120 (1871–1873)
43. Schubert, H.: Mathematische Mußestunden. Göschen'sche Verlagshandlung, Hamburg (1897) (English translation [44])
44. Schubert, H.: Mathematical Essays and Recreations. Open Court Publishing, Chicago (1898) (Translated from German by Thomas J. McCormack)
45. Schubert, H.: Mathematische Mußestunden, vols. I–III. Göschen, Leipzig (1900) (It appeared also in an abbreviated one-volume form in 1904)
46. Singmaster, D.: Walter William Rouse Ball, *Mathematical recreations and problems of past and present times*, first edition (1892). In: Grattan-Guinness, I. (ed.) Landmark Writings in Western Mathematics, 1640–1940, pp. 653–663. Elsevier, Amsterdam (2005) (Chap. 50)
47. Sylvester, J.J.: On recent discoveries in mechanical conversion of motion. In: Proceedings of the Royal Institution of Great Britain, vol. 7, pp. 179–198 (1873–1875) (Reprinted in The Collected Mathematical Papers of James Joseph Sylvester, vol. III, pp. 7–25. Cambridge, 1904–1912)
48. Tarry, G.: Géométrie de situation: nombre de manières distinctes de parcourir en une seule course toutes les allées d'un labyrinthe rentrant, en ne passant qu'une seule fois par chacune des allées. In: Association française pour l'avancement des sciences, Compte rendu de la 15e session, Nancy, 1886, vols. 1–2, pp. 49–53 de la 2e partie. Paris (1887) (Secrétariat de l'association, Georges Masson. Le résumé: p. 81 de la 1er partie; 2 feuilles de figures à la fin de la 2e partie)
49. Terquem, O.: Sur les polygones et les polyèdres étoilés, polygones funiculaires. Nouv. Ann. Math. **8**, 68–74 (1849) (D'après M. Poinsot, Sects. 1–10)
50. Vandermonde, A.-T.: Remarques sur des problèmes de situation. In: Histoire de l'académie royale des sciences. Année MDCCLXXI. Avec les Mémoires de Mathématique & de Physique, pour la même Année, Tirés des Registres de cette Académie, pp. 566–574. L'Imprimerie Royale, Paris (1771) (Printed in 1774; English translation in the book [16])

51. Wate-Mizuno, M.: The works of Kőnig Dénes (1884–1944) in the domain of mathematical recreations and his treatment of recreational problems in his works of graph theory. Ph.D. thesis, Université Denis Diderot-Paris 7 (2010)

M. Wate-Mizuno (✉)
Equipe REHSEIS, Laboratoire SPHERE, Université Paris 7, 5 rue Thomas Mann, Case 7093,
75205 Paris Cedex 13, France
e-mail: fotonoakvokampo@gmail.com