Rev.Phil.Psych. (2015) 6:49–67
DOI 10.1007/s13164-014-0212-5

The Mystery of Deduction and Diagrammatic Aspects of Representation

Sun-Joo Shin

Published online: 20 January 2015

Abstract Deduction is decisive but nonetheless mysterious, as I argue in the introduction. I identify the mystery of deduction as surprise-effect and demonstration-difficulty. The first section delves into how the mystery of deduction is connected with the representation of information and lays the groundwork for our further discussions of various kinds of representation. The second and third sections, respectively, present a case study for the comparison between symbolic and diagrammatic representation systems in terms of how two aspects of the mystery of deduction – surprise-effect and demonstration-difficulty – are handled. The fourth section illustrates several well-known examples to show how diagrammatic representation suggests more clues to the mystery of deduction than symbolic representation and suggests some conjectures and further work.

When we have a choice to get from given premises to a conclusion either by deduction or by induction, the choice is obvious: deduction is the way to go. With a correctly carried out deduction, the truth of the conclusion is guaranteed by the truth of the premises. Hence, after accepting premises to be true and checking each deductive step, we embrace the truth of the conclusion. It is one sure way to secure the certainty of the truth of a proposition. This is the strength of deduction. Hence, mathematics, the discipline where deduction plays a crucial role, has enjoyed its special status in the world of knowledge. However, the strength, some might say, can be a flip side of the weakness of deduction. If the truth of the premises guarantees the truth of the conclusion in a correctly carried out deductive process, it means that the information conveyed in the conclusion is already contained in the premises. All we are doing in deductive reasoning is extracting information (that we are interested in

S.-J. Shin (✉)
Department of Philosophy, Yale University, 344 College St, New Haven, CT 06511-6629, USA
e-mail: sun-joo.shin@yale.edu

articulating) from given information. If so, then deduction does not seem to tell us something new which is not conveyed in the premises.

Things are almost opposite in the case of induction: if we reach a conclusion by inductive reasoning, its truth is not 100 % certain. Nonetheless, there are important reasons why we need to adopt inductive reasoning in spite of the lack of certainty. Let's just imagine, if possible, that inductive reasoning is prohibited. Is science, as we know it, possible? Very unlikely. If certainty would be the only guide for our decision-making, there would be no decisions in our ordinary life, either.

The contrast between deductive and inductive reasoning outlined here highlights a trade-off between certainty and expansion of knowledge. The drawback on each side has been acknowledged: deduction does not expand our knowledge, and Hume's skepticism of the justification of induction is not decisively refuted.

This paper does not contradict these commonly accepted features surrounding the two different kinds of logical reasoning. Embracing the generally received picture, I would like to raise the following two questions about deductive reasoning which lead to the deduction mystery. Naturally we expect surprise at a conclusion reached by induction since the conclusion conveys more information than the premises. Then how about a conclusion of valid deductive reasoning? As noted, the truth of the conclusion is guaranteed by the truth of the premises. Nonetheless, it is more often than not that mathematical theorems have struck us as quite surprising and sometimes almost unbelievable. What is the source of this surprise? Some would respond "I did not know the information of the conclusion was contained in the information of the premises!" Why is it that difficult to recognize the containment relationship? I will call this the *surprise-effect*, which is the first aspect of the deduction mystery.

Another reaction to an unexpected consequence relation might be: "I do not believe the conclusion follows from the premises, and I need to see a proof!" This response invites us to the next topic I plan to take on in this paper: suppose we are told that this conclusion follows from these premises, and we are asked to prove the consequence relation by deduction. Why is this so difficult to show? Some theorems have to wait for decades and centuries to be proven. What is the source of this difficulty, which I will call the *demonstration-difficulty*?

These two issues, the surprise-effect and the demonstration-difficulty, I claim, comprise the mystery of deduction, and it is rather surprising (!) to realize how little attention has been paid to the deduction mystery even though we all have experienced it repeatedly. In the next section we will see how these two aspects of the deduction mystery are entangled with each other in the world of representation.

1 Information Represented and Transparency Compared

Valid deductive reasoning is about extracting a piece of needed information from given information. I claim that the two puzzles above, that is, the surprise-effect and the demonstration-difficulty, are directly related to the representation issue at several levels.

First of all, extracting information almost always involves manipulating information. Here is a subtle, but important, question to raise: Information is invisible, so how are we manipulating this invisible entity? Only in our mysterious mind in a mysterious way? Not always. More interesting and complicated cases require representation of information. (Some might argue that we always need representation, period.) Strictly speaking, manipulating information means manipulating units which represent the given information, and let me abbreviate this as 'manipulating representation.' Therefore, extracting information presupposes the representation of information.

One deduction mystery, the surprise-effect, can be viewed in terms of representation as well. The reason why we are more often than not surprised at a logical consequence is that it is not obvious at all to see that the information of the conclusion is contained in the information of the premises. What do I mean by 'see that the conclusion-information is contained in the premise-information'? Again, what we (literally) see is not information itself, but the representation of the information, and the relation we are interested in seeing (but which is not that obvious to observe) is the relation (that is, the consequence relation) between the premise-representation and the conclusion-representation. If the conclusion-representation is literally a proper part of the premise-representation, which is sometimes the case, the consequence relation is trivial and there is no surprise effect.

But when the conclusion-information is contained in the premise-information, that is, the conclusion is deductively derivable, the relation between the conclusion-representation and the premise-representation is not always clearly observable. In that sense, when we say deductive reasoning is the process of extracting information, the expression 'extracting' is close to a metaphor. If we have an intuition to keep 'extraction' in a literal sense, or we would like to spell out this metaphorical procedure, we may paraphrase the information-extracting process in the following way: in the case of deduction, we transform (!) the premise-representation so that we may more easily see the conclusion-representation in the transformed premise-representation in order to extract the conclusion-information from the premise-information. The entire process is called a proof or a demonstration of a logical consequence relation.

Unpacking (or paraphrasing) the surprise-effect in terms of representation, we have reached the topic of a proof, and it is time to relate the representation issue to our second feature of the deduction mystery, that is, the demonstration-difficulty. We have examined why it is almost always necessary to manipulate the premise-representation in order to show that the conclusion is derivable from the given premises, and as we all have experienced, the majority of derivation processes are quite difficult. There being more than one correct path from premises to a conclusion, in many interesting proofs there is no algorithm or menu to follow to find or create one of those correct paths. These paths are all about transforming one representation into another. In the midst of a lack of mechanical guidance, we believe practice, experience, and insight are required to cope with the difficulty involved here.

So far I have pinned down the two aspects of the deduction mystery in the context of information-representation. Before moving on to the next level, let me put the matter in a counterfactual way to make the same point slightly differently: if we could have represented logical consequence in a transparent way, then there would have been neither a surprise-effect nor a demonstration-difficulty. With a *transparent* *representation*, we may see that the conclusion-representation is a (physical or literal) part of the premise-representation without any manipulation, or, if transformations are needed, then finding a pathway from one representation to another representation is trivial or mechanical. This counterfactual statement tells us that we do not have a way to represent information, especially logical consequence relations, in such a transparent way that we do not need to manipulate given information at all, so that we could read off all of the logically implied conclusions from the given information. Let's call this fictional or utopian state an *absolutely transparent representation*.

This is not the end of the story of the deduction mystery and the representation issue, but only the beginning of a longer story I would like to unfold. Admitting that there is no absolutely transparent way to represent consequence relations, I would like to suggest that we compare transparency among different modes of representation. That is, could we say that some representations are more transparent than others, even though none of them is absolutely transparent? Is it possible that transparency comes on a spectrum? I believe this could be an extremely important and exciting project to pursue.

Before we start, let me say upfront that I do not want to claim that transparency is the only important criterion to evaluate or compare various forms of representation. Accuracy, accessibility, expressiveness, efficiency, etc. are as important as transparency when we assess representation. Also, I do not want to say transparency is a necessary condition for successful or good representation. First of all, as we acknowledged above, there is no absolutely transparent system, and moreover criteria for successful or good representation are context-driven, meaning that it all depends on what one wants to achieve by representation in a given context. In some cases, there might be a trade-off between transparency versus other desirable aspects (e.g. generality or accessibility), and we would have to make a choice. All I would like to focus on in this paper are (i) the relation between the deduction mystery and representation and (ii) transparency of representation as one main factor in the alleviation of the mystery. Hence, my comparison among representations will be limited to the degrees of their transparency.

We have been talking about transparency in the context of the surprise-effect and the demonstration-difficulty, and therefore our comparison between two forms of representation will look into two factors:

(1) Which form of representation exhibits the conclusion-representation in the premise-representation in a more obvious way so that we do not have to transform the premise-representation? (If both forms of representation require manipulations, this factor pushes us to the next factor.)
(2) Which representation guides us more easily in transforming the premise-representation into the conclusion-representation?

In the next section, I present a case study to illustrate how the first sense of transparency could play a role for the comparison between two logically equivalent representation systems, one symbolic and the other diagrammatic. In this case, I will argue that diagrammatic aspects of representation add transparency to representation.

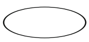
However, I strongly suspect that in many cases (1) could not be a deciding factor. The premise-representation could not possibly exhibit in an obvious fashion all of its logical consequences so that we may not need to transform it. Hence, the problem boils down to (2): Which representation makes it more obvious to see how to transform the premise-representation into the conclusion-representation? At the same time, the obviousness criterion reminds us of a complex aspect of a deductive system: almost every system is devised and presented to be used, and the role or ability of users is at the center of the issue of the deduction mystery. Especially the second sense of transparency above — ‘more easily’ or ‘more obvious’ — highlights not only the role of users but also the user-relative and the context-relative aspects of transparency.[1] Therefore, instead of attempting to reach a general conclusion about the issue, I would like to draw attention to an ongoing uncontroversial practice in the problem-solving process: we have been encouraged to adopt diagrammatic representation over symbolic notation, and we have done so over and over. The second section takes up this common aspect of the problem solving procedure in order to explore how diagrammatic aspects of representation have been considered to be more transparent.

2 One-dimensional versus Two-dimensional Representation

Charles Peirce’s Alpha Graphs are logically equivalent to a sentential language. Both are propositional logical systems, while one is diagrammatic and the other symbolic. As far as logical power goes, they are the same. Hence, I believe this is a good place to start our project of comparing different forms of representation in terms of specific properties we are interested in. Below, I investigate the relative transparency of a sentential language and Peirce’s Alpha system.

[Alpha Graphs]

Basic Vocabulary

1. Sentential Symbols: $A_1, A_2, A_3, \ldots$
2. Cut 

Meaning

1. Cut means negation.
2. Juxtaposition means conjunction.

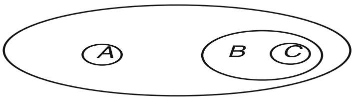
For a long time the Alpha language has been considered to be analogous to a sentential language where we have two connectives — negation (which corresponds to cut) and conjunction (which corresponds to juxtaposition). This view not only

[1] I thank an anonymous reviewer for the recommendation to emphasize the role of users.

misled us about the nature of the Alpha system itself, but also obscured interesting issues involving diagrammatic aspects of Alpha Graphs.

Being equipped with negation and conjunction, the Alpha system is truth-functionally complete. Even though it looks quite different from our familiar symbolic system, we know the two systems are logically equivalent to each other, which might be a relief. But, just as we would like to bring in more (redundant) connectives than negation and conjunction, many were not keen about using Alpha Graphs for carrying out reasoning, wrongly believing that Alpha Graphs have only these two connectives. This is one of the reasons why until recently Peirce's graphical systems were explored mainly by Peirce scholars, not by logicians or mathematicians. That is, the Alpha Graphs were considered to be a meaningful project within Peirce's philosophical framework, but not so much as its own free-standing project.

Some of my previous work showed in detail why the analogy between Peirce's Alpha Graphs and a sentential language with negation and conjunction is wrong from beginning to end. I suggested more visual features be read off of Alpha Graphs. This allows us to translate one and the same Alpha diagram into different sentences without manipulation![2] Let me illustrate the point by an example.



(1) $\neg(\neg A \wedge \neg(B \wedge \neg C))$
(2) $A \vee (B \wedge \neg C)$
(3) $\neg A \rightarrow (B \wedge \neg C)$
(4) $\neg(B \wedge \neg C) \rightarrow A$

When we incorrectly believed that Alpha Graphs had only two kinds of connectives, only sentence (1) was directly obtainable at first. In order to get the other logically equivalent sentences, we needed to transform (1) by using inference rules of sentential logic. However, when we pay attention to other visual features, we may be able to translate the graph directly into all of these sentences.

For reading (2): (i) Depending on whether a sentence symbol, say A, is written in an evenly enclosed or an oddly enclosed cut, we respectively read it off as A or $\neg A$, and (ii) depending on whether juxtaposition takes place in an evenly enclosed or an oddly enclosed cut, we respectively read it off as conjunction or disjunction. In the above example, three sentential symbols are read off as A, B, and $\neg C$; the juxtaposition between a cut of A and the cut of (B and $\neg C$) takes place in an area which is enclosed by an odd number of cuts; and the juxtaposition between B and $\neg C$ occurs in an area which is enclosed by an even number of cuts.

[2] Shin (2002), Section 4.1–4.3.

For readings (3) and (4): Peirce himself named the following form of nested cuts a *scroll*:[3]

A scroll can be read off as a conditional sentence. Depending on how we carve up a scroll, shown below, we get two different readings — (3) and (4):

As this example shows, Alpha Graphs are able to directly express more than negative- and conjunctive-forms of information, but disjunctive- and conditional-forms as well. Then is the Alpha system just like a sentential language with four connectives, that is, negation, conjunction, disjunction, and condition? We should resist a strong temptation to say "yes" to this question. If we equate Alpha Graphs with a four-connective sentential language, I claim that we would commit the same mistake as when we equated it with a two-connective sentential language.

Alpha Graphs could be directly read off as a negative, conjunctive, disjunctive, or conditional sentences, but that does not mean Alpha Graphs function in the same way as a four-connective sentential language. I would draw your attention to three related differences between them.

First of all, Alpha Graphs do not introduce any new syntactic device, whether we read them off with two connectives only (as the traditional method has directed us) or with four connectives (as presented in the above (1) –(4)). On the other hand, a four-connective sentential language has more vocabulary than a two-connective sentential language.

Hence, the next important difference follows: when we introduce a new item of logical vocabulary, we need new rules to tell us how to manipulate this new piece of vocabulary. For example, a natural deduction system presents two rules for each connective, one as an introduction rule and the other as an elimination rule. Accordingly, a two-connective sentential system has four rules of inference, and a four-connective system eight inference rules. But this is not the case with Alpha Graphs: when we treat the Alpha system as a four-connective language (by reading off more visual features), there is no change in its transformation rules, but it keeps the same set of transformation rules as when we take it as a two-connective language. At a syntactic level (both in vocabulary and in inference rules), there is no change whether we let an Alpha Graph be read off as a sentence with two connectives only or as a sentence with four connectives.

The third difference between the above Alpha graph and any of the above sentences (1) – (4) is the following: each sentence corresponds to one and only

[3]Here is C4 [Convention 4]: *The scroll is the sign of a conditional proposition de inesse* (that is, of material implication). (Peirce, Ms 450, p. 14, cited by (Roberts 1973), pp. 33–35.)

one form of information (negative, conjunctive, disjunctive or conditional forms of information). This is why we call sentence (1) a negation sentence, (2) a disjunctive sentence, and (3) and (4) conditional sentences. In order to prove that these different forms of sentences contain the same information — that is, that they are logically equivalent to one another — we need a proof using inference rules of a given system. In contrast, the graph refuses to be named as a specific form of information, but may be read off as any of the sentences (1)–(4), depending on which visual features are salient to the reader. No transformation of the graph is needed to see that the sentences (1)–(4) are equivalent, but we only need to carve up the graph differently.

Let me call this difference *unique versus multiple readability*. I claim unique vs. multiple readability is a main difference between symbolic and diagrammatic representation, and I relate this difference to the first kind of transparency we discussed in the previous section. We said one of the two mysteries of deduction has something to do with the fact that the conclusion-information is not always plainly exhibited in the premise-representation, and as a result we need to transform the premise-representation in order to realize that the conclusion-information is contained there. Surprise!

Let's see how this surprising effect is manifest in our example here. The implication among the four sentences should be proven, that is, a manipulation process is required. On the other hand, in the case of the Alpha graph, we may read off all of these four forms of information directly, without manipulation. That is, the information all of those four sentences express is plainly exhibited in the graph. Hence, there is a lesser degree of mystery in Alpha Graphs than in the case of sentences. Borrowing our terminology in the previous section, there is more transparency in Alpha Graphs than in sentences: all of the four forms of information may be directly represented in the graph, while the equivalence among the four sentences is far from transparent, which is why a proof is needed and a surprise may result.

Why is this the case? The immediate answer points to unique vs. multiple readability. When a unique reading is maintained, we cannot expect various forms of implied information to be read off directly from a given sentence, since one and only one form of information is read off. For example, sentence (1) is a negation sentence, period.[4] In order to show that the same information can be expressed in a conditional sentence as well, we need a proof. On the other hand, multiple readings of Alpha graphs allow one and the same graph to be read off in more than one way.[5] That is, one and the same graph carries information which could be expressed in different ways, for example, (1) – (4). Adopting multiple readings, we do not need to transform the above graph in order to show that what the negation sentence (1) represents is the same what the conditional sentence (3) represents. All we need to do is to carve up a given graph in the way we would like to obtain the information. For example, if one would like to obtain a negation form of information, read off the

[4]A construction history of a given sentence is unique, and in the case of (1) the first negation comes at the end of the construction process. For more technical discussions about unique readability see (Enderton 2001 [1972]), Section 1.4.

[5]For the algorithm of multiple readings, see Section 4.3, (Shin 2012).

entire content inside the outermost cut first and negate it. If one wants to express the same information in a conditional form, reading off a scroll at the end would do the job. Since the user is reading off one and the same graph, no proof is needed, and there is no surprise either. Unique vs. multiple readability leads to the following crucial difference between sentential and graphical representation: the relation among the sentences is not explicitly shown, and needs to be shown (i.e. proven), while the graph exhibits different forms of the same information and there is no need to derive one from another.

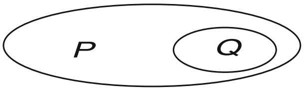
Then what is the root of unique vs. multiple readability? Why cannot we allow multiple readings for sentential languages? How can we justify multiple readings for diagrams? What would we lose if multiple readings were prohibited for diagrams? The contrast, I claim, is linked to one of the fundamental differences between symbolic and diagrammatic representation: symbolic representation is linear, that is, one-dimensional, while diagrammatic representation is two-dimensional. Linear representation cannot afford more than one reading for a given formula, since that could cause ambiguity. Therefore, we do not allow a string like 'P & Q ∨ R' to be grammatical: one could parse it either as a conjunctive or as a disjunctive sentence. Parentheses visually tell us how to enforce unique readability. Both '(P &Q) ∨ R' and 'P & (Q ∨ R)' are grammatical units, and the semantics tells us that their meanings are different from each other. Parentheses guarantee a unique construction history for each grammatical string, and visually it is quite natural and intuitive. Alternatively and in the same fashion, according to the prefix notation, the string '∨&PQR' should be read one and only one way, making a contrast with '&∨PQR.'

Interestingly enough, there is no visual device in the Alpha system analogous to the parentheses of a sentential language. That does not mean that we cannot enforce unique reading, and a traditional way to set up the semantics of Peirce's graphs allows only one reading – hence the sentence (1) only in the above example. I claim that unique readability of diagrams is too artificial and unnatural for users. Our experiences tell us how prevalent Gestalt phenomena are in the case of two-dimensional objects, for example, Necker's cube, Rubin's vase, etc. Depending on which part of the picture becomes a foreground or a background, how to group elements of a picture, or how to carve up a given picture, we get different interpretations of one and the same visual object. An Alpha graph, being two-dimensional, can also be carved up in more than one way. So why not allow multiple readings of it? In the case of a sentential reading of the string 'P & Q ∨ R,' multiple readability would be a disaster, which is why the string is not grammatical and no semantics is applied to it. An extremely interesting and important fact is that multiple readability does not cause an ambiguity problem for Alpha Graphs.[6] Then it is more intuitive and natural for the user to read off a given graph in the way he happens to carve it up. Taking advantage of two-dimensionality is the way to go.

When we do not pay attention to this fundamental difference, it is not surprising that one kind of approach to representation is unfairly imposed on another. The analogy between the Alpha System and a two-connective sentential language

[6]For a formal proof, see Section 4.2.2, (Shin 2002).

is a prime example of such an incorrect assessment with correspondingly undesirable consequences. Assuming that the Alpha System has only two connectives, Don Roberts emphasizes the importance of the order of negation and conjunction for the translation of the following Peircean diagram:[7]



> Notice that we do *not* read it: 'Q is true and P is false', even though Q is evenly enclosed and P is oddly enclosed [. . .] we read the graph from the outside (or least enclosed part) and we proceed inwardly [. . .]

Nested negation and conjunction sentences were obtained as translations of Alpha graphs, and for this reason much manipulation was expected to transform complicated-looking nested sentences into simpler-looking sentences. Under this kind of approach, one cannot hope that Alpha Graphs will yield new insights. It is quite instructive to learn that an analogy between Alpha Graphs and a four-connective sentential language is incorrect as well. No matter what, no sentential language can afford to abandon unique readability, while Alpha Graphs thrive on the multiple readability. Because Alpha Graphs utilize multiple readings, the containment of the conclusion-information in the premise-representation is manifest, and the system becomes transparent.

In light of this limited case study, my conclusion is that diagrammatic representation, being two-dimensional representation, may allow multiple readings depending on how one carves up a given graph. Multiple readability may make implied information more transparent than when a unique reading is enforced as in the case of symbolic representation. Hence, one aspect of the deduction mystery (which we called the surprise-effect) becomes somewhat manageable.

3 Mind Stimulated

The second aspect of the deduction mystery, the demonstration-difficulty, is an even bigger issue, since in many interesting and complicated cases we do not expect that the conclusion-information is easily readable from the premise-representation. Even after we are told that the conclusion logically follows from the premises, it may not be easy to find a proof to show the logical consequence. That is, in a deductively valid argument, the conclusion, in some sense, is contained in the premises, but often not in an obvious way. The main task is to transform the given premises, by manipulating the premise-representation, in order to make the containment relation obvious. Again,

[7] (Roberts 1973), p. 39.

I will argue that how information is represented could be at least as important an issue for the demonstration-difficulty.

Let P be premises and C a conclusion. Suppose we need to show the logical consequence relation between them. As said above, most probably we need intermediate stages to get to C from P. That is,

$$\mathrm{P} \longrightarrow \mathrm{D}_1 \cdots\cdots \longrightarrow \mathrm{D}_\mathrm{n} \longrightarrow \mathrm{C}$$

To simplify our story, let's collapse these n intermediate stages. So,

$$\mathrm{P} \longrightarrow \mathrm{D}^* \longrightarrow \mathrm{C}$$

Please note that we are not talking about a unique series $D_1, \cdots\cdots, D_n$, but one of the correct series. The mysterious demonstration-difficulty is about finding one of the "correct" $D*$s.

What do I mean by 'correct'? One necessary condition to be correct is that $D*$ should be a logical consequence of P. However, that would not be enough for a series to be a correct $D*$. There are many, possibly an infinite number of, sequences that are logical consequences of P, but not all of them would be a correct one to get to C. Similarly, we can think about many $D*$s from which C follows. But, again, many of them would not be directly helpful to get to C. That is, a given $D*$ is correct not only because it is a logical consequence of P and C is a logical consequence of $D*$, but also because it is helpful in bridging the gap between P and C. There is no menu or fixed algorithm to make a correct choice among the many alternatives, that is, the many logically consistent alternatives. This is the core of the mysterious demonstration-difficulty.

I would like to relate the current discussion to our common practice in searching for a right solution — adopting diagrams as a heuristic tool. Whether one accepts a sequence of diagrams as a rigorous proof or not, we often draw diagrams in sketching or outlining a proof. A strong recommendation for this practice is found in Polya's classical work "How to Solve it."[8]

> [E]ven if your problem is not a problem of geometry, you may try to *draw a figure*. To find a lucid geometric representation for your nongeometrical problem could be an important step toward the solution.

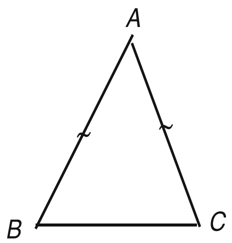
What is even more intriguing, we do not often encounter the practice of running in the opposite direction, that is, converting diagrammatic problems into symbolic representation. There must be something special about diagrammatic representation which helps with the demonstration-difficulty of deduction, and we would like to pin down what it is about diagrammatic representation which symbolic representation lacks.

[8](Polya 1973 [1971]), pp. 107–8.

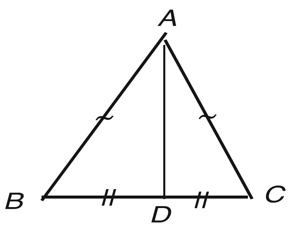
Let me start with a simple standard textbook proof. This is not Euclid's corresponding original proof, but a simpler one which will suit our needs. For easy reference in our further discussions, let me number the steps.

Proposition 1 *If two sides of a triangle are congruent, the angles opposite them are congruent.*

Proof Given a triangle, ABC, we want to prove $\angle ABC = \angle ACB$.



(1) We draw a line from A to the center of the line BC, say D.

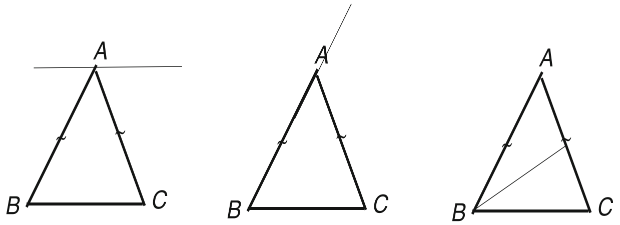


(2) Then, $\triangle ABD$ is congruent to $\triangle ACD$.
(3) Therefore, we know $\angle ABC = \angle ACB$.

□

As our experience of this short proof tells us, the demonstration-difficulty lies in the step where we bring in a line AD — step (1). Once the line AD is drawn, the rest of the proof seems to be quite smooth: we see two congruent triangles, $\triangle ABD$ and $\triangle ACD$, emerge. Then, by the definition of congruency, we know the two angles, $\angle ABC$ and $\angle ACB$, are congruent to each other.

What is special about step (1), compared with steps (2) and (3)? First, in step (1) we are drawing an auxiliary line, which means we are introducing a new object into a proof. On the other hand, in steps (2) and (3), all we do is observe and recall a definition. Here, action (like drawing) contrasts with observation or recollection. Second, there is no specific guideline for the action of (1), except that it should be approved by axioms or by previously proven theorems. In this case, the issue is not whether we can draw line AD or how we can find a center of line BC, etc. The heart of the matter is this: How do we know we need to draw a line from A to the center of the line BC to get to the conclusion? The fact is that there are many other logically valid steps available to us. For example,



All of these new objects could have been introduced and none of them is illegitimate. However, not all of them would make the rest of the proof go smoothly. That is, steps (2) and (3) do not get prompted by these other lines unlike our line in step (1) above. Among logically valid multiple choices, we need to find an object which fits the bill, and there is no algorithm or menu for a right choice. Hence, the demonstration-difficulty in this case amounts to the difficulty of introducing a correct new object into a proof. It requires ingenuity!

Back to our main topic, the relation between the deduction mystery and forms of representation: Could different forms of representation make a difference in this ingenious stage? I would like to propose that diagrammatic representation stimulates the mind in such a way as to prompt a new correct auxiliary object more easily than symbolic representation does. Why is this the case?

There have been warnings against the use of diagrams in a rigorous proof. The main reason is that we might rely on accidental properties of a diagram we happen to draw. That is, diagrams could mislead us. Can we overcome this misleading aspect of diagrammatic representation? We know we cannot get rid of accidental properties of a token-diagram, since a diagram, once being drawn, has both universal and particular properties. An infamous ridicule by Berkeley of Locke's abstract triangle could be viewed as a dilemma facing diagrammatic representation.

> If any man has the faculty of framing in his mind such an idea of a triangle as is here [quoted passage about the general idea of a triangle out of *An Essay on Human Understanding*, (IV. vii. 9)] described, it is in vain to pretend to dispute him out of it, nor would I go about it. All I desire is that the reader would fully and certainly inform himself whether he has such an idea or no. And this, methinks, can be no hard task for anyone to perform. What more easy than for anyone to look a little into his own thoughts, and there try whether he has, or can attain to have, an idea that shall correspond with the description that is here given of the general idea of a triangle, which is "neither oblique nor rectangle, equilateral, equicrural nor scalenon, but all and none of these at once?"[9]

[9](Berkeley 1920 [1776]), Introduction Section 13

How could a triangle be drawn without its particular shape, size, etc.? Particularity is inevitable in the case of diagrammatic representation. This is a main reason why diagrams have been avoided in mathematical or logical proofs, where generality or universality is especially important. Symbolic representation comes at the other end in terms of the particularity-versus-generality spectrum.[10] When we say 'let ABC be a triangle,' we do not attribute any specific size to the triangle, and there is no particularity whatsoever. ABC is a triangle in general. Hence, it seems to be quite natural that mathematical proof has been dominated by symbolic notation.

Embracing the ongoing practice and understanding the rationale behind the dominance of symbolic representation, let us go back to the issue we raised at the beginning of the section: Why, nonetheless, has it been the case that diagrams are used as a heuristic tool? I argue that the answer has something to do with the particularity of diagrammatic representation. Ironically, we have seen that particularity, which causes a lack of generality, has prevented diagrams from getting into proofs. Then how could this property help us adopt diagrams at the stage of finding a proof?

In the above simple proof example, we identified the step where the mystery of deduction lies, that is, the root of the demonstration-difficulty: we need to bring in a new object, and for this activity there is no mechanical algorithm to follow. Hence, it requires a creative and ingenious mind to search for a correct auxiliary object so that the object may provide us with a new way of viewing the issue we are working on. Then let's compare the figure drawn above as a triangle and the symbolic notation ABC as an isosceles triangle. Both kinds of notation work as the representation of an isosceles triangle in general, but in strikingly different ways. A token of an isosceles triangle drawn is a particular isosceles triangle while ABC is declared to be an isosceles triangle in general.

As we know, Euclidean proofs can be arithmetized, and hence we may present our example proof as a symbolic deduction sequence. In a step corresponding to where the auxiliary line AD is drawn, we may find an existential introduction inference rule applied to an axiom, a definition, or a previously proven theorem. A crucial question is: Which representation helps us to find an auxiliary line so that two congruent triangles may emerge, leading us to conclude that the angles we are interested in are congruent to each other? Obviously, the particular diagram drawn is the answer. It is not because the drawn triangle itself has a specific size and a particular shape, but because the drawn triangle exhibits essential features of an isosceles triangle in general. In contrast, symbolic representation ABC removes all of the particularity which any diagrammatic representation could have, and therefore it successfully acquires generality which is a powerful property of mathematics. There is a price, though. ABC represents an isosceles triangle by stipulation and shows nothing any isosceles triangle has, for example, having three sides two of which have the same length. That is, while it does not have any particular property that any particular token of an isosceles triangle could possess, symbolic representation ABC does not have any

[10]For further discussions of this topic, refer to Shin (2012).

property every isosceles triangle shares either. It is an extreme case of a lack of particularity.

This deficiency, I claim, makes the demonstration-difficulty mystery even more difficult for symbolic than for diagrammatic representation. Diagrammatic representation displays common features of every isosceles triangle, along with specific accidental features of an isosceles triangle-token. Visually vivid features stimulate the mind to create a new object more easily than when no visual feature is available. Without these visual features we need to recall the definition of an isosceles triangle to mind and search for a right axiom or previously proven theorem to be used here. With symbols, nothing is available to stimulate the mind to narrow down choices of a new object, but we have only our memory to rely on.

Examining the issue carefully, we realize that it is not correct to say that diagrams lack generality. Instead, the risk is that generality might be blurred or overridden by particularity. As our practice shows, when needed, visually represented generality activates the mind in a more lively fashion so that we may experiment on new objects more efficiently. No wonder at the brainstorming stage we draw and draw! Furthermore, noting that generality is represented in diagrams, we may set up guidelines to clarify which represented properties are essential, and not accidental, properties. Then we can avoid misusing diagrams, so that diagrammatic formalism is justified, as recent literature shows.[11] If so, there is no need to limit the language in a proof to symbolic representation. This is another important topic, but beyond the scope of this paper.

4 More Examples and Further Work

The two case studies, Alpha Graphs in Section 2 and a geometry proof in Section 3, respectively illustrated a comparison between diagrammatic and symbolic representation in terms of two aspects of the deduction mystery — the surprise-effect and the demonstration-difficulty. In both cases, diagrammatic representation makes the deduction process more transparent than symbolic representation. The two dimensionality of Alpha graphs lets the reader carve up one and the same diagram in more than one way, which reduces the surprise-effect, and thereby makes the deduction mystery somewhat manageable. The particularity of figures in a Euclidean proof stimulates our mind to introduce a new object to the discourse so that new figures may be constructed to find a right solution — helping to solve the demonstration-difficulty. Some might pause here with the following comments: a Euclidean proof is about geometric figures. Hence, we should not be surprised to see that diagrammatic representation is more advantageous to the reasoning process than symbolic notation.

[11] (Anderson et al. 2000), (Barwise 1993), (Barwise and Allwein 1996), (Barwise and Etchemendy 1989), (Barwise and Etchemendy 1991), (Barwise and Etchemendy 1995), (Barwise and Hammer 1994), (Chandrasekaran et al. 1995), (Hammer 1995), (Harel 1988), (Miller 2008), (Mumma 2006), (Shin 1994), (Shin 2002), (Sowa 1984) and http://www.cmis.brighton.ac.uk/research/vmg.

Taking this comment seriously, this section brings in non-geometric examples to confirm the claims that have been made in previous sections. First, let me present three different styles of a proof for a well-known non-geometric proposition and compare them. I would also like to briefly introduce a familiar diagrammatic representation in order to encourage further work on the relation between representation forms and our reasoning process.

Proposition 2 $\forall n\ (1 + \cdots + n = [n \cdot (n+1)]/2)$.

Let's consider three different ways to prove the proposition.

Proof by induction on n This proof does not bring in any diagrammatic representation.

(Base step) It is when $n = 1$. $1 = [1 \cdot 2]/2$.

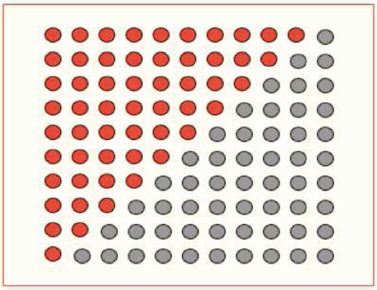
(Inductive Step)

Inductive Hypothesis: Assume that $1 + \cdots + k = [k \cdot (k+1)]/2$ for some k.

[Show $1 + \cdots + (k+1) = [(k+1) \cdot ((k+1)+1)]/2$.]

$$\begin{aligned} 1 + \cdots + (k+1) &= 1 + \cdots + k + (k+1) \\ &= [k \cdot (k+1)]/2 + (k+1) \quad \text{by IH} \\ &= [k \cdot (k+1)]/2 + 2(k+1)/2 \\ &= [k \cdot (k+1) + 2(k+1)]/2 \\ &= [(k+1) \cdot (k+2)]/2 \\ &= [(k+1) \cdot ((k+1)+1)]/2 \end{aligned}$$

Proof by pebbles The following diagram is a key to the proof.



When $1 + \cdots + n$ pebbles are copied, we see n rows and $(n+1)$ columns, that is, $n \cdot (n+1)$ pebbles in the square. Hence, we obtain $1 + \cdots + n = [n \cdot (n+1)]/2$.

Proof by two-dimensional presentation This proof does not use a diagram, but intuitively it is closer to the pebble argument than to the inductive proof. Why is this the case?

Let's write up the proposition $(1 + \cdots + n)$ in two slightly different ways:

$$\begin{array}{ccccc} 1 + & 2 & + \cdots + & (n-1) & + n \\ n + & (n-1) & + \cdots + & 2 & + 1 \end{array}$$

Now a new form emerges and we could write up one more line:

$$\begin{array}{ccccc} 1 + & 2 & + \cdots + & (n-1) + & n \\ n + & (n-1) & + \cdots + & 2 + & 1 \\ \hline (n+1) + & (n+1) & + \cdots + & (n+1) + & (n+1) \end{array}$$

Then the last line amounts to $n \cdot (n+1)$ and the original line we aim to compute is $[n \cdot (n+1)]/2$. Hence, $1 + \cdots + n = [n \cdot (n+1)]/2$. A key step here is to represent two formulas in a two-dimensional way so that we may come up with a new way to add $1+n$, $2+(n-1)$, etc. This, I believe, is a main reason why the third proof could be classified as a diagrammatic proof, even though there is no explicit diagram used. In that sense, I would like to categorize both the pebble argument and this argument as diagrammatic.

This is a classic example to show that the (heuristic) power of diagrams is not limited to geometric problems and that a non-geometric problem can benefit from diagrammatic representation in order to gain insight. However, it would be too hasty to conclude that one form of representation is better than another. On the contrary, I would like to say that the choice of the right kind of representation is highly context-dependent. Let me discuss one more example in order to illustrate further aspects of various forms of representation.

Representation of a Turing Machine

All of (a), (b), and (c) represent a Turing Machine which writes three 1's on a blank tape:[12]

(a) Set of quadruples

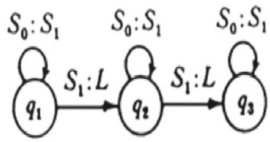
$\langle q_1, S_0, S_1, q_1 \rangle$, $\langle q_1, S_1, L, q_2 \rangle$, $\langle q_2, S_0, S_1, q_2 \rangle$, $\langle q_2, S_1, L, q_2 \rangle$, $\langle q_2, S_0, S_1, q_3 \rangle$

(b) Machine table

State	Scanning	
	$S_1 q_1$	$L q_2$
	$S_1 q_2$	$L q_3$
	$S_1 q_3$	

[12](Boolos et al. 2002), p. 26. Symbol S_0 stands for a blank and S_1 for 1.

(c) Flow graph



Representation (a), being linear, is symbolic, while representation (c) looks diagrammatic. How about the machine table (b)? Intuitively, it looks diagrammatic, but it is important to clarify our intuition: In what sense is (b) closer to (c) than to (a)?

Suppose one tries to add an instruction, say, if the machine scans a blank in state 2, it is to go to the left cell and stay in the same state. This instruction contradicts existing ones. In the case of a flow chart, when one draws an arrow from q_2 to itself, written S_0:L, a contradiction between this new arrow and S_0:S_1 can be easily caught. Quadruple representation (a), mainly because of its linearity, would allow one more quadruple $\langle q_2, S_0, L, q_2 \rangle$ to be added more easily without perceiving a contradiction. How about the machine table representation (b)? Go to where the column for blank (which is the first one) and the row for state 2 (which is the second row) meet and write an entry Lq_2. Then we will see that the spot is already occupied by the entry S_1q_2, which means that a new instruction contradicts an existing one. That is, representation (b) also makes adding contradictory pieces of instruction more transparent than quadruple representation (a). In that sense, both (b) and (c) are two-dimensional representations, while (a) is one-dimensional.

On the other hand, a loop, if any, is represented much more graphically and clearly in (c) than in (a) or (b). If an arrow is drawn from q_3 to q_1 with S_1:L, we perceive a loop, period. In the case of (b), the empty spot would be filled with Lq_1, but a loop is not clearly visualized as in (c). How many of us would try to design a Turing Machine using quadruples or even a machine table? The flow chart is clearly the easiest to read off and hence the most popular way to design Turing machines. However, when we prove that Turing machines are enumerable, understanding a Turing machine as a linear concatenation of vocabulary is crucial. The quadruple representation can easily convince us that Turing machines (as a set of finite words) are enumerable, and hence there must be functions Turing machines cannot compute. Why? We have a proof that functions are not enumerable.[13] This is a good example to show that each form of representation has its own strengths and weaknesses, and I believe more work can be done in the future on this topic.

Several hunches in random order: Two-dimensional representation has more space freedom than one-dimensional linear representation. Related to this greater freedom, a new configuration more easily emerges when new information is added. That is, Gestalt phenomena are prevalent in two-dimensional representation. This might be one of the weakest aspects of linear representation. Since every piece of information is linearly added, in order to figure out relations among given pieces of information,

[13] Non-enumerability of functions can be proven using a diagonalization, and this is another well-known place where diagrammatic representation is more intuitive.

linear representation usually requires serious manipulation. This is one of the reasons why deduction is difficult even though the conclusion-information is contained in the premise-information. On the other hand, there must be a reason why most of the historical record has been kept in an overwhelmingly linear form. Linearity of time is one factor, some have suggested. Other practical advantages — that it takes up less space, that it is easier to add things one by one, etc. — are hard to ignore. Overall, I believe that different forms of representation are complimentary to one another. It will be extremely interesting and important to look into more accurate descriptions of the relationships among various kinds of representation so that an appropriate kind of representation may be chosen for any given purpose.

References

Anderson, Michael, Peter Cheng, and Volker Haarslev (Eds.) 2000. *Theory and Application of Diagrams*. Berlin: Springer.

Barwise, Jon. 1993 In G. Mineau, B. Moulin, and J. Sowa (Eds.) *Heterogeneous reasoning In Conceptual Graphs for Knowledge Representation*, (pp. 64–74). Berlin: Springer-Verlag.

Barwise, Jon, and Gerard Allwein (Eds.) 1996. *Logical Reasoning with Diagrams*. New York: Oxford University Press.

Barwise, Jon, and John Etchemendy. 1989. Information, Infons and Inference In Cooper, Mukai, and Perry (Eds.) *Situation Theory and its Applications I. In: vol. 1*, (pp. 33–78). Stanford, California: Center for the Study of Language and Information.

Barwise, Jon, and John Etchemendy 1991. Visualization in Teaching and Learning Mathematics W. Zimmerman, and S. Cunningham (Eds.), Mathematical Assoication of America, Washington, DC.

Barwise, Jon, and John Etchemendy. 1995. Heterogeneous reasoning. *In Chandrasekaran* (pp. 211–234).

Barwise, Jon, and Eric Hammer 1994. What is a Logical System D. M. Gabbay (Ed.), Oxford University Press, New York.

Berkeley, Geroge. 1920 [1776]. *A Treatise Concerning The Principles of Human Knowledge*. Chicago: Open Court.

Boolos, George, John Burgess, and Richard Jeffrey. 2002. *Computability and Logic*. Cambridge, United Kingdom: Cambridge University Press.

Chandrasekaran, B., J. Glasgow, and N.H. Narayanan 1995. Diagrammatic Reasoning: cognitive and computational perspective. Menlo Park, Calif.: AAAI Press; Cambridge, Mass.: MIT Press.

Enderton, Herbert B. 2001 [1972]. *A Mathematical Introduction to Logic*. New York: Academic Press, Inc.

Hammer, Eric. 1995. *Logic and Visual Information Stanford*. CA: Center for the Study of Language and Information.

Harel, David. 1988. On visual formalisms. *Commun ACM* 13(5): 514–530.

Miller, Nathaniel 2008. Euclid and his Twentieth Century Rivals: Diagrams in the Logic of Euclidean Geometry. CSLI Studies in the Theory and Applications of Diagrams.

Mumma, John 2006. Intuition Formalized: Ancient and Modern Methods of Proof in Elementary Geometry PhD thesis, Carnegie Mellon University.

Polya, G. 1973 [1971]. *How to Solve It: A New Aspect of Mathematical Method*. Princeton, New Jersey: Princeton University Press.

Roberts, Don 1973. The Existential Graphs of Charles S. Peirce. The Hague: Mouton.

Shin, Sun-Joo. 1994. *The Logical Status of Diagrams*. New York: Cambridge University Press.

Shin, Sun-Joo. 2002. *The Iconic Logic of Peirce's Graphs*. Cambridge, Massachusetts: MIT Press (Bradford).

Shin, Sun-Joo. 2012. The forgotten individual: diagrammatic reasoning in mathematics. *Synthese* 186(1): 149–168.

Sowa, John 1984. Conceptual Structure: Information Processing in Mind and Machine. Reading, Mass.: Addison-Wesley.

Visual Modelling Group, University of Brighton. http://www.cmis.brighton.ac.uk/research/vmg.