What's in a Diagram?

On the Classification of Symbols, Figures and Diagrams

Mikkel Willum Johansen

Abstract In this paper I analyze the cognitive function of symbols, figures and diagrams. The analysis shows that although all three representational forms serve to externalize mental content, they do so in radically different ways, and consequently they have qualitatively different functions in mathematical cognition. Symbols represent by convention and allow mental computations to be replaced by epistemic actions. Figures and diagrams both serve as material anchors for conceptual structures. However, figures do so by having a direct likeness to the objects they represent, whereas diagrams have a metaphorical likeness. Thus, I claim that diagrams can be seen as material anchors for conceptual mappings. This classification of diagrams is of theoretical importance as it sheds light on the functional role played by conceptual mappings in the production of new mathematical knowledge.

1 Introduction

After the formalistic ban on figures, a renewed interest in the visual representation used in mathematics has grown during the last few decades (see e.g. [11–13, 24, 27, 28, 31]). It is clear that modern mathematics relies heavily on the use of several different types of representations. Using a rough classification, modern mathematicians use: written words, symbols, figures and diagrams. But why do mathematicians use different representational forms and not only, say, symbols or written words? In this paper I will try to answer this question by analyzing the cognitive function of the different representational forms used in mathematics. Especially, I will focus on the somewhat mysterious category of diagrams and

M. W. Johansen (✉)
Københavns Universitet, Copenhagen, Denmark
e-mail: mwj@ind.ku.dk

L. Magnani (ed.), *Model-Based Reasoning in Science and Technology*,
Studies in Applied Philosophy, Epistemology and Rational Ethics 8,
DOI: 10.1007/978-3-642-37428-9_6,

explain why diagrams can be seen as material anchors for conceptual metaphors and blends. However, in order to identify what is special about diagrams, we will have to analyze the other representational forms as well. Thus, I will begin by explaining the cognitive significance of symbols and figures, and conclude with the analysis of diagrams.

2 The General Function of External Representations

In modern cognitive science there is a growing understanding of the fact that human cognition cannot be understood only by looking at the processes going on inside the human brain. As it is, our cognitive life seems to involve the external environment in several important ways.

From this point of view, external representations have several different cognitive functions. They obviously serve communicative purposes and reduce demands on internal memory. Furthermore, as noticed by Andy Clark [5], when thoughts are represented in an external media they are in non-trivial ways turned into objects. This objectification of thoughts allows us—and others—to inspect and criticize the thoughts. We can carefully scrutinize each step of a complicated argument and even form meta-thoughts about our own thinking. It seems that the use of external representations of mental content is a prerequisite for the formation of high-level cognitive processes. This function is clearly relevant in the case of mathematics, both in regard to the everyday practice of working mathematicians and in regard to the formation of meta-mathematics; in fact, meta-mathematics is exactly the kind of high-level cognitive processes that depends crucially on our ability to represent mathematical thoughts in an external media.

The last general function of external representations I will go through here is their ability to serve as material anchors for complex thoughts. As noticed by Edwin Hutchins [15], our ability to perform reasoning involving complex conceptual structures depends on our ability to represent such structures in an appropriate way. When we reason with a complex structure, we manipulate parts of the structure while the rest is kept stable. Unfortunately, our ability to do this mentally is limited. Consequently, human reasoning can be facilitated by the use of external representations that allow us to anchor some of the elements of the conceptual structure in physical representations that are globally stable, but locally manipulable. The combination of stability and manipulability of some external representations allows us to focus on the part of the structure manipulated on, while the rest of the structure is kept stable by the external media. Consequently, we might be able to work on more complex conceptual structures and perform more complex manipulations if we use an appropriate external representation. As we shall see, this anchoring property plays an important part in the function of both figures and diagrams.

It should be noted that external representations have several other important cognitive functions, see e.g. [20] for an overview.

3 Symbols and Words

3.1 Symbols as Semantic and Syntactic Objects

It is a well-known fact that mathematical symbols can be treated as both semantic and syntactic objects (see e.g. [7, 8]). As semantic objects, symbols carry mathematical content or meaning. With a few rare exceptions modern mathematical symbols carry meaning only by convention; the symbols are abstract and have no likeness with their referents. The symbol "·" for instance, does not resemble the arithmetic operation of multiplication, and the Hindu-Arabic numeral "8" does not have any likeness with the quantity eight (such as for instance the Egyptian symbol "IIIIIIII", where the token "I" is repeated eight times).

As syntactic objects, symbols are objects of formal transformations. When symbols are treated as syntactic objects, the meaning of the symbols is suspended, and the problem at hand is solved by manipulating the symbols following purely formal rules. As an example, consider how the points of intersection between a circle and a straight line are found using analytic geometry. If we are to find the intersection points between a straight line l with slope $-1/2$ going through (0;10) and a circle C with center (4;3) and radius 5, we will at first have to find the equations of the two objects ($y = -1/2x + 10$ and $(x-4)^2 + (y-3)^2 = 5^2$, respectively). Then we must substitute the y in C's equation with the expression $-1/2x + 10$, simplify and solve the resulting quadratic equation. During the solution process, we are not interested in—and we do not use—the geometric interpretation of the symbols. We are only concerned with the actual symbols on the paper before us. There is no reference to the meaning or content of the symbols, only to the symbolic forms and the transformations we make on them; we talk about "substituting" one symbolic form with another and "simplifying" other expressions. The meaning of the symbols is only restored when the solution is found and given a geometric interpretation in the form of intersection points.

Arguably, this is a simple example taken from high-school mathematics, but the same dialectic between the use of symbols as semantic and syntactic objects can be found in more advanced mathematical texts as well (for analysis of more elaborate examples see [18, p. 139]).

In this case the symbols are in cognitive terms used as *cognitive artifacts* allowing computations to be performed as *epistemic actions* (cf. [7, 18, 21]). For those not familiar with distributed cognition, this might call for some explanation. An epistemic action is an actions taken in order to get information or solve computational tasks, and not in order to reach a pragmatic goal. The use of epistemic actions is a well-known cognitive strategy. When, for instance, we solve a jigsaw puzzle, we rotate and manipulate the external, physical pieces of the puzzle in order to *see* where they fit. In other words, we solve a computational problem by performing motor actions and perceiving the results of those actions. In theory we could solve the problem mentally by make internal models of the pieces and think out a solution, but we rarely do that. The reason is simple.

As David Kirsh has put it: "Cognitive processes flow to wherever it is cheaper to perform them [20, p. 442]", and for humans it is cognitively cheaper, faster and more reliable to solve jigsaw puzzles by performing epistemic actions on the external puzzle pieces rather than thinking out a solution.[1]

In other cases, however, our physical environment does not support the externalization of a problem. For this reason we sometimes produce special artifacts—cognitive artifacts—that allow us to externalize the problem and solve it by performing epistemic actions. When treated as syntactic objects, mathematical symbols are exactly such cognitive artifacts: They allow us to substitute mental computations with epistemic actions, and that is the cognitively cheapest way of solving some mathematical problems.

3.2 Symbols as Physical Objects

As mentioned above, it is well-known that symbols can be used as semantic and syntactic objects. I would, however, like to point out that symbols can also play a third and qualitatively different role in mathematical cognition, and that is the role as plain physical objects. This use of symbols is often manifested in pen-and-paper calculations. When for instance we multiply two numbers using pen and paper, the results of the sub-calculations are carefully arranged in columns and rows, and we use the previously written results as visual cues as to where we should write the next sub-result. In other words, the physical layout of the symbols is used as a way to guide the (epistemic) actions we perform on them (see [18, p. 125] for more elaborate examples).

The use of symbols as objects is also clear in the case of matrix multiplication. Here, the usual arrangement of the elements of matrices in columns and rows is a considerable help when we have to locate the elements we are about to operate on in a particular step in the process (cf. [11, p. 242]) . Notice, that the algebraic structure of matrix multiplication is completely independent of the usual physical layout of the symbols; the product of an $m \times n$ matrix $A = [a_{ij}]$ with an $n \times p$ matrix $B = [b_{jk}]$ can simply be defined as the $m \times p$ matrix, whose *ik*-entry is the sum:

$$\sum_{j=1}^{n} a_{ij} b_{jk}$$

(see e.g. [29, p. 178]). So in theory, it would be possible to perform matrix multiplication on two unsorted lists A and B of indexed elements. In that case, it

[1] It should be noted that there is an interesting parallelism between the concept of *epistemic actions* developed in [21], and the concept of *manipulative abduction* developed by Lorenzo Magnani (e.g. [25, 26]). Magnani's concept is however developed in a slightly different theoretical framework, and it would take us too far astray to explore the parallelism further.

would however pose a considerable task to find the right elements to operate on. By arranging the elements of the matrices in columns and rows in the usual way, the cost of this task is markedly reduced: The sum given above is simply the dot product of the ith row of A and the kth column of B, and if you know that, it is easy to find the elements you need. Thus, the multiplication process is clearly guided by the physical layout of the matrices.

It has also been suggested that the actual physical, or rather: *typographical* layout of symbolic representation of mathematical content has in some cases inspired new theorems and theoretical developments. Leibniz' derivation of the general product formula for differentiation is a case in point. Using the standard symbolism, the formula can be stated as:

$$d^n\overline{(xy)} = d^n x d^0 y + \frac{n}{1} d^{n-1} x d^1 y + \frac{n(n-1)}{1 \cdot 2} d^{n-2} x d^2 y \ etc. \tag{1}$$

It has been suggested that Leibniz' derived the formula by making a few, inspired substitutions in Newton's binominal formula:

$$^n\overline{(x+y)} = x^n y^0 + \frac{n}{1} x^{n-1} y^1 + \frac{n(n-1)}{1 \cdot 2} x^{n-2} y^2 \ etc. \tag{2}$$

(see [18, 24, p. 155] for further elaboration and more examples).

3.3 Words as Abstract Symbols

Finally, we might compare the use of abstract mathematical symbols with the use of written words. It should be noted that words are also abstract symbols (at least in alphabetic systems). In general, the physical appearance of a word has no likeness with the object, the word is supposed to represent; The word-picture "point", say, does not look like a point, and the word-picture "eight" does not have any more likeness with eight units than the abstract number-symbol "8".

Furthermore, words can carry mathematical content just as well as mathematical symbols. Of course, in general symbols allow a much shorter and more compact representation of a given content, but that is, in my view, only a superficial difference between the two representational forms. The important difference between written words and symbols is the fact that mathematical symbols, besides their role as bearers of content, can also be treated as syntactic and physical objects. With a few rare exceptions (such as avant-garde poetry), written words are never used as more than semantic objects; they cannot be used for purely syntactic transformations or as purely physical objects. For this reason, there are qualitative differences between written words and written mathematical symbols. We can simply do more with symbols than we can do with written words (cf. [18, p. 136]) .

4 Figures

4.1 Figures as Anchors of Conceptual Structures

Let me start with an example from classical geometry. In Heath [14, p. 197] Euclid's proof of Pythagoras' theorem (Euclid I.47) is accompanied by a figure similar to Fig. 1. The figure represents a particular construction that is used in the proof. Interestingly, the figure is not the only representation of the construction. During the course of the proof we are also give a full verbal description of the construction (see caption of Fig. 1). So why do we need the figure?

Most external representations can be said to anchor conceptual content, but the thing is that figures anchor content in a qualitatively different way than rhetoric and symbolic representations. Figures are holistic objects that present themselves as immediately meaningful to us. Thus, figures do not only provide a material anchor for the conceptual structure at hand; they provide an anchor that grounds our understanding of the conceptual structure in every-day sensory-motor experience of the physical world. In order to understand and give meaning to the content of the proof of Euclid I.47 we simply need a figure such as Fig. 1. We could of course imagine the figure or construct it in our mind's eye, but by doing so the limits of our short-term memory would pose limits to the complexity of the proofs we were able to understand. To use the example at hand, most people would, I believe, find it hard to keep track of the twelve individual points involved in the construction used in Euclid I.47, if they were only given a verbal description. By drawing a figure using a (semi)-stable medium we reduce the demands on short-term memory and are thus able to increase the complexity of the conceptual

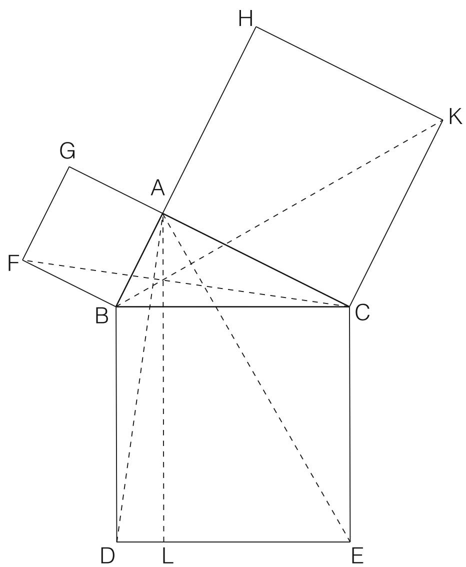


Fig. 1 Figure form the proof of Pythagoras' theorem (Euclid I.47) (redrawn from [14, p. 197]). In the proof, the following verbal description of the construction is given: "Let *ABC* be a right-angled triangle having the angle *BAC* right. [... L]et there be described on *BC* the square *BDEC*, and on *BA*, *AC* the squares *GB*, *HC*; through *A* let *AL* be drawn parallel to either *BD* or *CE*, and let *AD*, *FC* be joined. From this construction the proof proceeds ([14, p. 197])"

structures we are able to handle. That is one of the reasons why we use figures as a material anchor for complex conceptual structure, such as the structure involved in Euclid I.47.

4.2 Knowledge Deduced from Figures

Apart from making it easier for us to grasp a given mathematical construction, figures can also in some cases be used to deduce information about the mathematical objects represented in the figure. As a first example, we can return to the circle *C* and line *l* discussed in Sect. 3 above. There, I gave an outline of how the intersection points between *C* and *l* could be determined by analytic means. Another—and perhaps more direct—strategy would be to draw the two objects and simply read the intersection points off from the Figure (see Fig. 2). As we do not need to use information about the intersection points in order to construct such a figure, the figure clearly allows us to deduce genuinely new information about the objects it represents.

As another and slightly different example, we can look at the very first proof of *The Elements*. Here, Euclid shows how to construct an equilateral triangle on a given line segment *AB* (see Fig. 3). In order to achieve this result Euclid constructs two circles, one with center *A* and radius *AB* and another with center *B* and radius *AB*. Euclid then proves that the wanted equilateral triangle can be constructed by using the intersection point *C* of the two circles as the third vertices in the triangle. From a philosophical point of view, the interesting thing about this proof is the fact that Euclid does not *prove* that the two circles have an intersection point (we now know that the existence of this intersection point cannot be proven from Euclid's axioms). However, on the figure we can clearly *see* that the two circles have an intersection point, as any circles constructed in this way must have. So once more, we can deduce more information from the figure than we used in its construction, and apparently this was precisely what Euclid did. So here, information deduced from a figure plays an indispensable part of a mathematical proof.

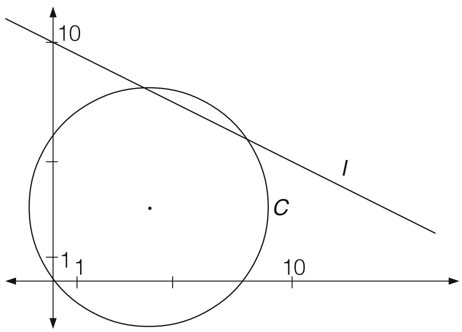


Fig. 2 Figure representing intersecting line and circle

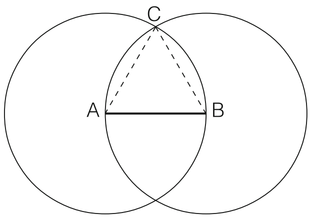


Fig. 3 Euclid I.1: How to construct an equilateral triangle on a given line segment *AB*

The use of figures in mathematical reasoning is hotly debated, and there are several things to discuss in connection to the three examples given above. At first, it should be noted that the use of figures as an aid to grasp mathematical content (such as the figure accompanying Euclid I.47) is largely recognized and supported, even by formalistically inclined mathematicians. Moritz Pasch for instance readily admits that figures "can substantially facilitate the grasp of the relations stated in a theorem and the constructions used in a proof [30, p. 43], my translation." So, even by the standards of Pasch it is legitimate to use a figure to anchor the conceptual content of a mathematical construction.

The controversy only begins when we move to Figs. 2 and 3 above. Here, the figures are used not only to illustrate, but also to infer new mathematical knowledge. The question is whether we can trust this knowledge. What is the epistemic status of knowledge deduced from a figure?

Let us begin by discussing the quality of the knowledge deduced from Fig. 2. If we compare the analytic and the pictorial method of determining intersection points, it is clear that the information deduced from the figure is not as precise as the information obtained by analytic means. Consequently, this use of figures is mainly considered a heuristic tool, and any information deduced from a figure should be tested by more reliable (i.e. analytic) means. So for instance, in the Danish high-school system a figure such as Fig. 2 is considered a valid method of finding intersection points between two curves, but the solutions read off from the figure should be tested (by means of the analytic expressions of the curves in question) if the solution is to count as a satisfying answer to the problem.

If we move to Fig. 3, the negative evaluation of the knowledge deduced from the figure is even stronger. Here, the knowledge deduced from the figure is used as an essential step in a mathematical proof, and that is—by several parties—considered an illegitimate use of pictorial knowledge. Pasch for instance continues the quote given above by stating that:

> If you are not afraid to spend some time and effort, you can always omit the figure in the proof of any theorem, indeed, a theorem is only really proved if the proof is completely independent of the figure. [...] A theorem cannot be justified by figure considerations, only by a proof; any inference that appears in the proof must have its counterpart in the figure, but the theorem can only be justified by reference to a specific previously shown theorem (or definition), and not by reference to the figure ([30, p. 43], my translation).

So Pasch would not accept Euclid's proof of Euclid I.1 as legitimate because a vital step in the proof depends on knowledge deduced from a figure. Pasch was not alone in this assessment of figures. It is well-known that David Hilbert shared

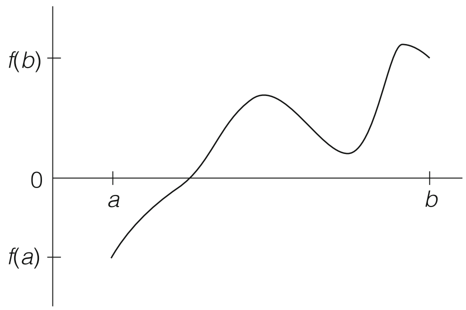


Fig. 4 Picture proof of the intermediate zero theorem. The theorem states that if a continuous function $f(x)$ defined on the interval $[a;b]$ takes both positive and negative values on the interval, then there exists a $c \in [a;b]$ such that $f(c) = 0$

Pasch's viewpoint on this matter (see e.g. [27]), and in the formalist movement a proof is in general considered "a syntactic object consisting only of sentences arranged in a finite and inspectable array [34, p. 304]". It goes without saying that figures do not have a place in such an array.

In recent years this negative evaluation of knowledge deduced from figures has been challenged by amongst others by Marcus Giaquinto [11, 12], Brown [4] and Davis [6]. Thus, Giaquinto accepts Euclid's use of a figure in the proof of Euclid I.1 on the ground that: 1) the inferences drawn from the figure does not depend on exact properties of the figure and 2) the subject matter of the proof is a homogenous class of mathematical objects (circles) that have a close relationship to a perceptual concept (perceptual circles, as the ones seen in the figure). Brown and Davis are both even more liberal in their use of figures. Brown [4, p. 25] considers Fig. 4 an adequate proof of the intermediate zero theorem (something explicitly rejected by Giaquinto), and Davis considers Fig. 5 to constitute a valid proof of the theorem that you cannot cover a circle with a finite number of smaller, non-overlapping circles [6, p. 338].

4.3 Figures and Objects

There are a number of well-known problems connected to the use of figures in mathematical deductions. As noted above, figures might not have the necessary precision, and consequently proofs based on figures can be misleading (as Rouse

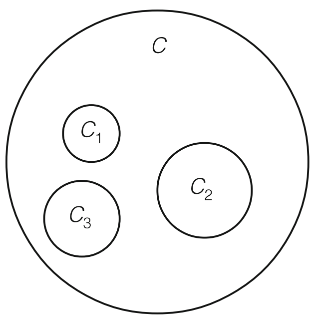


Fig. 5 Picture proof that you cannot cover a circle with a finite number of smaller, non-overlapping circles (redrawn from [6, p. 338])

Ball's famous proof that all triangles are isosceles potently illustrates [2, p. 38]). Furthermore, figures are in some cases over-specified (i.e. you cannot draw a general triangle, only a specific one) and in others they lack generality (see [11, p. 137] for more). However, none of these problems can, in my view, justify a complete ban on the use of knowledge deduced from figures. They merely impose limitations that should be observed (see also [3]).

To my mind, the main problem concerning the use of figures is connected to another and cognitively more interesting question: Why can we apparently use figures to deduce knowledge about mathematical objects *at all*?

The simple and straightforward answer is that figures somehow resembles the mathematical objects they represent; the circles drawn in Fig. 3 simply have a likeness with mathematical circles. Unfortunately, this intuitive idea is faced with several problems. Firstly, it seem to presuppose the existence of mathematical objects, or in other words Platonism, and secondly even if this presupposition is granted, it is not clear what it would mean for a physical drawing to resemble a platonic object. As a way to avoid these problems, I suggest that we see things slightly differently. Some mathematical entities such as circles and triangles are not pre-existing objects, but rather concepts created by us. They are not created at random, but are rather abstractions from and idealizations of classes of perceptual objects and shapes (see [18, p. 163]. See also [12] for a cognitively realistic account of how such an abstraction process might in fact be carried out).

Seen in this light, the longer and more correct way of explaining the relationship between a figure and the mathematical objects it represents, is the following: The figure has a direct likeness with the members of the general class of perceptual objects that provide the abstraction basis for the mathematical objects, the figure represents. Or better still: One could say that some mathematical objects such as circles and triangles are attempts to model certain aspects of physical reality, and that the shapes we see on the figures above have a direct likeness with the physical objects, the mathematical concepts are used to model. Of course not all mathematical objects have such a direct connection to sense perception, but then again: not all mathematical objects are naturally represented using figures.

If we see the relationship between perceptual figures and mathematical objects as suggested above, the real epistemological problem connected to the use of figures becomes clear. If I use information deduced from a perceptual figure in order to prove a theorem, I have proved the theorem for the wrong kind of objects. I have proven that the theorem holds good for perceptual objects, but not for the corresponding mathematical ones. Although the mathematical objects are supposed to model the perceptual objects, they might not do so perfectly; there might be a mistake in the model. In the cases above, we might accidentally have defined mathematical circles and functions in ways that would allow them to have holes where the corresponding perceptual figures have intersection points—a function such as $f(x) = x^2 - 2$ defined only on the rational numbers is a very potent example of such an object (as also pointed out by [12]). The function changes sign on the interval $[0;2]$, but there is no rational number $c \in [0;2]$ such that $f(c) = 0$. So the intermediate zero theorem does not hold good for this particular function. So it seems that we should be careful when we draw conclusions about mathematical objects on the basis of a perceptual figure. At least, we should make sure that the mathematical objects model the relevant properties of the figure in the right way.

This observation on the other hand does not show purely deductive proof to be epistemologically primary to proofs relying on figures. In my view, a figure can provide ample proof that a theorem holds good for a class of perceptual figures, and consequently the theorem *ought* to hold good for the mathematical objects modeling the class of figures as well. One can to a certain extend see the rigorization and axiomatization of mathematics during the 19th century as an attempt to make this come true. Thus, Hilbert's and Pasch's work on geometry was not an attempt to overthrow Euclid, but rather an attempt to make explicit all of the axioms needed in order to give rigorous proofs of all of the theorems of Euclidian geometry. Furthermore, as pointed out by Brown [4, p. 25] , something similar can be said about the proof of the intermediate zero theorem. As it is, the rigorous, formal proof accepted today presupposes the completeness of the real number system (amongst other things). However, this property of the real number system was not simply discovered or chosen at random. When the real number system was rigorously constructed during the 19th century, it was given the property of completeness exactly in order to make it possible to give deductive proofs of the intermediate zero theorem and other theorems presupposing the 'gaplessness' of real numbered function. Thus, proofs about perceptual figures do not only serve as a heuristic tool helping mathematicians to identify theorems they subsequently can give deductive proofs. Proofs about figures can serve—or at least have historically served as—a way to point out some of the properties we want our mathematical objects and deductive systems to have.

5 Diagrams

Before we begin, it should be realized that the word 'diagram' is used to describe a multitude of different external representations. This class of representations is not homogenous, in part because the terms 'diagram' and 'figure' are often treated as synonymous. As I see it, there are qualitative differences between figures diagrams, and part of my goal with the following analysis is to explain these difference in order to introduce a classification of mathematical representations that is more in line with their cognitive function.

It should also be noted that a lot of the work previously done on the use of diagrams in mathematics has focused on the logically soundness of diagram based reasoning. This line of work goes back at least to C. S. Peirce (1839–1914), and was revived in the mid 1990's, in part by researchers connected to the development of artificial intelligence (see e.g. [1, 13, 32]. See also [33] for a historical overview). The main goal of this program is to create a *diagrammatic calculus*, that is: a diagram based system of representations and formal transformation rules that allows for logically valid reasoning. Judged by its own standards, this program has been a great success. Several logically valid diagram based reasoning systems have been produced and some even implemented in computer based reasoning systems. The success however, has come at a price. Most mathematicians use diagrams as a heuristic tool, but instead of describing and understanding how diagrams fulfill this role, the program has focused on creating a new and different role for diagrams by turning them into a tool for logically valid, formal reasoning. To use the terminology of Sect. 3, the program wants to use diagrams as *syntactic objects*, similar to the way symbols are (in part) used. Although valuable in itself, this largely leaves the heuristic power of diagrams unexplained. In my view, there are qualitative differences between diagrams and symbols, just as there are between diagrams and figures, and if we want to understand the role diagrams play in human reasoning, we should acknowledge these differences and not shape diagrammatic reasoning into the paradigm case of logically valid, formal deduction.

Thus prepared we should have an example. Fig. 6 is a Venn diagram. Such diagrams are typically used to represent sets, and in this case three sets, *A*, *B* and *C*, are represented. Interestingly, each set is represented as a circle, or rather: a bounded area, and such areas do not have any apparent likeness with sets. Mathematical sets are abstract objects, and as such they do not have any inherent spatial characteristics. So what is the precise relationship between a diagram and the objects it represents? In the case of figures we solved the similar problem by pointing out that a figure has a direct likeness, not with the mathematical object it represents, but with the class of objects that forms the abstraction basis for the mathematical object. Unfortunately, this strategy does not seem to be viable in the case of diagrams. If we look to the objects that form the abstraction basis for, say, the set of real numbers or the set of continuous functions, it is not clear that they are arranged in anything like a bounded area. Even in more relaxed and non-technical examples (e.g. 'the sets of nouns' or 'the set of chairs') is it in general not possible to see a likeness with bounded areas. So it seems that diagrams such as Fig. 6 represents mathematical objects in a qualitatively different way than figures. The question is: How? Why do we see Venn diagrams as a representations of sets at all?

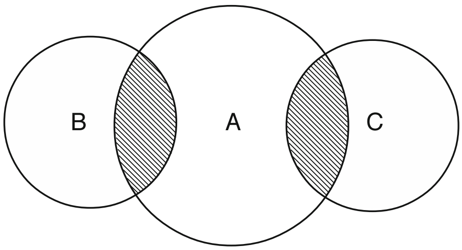


Fig. 6 A typical Venn diagram (redrawn from [16, p. 4])

In the diagrammatic calculus-program mentioned above it has been suggested that there is (or rather: ought to be) a homomorphism between a diagram and the objects it represents (e.g. [3]). The main function of this description is to explore the validity of diagrammatic reasoning, not to explore its cognitive nature. Consequently, the description does not explain why we effortlessly see certain diagrams as representations of certain mathematical objects, nor does it explain the heuristic power of diagrams. If we want to understand how diagrams represent objects, we must in my view explore the cognitive function of diagrams.

As I see it, we combine two different cognitive strategies when we use diagrams. Diagrams are external representations, and as such they can be seen as an expression of the general cognitive strategy of externalization of mental content, described above. All of the representational forms considered in this paper are expressions of this strategy. In other words, it is the second strategy that sets diagrams apart, and that strategy is *conceptual mapping*.

Conceptual mapping is a general cognitive mechanism where either one conceptual domain is mapped onto another, or a third, fictive domain is created by integrating two different conceptual domains. The first mechanism is usually referred to as 'conceptual metaphor' and the second as 'conceptual blending'. Both mechanisms are well described in the literature (see e.g. [10, 22]), and their cognitive function in mathematics has been discussed (e.g. [18, 23]), although focus has mainly been on linguistic expressions of such mappings. The main focus and contribution here will be to expand the analysis to cover non-linguistic expressions of conceptual mappings. I will do this by analyzing the role played by conceptual metaphors in our use of diagrams.

5.1 Elements of the Container Metaphor

In modern mathematics a multitude of different conceptual metaphors are in use. One of these is the Sets Are Containers-metaphor, where sets are conceptualized as containers. It is not hard to find linguistic expressions of the metaphor; most introductions to set theory will one way or the other conceptualize sets as containers. As an example, we can look at the textbook *Basic Algebra* by Nathan Jacobson. Here, sets are introduced as arbitrary collections of elements, and the basic properties of and relations between sets are described in the following way:

> If A and $B \in \mathcal{P}(S)$ (that is, A and B are subsets of S) we say that A is *contained in* B or is a *subset of* B (or that B *contains* A) and denote this as $A \subset B$ (or $B \supset A$) if every element a in A is also an element in B. [...] If A and B are subsets of S, the subset of S of elements c such that $c \in A$ and $c \in B$ is called the *intersection* of A and B. We denote this subset as $A \cap B$. If there are no elements of S contained in both A and B, that is, $A \cap B = \emptyset$, then A and B are said to be *disjoint* (or *non-overlapping*) ([16, pp. 3–4]. All emphasis form the original).

Here, sets are clearly described as containers: A set B can *contain* another set A, and both sets can *contain* elements *etc*. However, as sets are arbitrary collections of objects, they cannot literally contain anything. Thus, the description of sets as containers must be metaphorical. To be more precise, it is an example of a common conceptual metaphor, where properties from one domain—containers—are mapped onto another domain—sets: in more detail, sets are understood as containers, subsets as containers located inside a container, and the elements of a set as objects contained in a container (see [18, p. 174] for further details).

From a cognitive point of view, the main function of this conceptual mapping is to allow us to ground our understanding of sets in our experience of containers. As it is, most humans have constant experiences with containers. We use containers such as bottles and boxes an a daily basis, we move containers around and contain them in other containers (as when we put a bottle in a bag or a Tupperware in the refrigerator). We are ourselves contained in clothes and buildings and contain the food and liquid we consume. So in sum, we know a lot about containers from direct experience (cf. [19, p. 21]).

Mathematical sets on the other hand are abstract objects. We cannot experience sets directly in any way, but the conceptualizing of sets as containers allows us to map our experiences with containers onto the domain of sets. This move gives us an intuitive grasp of sets and—more importantly—allows us to recruit our knowledge about containers when we reason about sets. Conceptual metaphors are inference preserving, so if we know something to be true about containers from direct experience, we can simply activate the metaphor and map the conclusion onto the domain of sets. When sets are understood as containers, it is for instance easy to see that if a set A is enclosed in another set B, all the elements of A will also be elements of B, because we know this to be true of the content of a container A enclosed in another container B.

5.2 Diagrams as Material Anchors for Conceptual Mappings

It is now time to return to Fig. 6, were three sets were represented as bounded areas. As noted above, bounded areas do not have any direct likeness with sets. Consequently, the circles making up the bounded areas could be seen as abstract or conventional representations, similar to abstract symbols. There is however something more at play in the diagram. Circles, or rather: bounded areas constitute a special type of containers. Unlike some containers, bounded areas can overlap, but otherwise they are encompassed by the same basic logic as containers in general. Now, if we use the SETS ARE CONTAINERS-metaphor to conceptualize mathematical sets as containers, we can see that the circles making up the bounded areas of Fig. 6 are more than arbitrary representations. The circles might not have a direct likeness with mathematical sets, but they do have a direct likeness with containers, and when we conceptualize sets as containers, the diagram gets an indirect or *metaphorical* likeness with mathematical sets as well.

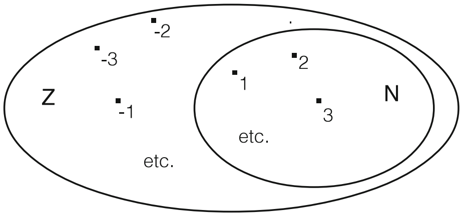
In order to understand and use the diagram we must in other words conceptualized the mathematical objects, the diagram represents, using a particular conceptual mapping. This, in my view, marks the principal qualitative difference between diagrams and the other representational forms discussed above: In contrast to symbols, diagrams are not abstract representations, and in contrast to figures, diagrams only have an indirect likeness with the objects, they represent.

Furthermore, we should be aware that Fig. 6 is not only a representation of three sets. It is in fact a diagrammatic proof of the distributive law for set operations $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (cf. [16, p. 4]). It can easily be seen that the shaded areas of the diagram corresponds to both $A \cap (B \cup C)$ and $(A \cap B) \cup (A \cap C)$. Thus, the identity holds good for bounded areas of the plane, and by using the Sets Are Containers-metaphor, we can map this inference onto the domain of sets to get the corresponding set-theoretical identity.

The use of a diagram to support this kind of picture proof draws attention to the double nature of diagrams. Diagrams are an expression of our use of conceptual maps, but they are also external representations, and as such shares some of the properties of figures and symbols. In particular, diagrams are, similar to figures, able to function as material anchors for conceptual structures. In a diagram, a conceptual structure established via a metaphor or conceptual blend is mapped onto an external, physical structure, whose individual elements serve as proxies for elements of the conceptual structure. Furthermore, the physical structure is globally stable, but locally manipulable. When a diagram is drawn, we can for instance add new bounded areas to it or, as it is the case in Fig. 6, shade certain areas without altering the overall structure of the diagram. By manipulating and inspecting a diagram we can, due to the underlying conceptual mapping, draw inferences about the objects represented by the diagram (or rather: about the objects as they are conceptualized under the given metaphor or blend). In the case above, the distributive law was for instance verified simply by drawing and inspecting Fig. 6. As in the case of figures, this anchoring property allows us to increase the complexity of the conceptual structures we are able to work on. So to give the full cognitive characterization, diagrams are material anchors for conceptual mappings. In this way, diagrams are—from a cognitive point of view—highly complicated cognitive tools, and they are clearly qualitatively different from both figures and symbols.

As noted above, inferences based on the inspection of a figure can for several reasons lead to false conclusions. When we turn to inferences based on diagrams, we must add a new entry to the list of possible errors, and that is: inadequacy of the metaphor. When we use a diagram to draw inferences, we reason about the mathematical objects taken under a particular metaphorical conception, but unfortunately metaphors can be misleading (in fact, some authors claim they always are (see e.g. [17]). In the case of Venn diagrams, the underlying container metaphor is clearly inadequate in several respects. Firstly, the metaphor does not capture the mathematically important difference between a set A being a subset of another set B and A being an element of B. Secondly, the metaphor might lead to false conclusions about the relative size of infinitely large sets. If for instance the natural numbers and the integers are represented in a Venn style diagram it would seem from the diagram that the set of integers is larger than the set of natural

Fig. 7 The natural numbers and the integers



numbers (see Fig. 7). However, from a mathematical point of view, it isn't. This observation is of course extremely important to keep in mind when one is reasoning with diagrams.

5.3 Commutative Diagrams

So far, the analysis has only been based on one type of diagrams: Venn diagrams. Such diagrams are well-suited as examples because they are easy to understand and can be used to represent mathematically basic results. On the other hand, working mathematicians rarely use Venn diagrams, so in order to show the generality of the analysis, we should also cover a mathematically more realistic example. For this reason I will make a brief analysis of one of the diagrammatic forms most commonly used in modern mathematics: the commutative diagram (see also for more examples [18]).

In short, commutative diagrams are diagrammatic representations of maps between sets (where the exact type of maps and sets are typically specified further). Following Jacobson once more, from a strictly formal point of view a *map* consists of three sets: a domain S, a codomain T and a set α of ordered pairs (s, t), with $s \in S$, $t \in T$ and such that:

1. for any $s \in S$ there exists a pair $(s, t) \in \alpha$ for some $t \in T$.
2. if $(s, t) \in \alpha$ and $(s, q) \in \alpha$ then $t = q$ [16, p. 5].

Less formally stated, a map is a set of relations between the elements of two sets.

In modern mathematics maps are commonly represented as an arrow going from one symbol representing the domain to another symbol representing the codomain. Thus, the map f with domain A and codomain B can for instance be represented as: $A \xrightarrow{f} B$. This representation draws on several conceptual metaphors. In order to understand the diagram, the two sets must be conceptualized as locations in space, and the map must be understood as a movement along a directed path from one location to the other. As sets are not literally locations in space and maps are not literally movements, these conceptualizations are clearly metaphorical. Now, an arrow does not have a direct likeness with directed movement in space. An arrow is a (more or less) arbitrary symbol used by convention to signify movements in space. So in contrast to Fig. 6, the diagram above contains both conventional symbols and figural elements that have a direct likeness to the corresponding elements of the metaphor. Consequently, it is what I will call a *mixed anchor*.

The representation of maps as arrows between sets is especially useful when several maps between several sets are involved. So, three maps between three sets can be represented by the triangle displayed in Fig. 8.

Clearly, Fig. 8 is also a mixed anchor containing both conventional symbols and figural elements. Notice also that the symbols A, B and C are used not only as semantic objects designating the three sets, but also as purely physical objects marking the metaphorical location of the sets in the diagram.

An interesting aspect of this kind of diagrams is the fact that the conceptual metaphor embodied in the diagram is inadequate in an important respect. In the diagram, a map is represented as a movement between two sets, but from a formal point of view, a map is not a relation between two sets, but between the elements of two sets. This inadequacy has important consequences. According to Fig. 8, I can get from A to C in two ways: I can either go by the f—and then by the h-arrow, or I can go directly by taking the g-arrow. From this, it seems that the composition $h \circ f$ of f and h is equal to g. However, this might not hold good. From a mathematical point of view the composition of f and h is only equal to g if the composition, for any elements in A, will take me to the same element in C as g (i.e. $\forall x \in A : (h \circ f)(x) = g(x)$). If this is the case, the diagram is said to commute.

In order to complete the cognitive analysis, we should also notice that commutative diagrams are locally manipulable, but globally stable. We can easily add new locations (i.e. sets) or new arrows (i.e. maps) to the diagram without disturbing the overall stability of the representation.

Furthermore, commutative diagrams can be used to infer new knowledge about the objects, they represent. If we know, say, that both square $ABDE$ and square $BCEF$ in Fig. 9 commutes (i.e. $j \circ f = m \circ h$ and $k \circ g = n \circ j$), then from the diagram we can easily infer that the whole square commutes as well (intuitively, that you can go form A to F by any route you want).

Finally, commutative diagrams can also be used to anchor very complicated conceptual structures. A good example is the so-called five lemma. I will not go into the mathematical details, as they are inconsequential for our purposes (see e.g. [9]). The lemma states that for a certain types of sets (such as abelian groups), the map n is of a particular type, if a number of conditions are met (the rows must be exact and the maps k, m, p and q must be of particular types). The content of this lemma is often illustrated with a commutative diagram similar to Fig. 10.

Fig. 8 Representation of three maps between three sets

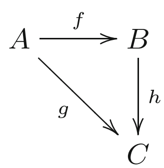


Fig. 9 Two commutative squares

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C \\ \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle j} & & \downarrow{\scriptstyle k} \\ D & \xrightarrow{m} & E & \xrightarrow{n} & F \end{array}$$

Fig. 10 The five lemma

$$\begin{array}{ccccccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C & \xrightarrow{h} & D & \xrightarrow{j} & E \\ \downarrow{\scriptstyle k} & & \downarrow{\scriptstyle m} & & \downarrow{\scriptstyle n} & & \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle q} \\ A' & \xrightarrow{r} & B' & \xrightarrow{s} & C' & \xrightarrow{t} & D' & \xrightarrow{u} & E' \end{array}$$

Notice, that in this case no less than 23 different objects (10 sets and 13 maps) are involved. Although the usual conceptual metaphor (where sets are understood as locations in space and maps as paths between them) allows us to get a more intuitive grasp of the situation, it still poses considerable demands on short-term memory to keep track of all of the elements involved in the lemma. The conceptual structure is so complicated that we simply need an external anchor in the form of a physical diagram in order to stabilize it. In this sense the diagram above is similar to the figure accompanying Euclid I.47 (Fig. 1). They both illustrate how the use of external anchors allows us to increase the complexity of the conceptual structures we are able to work with. Only, Fig. 10 has an indirect or metaphorical likeness with the objects it represents, whereas Fig. 1 has a direct likeness.

6 Conclusion

In this paper I have categorized the different representational forms used in mathematics from a cognitive point of view. The analysis suggests a more careful use of language, especially in connection to the words ‘figure’ and ‘diagram’. However, the main aim of the categorization is not to police the use of language, but rather to draw attention to the fact that mathematical reasoning depends on a multitude of qualitatively different representational forms. From a cognitive point of view there are qualitative differences between written words, symbols, figures and diagrams. The different representational forms involve different cognitive processes and they play different roles in the reasoning process.

This appeal to recognize the differences between the various representational forms is also an appeal to recognize the complexity and diversity of mathematical reasoning. This complexity has not always been acknowledged. The formalistic movement clearly failed to recognize the complexity of mathematical reasoning by identifying mathematics with the use of a specific cognitive tool: symbols.

Although the formalistic movement has been challenged in recent decades, the exact quality of and difference between the various representational forms used in mathematics is still not well understood. From this paper it should be clear that there are several important differences between written words, symbols, figures and diagrams, and the main attraction of the cognitive perspective applied here is exactly the fact that it makes it possible for us to see and understand these differences.

References

1. Allwein, G., Barwise, J. (eds.): Logical Reasoning with Diagrams. Oxford University Press, Oxford (1996)
2. Ball, W.W.R.: Mathematical Recreations and Essays, 4th edn. Maxmillan and Co, London (1905)
3. Barwise, J., Etchemendy, J.: Heterogeneous logic. In: Glasgow, J., Narayanan, N.H., Chandrasekaran, B. (eds.) Diagrammatic Reasoning: Cognitive and Computational Perspectives, pp. 211–234. MIT Press, Cambridge (1995)
4. Brown, J.R.: Philosophy of Mathematics: An Introduction to a World of Proofs and Pictures. Routledge, London (1999)
5. Clark, A. Magic words: how language augments human computation. In: Carruthers, P., Boucher, J. (eds.) Language and Thought: Interdisciplinary Themes, PP. 162–183. Cambridge University Press, Cambridge (1998)
6. Davis, P.J.: Visual theorems. Educ. Stud. Math. **24**(4), 333–344 (1993)
7. De Cruz, H.: Mathematical symbols as epistemic actions—an extended mind perspective. Unpublished on-line working paper (2005)
8. de Cruz, H.: Innate Ideas as a Naturalistic Source of Mathematical Knowledge. Vrije Universiteit Brussel, Brussel (2007)
9. Eilenberg, S., Steenrod, N.: Foundations of Algebraic Topology. Princeton University Press, Princeton (1952)
10. Fauconnier, G., Turner, M.: The Way We Think: Conceptual Bending and the Mind's Hidden Complexities. Basic Books, New York (2003)
11. Giaquinto, M.: Visual Thinking in Mathematics:an Epistemological Study. Oxford University Press, Oxford (2007)
12. Giaquinto, M.: Crossing curves: a limit to the use of diagrams in proofs. Philosophia. Mathematica. **19**(3), 281–307 (2011)
13. Glasgow, J., Narayanan, N.H., Chandrasekaran, B.: Diagrammatic Reasoning: Cognitive and Computational Perspectives. AAAI Press, Cambridge (1995)
14. Heath, T.L.: The Thirteen Books of Euclid's Elements. Barnes & Nobel, Inc., New York (2006)
15. Hutchins, E.: Material anchors for conceptual blends. J. Pragmatics **37**(10), 1555–1577 (2005)
16. Jacobson, N.: Basic Algebra, 2nd edn. W.H. Freeman and Company, New York (1985)
17. Jensen, A.F.: Metaforens magt. Fantasiens fostre og fornuftens fødsler. Modtryk, Aarhus C (2001)
18. Johansen, M.W.: Naturalism in the Philosophy of Mathematics. Ph.D. thesis, University of Copenhagen, Faculty of Science, Copenhagen (2010)
19. Johnson, M.: The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. University of Chicago Press, Chicago (1990)
20. Kirsh, D.: Thinking with external representations. AI & Soc. **25**(4), 441–454 (2010)

21. Kirsh, D., Maglio, P.: On distinguishing epistemic from pragmatic action. Cogn. Sci. **18**(4), 513–549 (1994)
22. Lakoff, G., Johnson, M.: Metaphors We Live By. University of Chicago Press, Chicago (1980)
23. Lakoff, G., Núñez, R.: Where Mathematics comes from: How the Embodied Mind Brings Mathematics into Being. Basic Books, New York (2000)
24. Larvor, B.: Syntactic analogies and impossible extensions. In: Löve, B., & Müller, T. (eds.) PhiMSAMP. Philosophy of Mathematics: Sociological Aspects and Mathematical Practice. Texts in Philosophy, vol. 11, pp. 197–208. College Publications, London (2010)
25. Magnani, L.: Conjectures and manipulations: external representations in scientific reasoning. Mind Soc. **3**(1), 9–31 (2002)
26. Magnani, L.: External diagrammatization and iconic brain co-evolution. Semiotica **186**, 213–238 (2011)
27. Mancosu, P.: Visualization in logic and mathematics. In: Mancosu, P., Jørgensen, K.F., Pedersen, S.A. (eds.) Visualization, Explanation and Reasoning Styles in Mathematics, pp. 13–30. Springer, Dordrecht (2005)
28. Manders, K.: Diagram based geometric practice. In: Mancosu, P. (ed.) The Philosophy of Mathematical Practice, pp. 65–79. Oxford University Press, Oxford (2010)
29. Messer, R.: Linear Algebra: Gateway to Mathematics. HarperCollins College Publishers, New York (1994)
30. Pasch, M., ehn, M.: Vorlesungen über neuere Geometrie. Die Grundlehren der mathematischen Wissenschaften, vol. 23. Springer, Berlin (1882/1926)
31. Reviel, N.: The Shaping of Deduction in Greek Mathematics: a Study in Cognitive History. Ideas in Context. Cambridge University Press, Cambridge (1999)
32. Shin, S.J.: The Logical Status of Diagrams. Cambridge University Press, Cambridge (1994)
33. Shin, S.J., Lemon, O.: Diagrams. In: Zalta, Edward N. (ed.) The Stanford Encyclopedia of Philosophy, winter 2008 edn. http://plato.stanford.edu/archives/win2008/entries/diagrams/ (2008)
34. Tennant, N.: The withering away of formal semantics? Mind & Lang. **1**(4), 302–318 (1986)