

Rationale of the Mathematical Joke

Andrew Aberdein¹

Abstract. A widely circulated list of spurious proof types may help to clarify our understanding of informal mathematical reasoning. An account in terms of argumentation schemes is proposed.

1 INTRODUCTION

A long-standing complaint of researchers into scientific practice is that the published works of scientists do not reveal, indeed willfully conceal, the cognitive processes that led to their results. R. K. Merton traces this complaint as far back as Francis Bacon [14, p. 4]. Researchers into mathematical practice must therefore have resort to less obvious sources. One neglected but potentially revealing object of inquiry is the folklore which circulates within the mathematical community: slang, proverbs, legends, and jokes [17, p. 24]. As Freud observed long ago, the last of these in particular can prove inadvertently revealing: a ‘suppressed purpose can, with the assistance of the pleasure from the joke, gain sufficient strength to overcome the inhibition, which would otherwise be stronger than it’ [6, p. 187]. Or, as the folklorist Gershon Legman puts it in the introduction to his *Rationale of the Dirty Joke*: ‘What is suggested here is not that all these stories are necessarily true ... What is meant is that these stories and individuals do personify ... real but inexplicable peculiarities of human behavior, which they are attempting somehow to fit into a rational view of the world, whether as horror or as humor’ [12, p. 22].

This paper explores one notable source of mathematical horror and/or humor: a collection of spurious proof types first compiled by Dana Angluin as [3], and reprinted in full (with two additions, but no attribution) in [17, pp. 28 f.]. The explicit and implicit arguments they embody will be analyzed in terms of ‘argumentation schemes’: stereotypical reasoning patterns, often accompanied by ‘critical questions’, which itemize possible lines of response. The most thorough presentation of this methodology to date is that of Douglas Walton and his collaborators [21]. Although argumentation schemes have been developed primarily for non-mathematical contexts, many of them lend themselves readily to mathematical application. For further discussion of this application, and some additional examples of argumentation schemes applied to mathematics, see [1, 2, 4]. The classification of argumentation schemes is an open problem, complicated, perhaps intractably, by the multiple dimensions of similarity between schemes. Walton proposes a threefold division between reasoning, source-based arguments, and the application of rules [21, pp. 348 f.]. This order is broadly followed below, but some of the categories are subdivided to highlight problematic aspects of mathematical argumentation.

2 RETRODUCTION

Much mathematical reasoning is sound deductive inference, secure from the anxieties that jokes such as Angluin’s reflect. But not all of it. One of the most ancient techniques of mathematical problem-solving is retroduction, or working backwards. In one of its aspects, at least, this was known to ancient mathematicians as analysis, a term which has, of course, acquired other senses. Pappus defined it as follows: ‘in analysis we assume that which is sought as if it were (already) done, and we inquire what it is from which this results, and again what is the antecedent cause of the latter, and so on, until by so retracing our steps we come upon something already known, or belonging to the class of first principles, and such a method we call analysis as being solution backwards (*ἀνάλυσις λύσεων*)’ [8, p. xlvii]. The steps which the analysis traces should be deductively sound, and once the process is complete the order of the steps may be reversed to yield a conventional deductive proof. This is familiar in the strategy of reducing an open problem to one that has already been solved. Of course, the target must be chosen carefully, as one of Angluin’s spurious proofs indicates:

Proof by reduction to the wrong problem ‘To see that infinite-dimensional coloured cycle stripping is decidable, we reduce it to the halting problem.’ [3, p. 17]

Multi-step argumentation, often of substantial length or complexity, is characteristic of, but not unique to mathematical reasoning. It may be understood in terms of the following scheme:

2.1 Argument from Gradualism

Premise Proposition A is true (acceptable to the respondent).

Premise There is an intervening sequence of propositions, $B_1, B_2, \dots, B_{n-1}, B_n, C$ such that the following conditionals are true: If A then B_1 ; if B_1 then B_2 ; ...; if B_{n-1} then B_n ; if B_n then C .

Premise The conditional ‘If A then C ’ is not, by itself, acceptable to the respondent, nor are shorter sequences from A to C (than the one specified in the second premise) acceptable to the respondent.

Conclusion Therefore, the proposition C is true (acceptable to the respondent). [21, p. 339]

When a problem is to be reduced to a previously proved result, the task of proving C becomes that of finding a suitable theorem A , and the intervening sequence of propositions, $B_1, \dots, B_n$ required for this scheme.

Uncontentious deductive examples are not the only varieties of backward reasoning to arise in mathematics:

Proof by importance A large body of useful consequences all follow from the proposition in question. [3, p. 17]

This piece of reasoning is interestingly ambiguous. On the one hand, it might be perceived as consequentialist practical reasoning:

¹ Humanities & Communication, Florida Institute of Technology, U.S.A., email: aberdein@fit.edu

2.2 Argument from Positive Consequences

Premise If A is brought about, then good consequences will plausibly occur.

Conclusion Therefore, A should be brought about.

Critical Questions:

1. How strong is the likelihood that the cited consequences will (may, must) occur?
2. What evidence supports the claim that the cited consequences will (may, must) occur, and is it sufficient to support the strength of the claim adequately?
3. Are there other opposite consequences (bad as opposed to good, for example) that should be taken into account? [21, pp. 332 f.]

Such reasoning is employed in defence of contested axioms or principles, such as the axiom of choice or the law of excluded middle, but precisely because as axioms they defy proof. As a spurious proof, ‘proof by importance’ seems rather to work by suggesting that a widely employed but unproven conjecture must be true, because of its extensive coherence with other results. This sort of hunch was employed in defence of, for example, the Taniyama-Shimura hypothesis prior to its confirmation by Andrew Wiles.

But reasoning backwards from the consequences of a conjecture may also be understood as abductive reasoning. George Pólya discusses the importance of abduction for Euler in convincing himself of the correctness of his novel proof that $\sum 1/n^2 = \pi^2/6$ through representation of $\sin x$ as an infinite product: ‘Euler does not reexamine the grounds for his conjecture ... he examines only its consequences. He regards the verification of any such consequence as an argument in favor of his conjecture ... In scientific research as in ordinary life, we believe, or ought to believe, a conjecture more or less according as its observable consequences agree with the facts. In short, Euler seems to think the same way as reasonable people, scientists or non-scientists, usually think’ [16, vol. 1, p. 22]. Of course, despite the importance of this informal procedure for Euler, he never mistook it for proof. Indeed he kept worrying away at the problem until he found an independent proof in more conventional terms.

Euler’s inference, and some cases of Angluin’s ‘proof by importance’, might thus be characterized by the positive form of the following abductive argumentation scheme:

2.3 Argument from Evidence to a Hypothesis

Premise If A (a hypothesis) is true, then B (a proposition reporting an event) will be observed to be true.

Premise B has been observed to be true [false], in a given instance.

Conclusion Therefore, A is true [false].

Critical Questions:

1. Is it the case that if A is true, then B is true?
2. Has B been observed to be true (false)?
3. Could there be some reason why B is true, other than its being because of A being true? [21, pp. 331 f.]

The falsification variant is *modus tollens*, and thus uncontroversial. From a strict falsificationist perspective, the best that can be said of any hypothesis is that it has not (yet) been falsified. The defeasible acceptance of a hypothesis on the grounds of non-falsification is another application of the positive variant of this scheme. For empirical science, this may come close enough to confirmation for practical cases. In mathematics the effect is less convincing:

Proof by accumulated evidence Long and diligent search has not revealed a counterexample. [3, p. 17]

Of course this is no sort of ‘proof’. If the scope of the search is explicitly stated, a more limited result could be confirmed. But counterexamples revealed only by *extraordinarily* long and diligent searches are common enough to undercut any hope that such reasoning may hold in general. For example, Euler suggested in 1769 that $a^4 + b^4 + c^4 = d^4$ has no integer solutions, but it was not until 1986 that a counterexample in the millions was uncovered [20, p. 293]. It is because of such counterexamples, which have few empirical analogues, that the third critical question can only be fully answered in mathematical cases by proving the hypothesis.

3 CITATION

Source-based arguments are a rich source of mathematical anxiety and humour. The most numerous cases on Angluin’s list are those which turn on deviant citation practices. As Paul Renteln and Alan Dundes note, this prevalence ‘hints at the anxieties felt by many mathematicians regarding the degree to which mathematical truth is dependent upon the trustworthiness of previous results. This anxiety is exacerbated by the fact that some mathematicians have a less rigorous proof style than other mathematicians’ [17, p. 28]. Several subvarieties may be discerned, many of them painfully familiar, in which citation fails for distinct reasons. Firstly there can be references to things which do not exist at all, or which do not say what they are supposed to:

Proof by wishful citation The author cites the negation, converse, or generalization of a theorem from the literature to support his claim.

Proof by ghost reference Nothing even remotely resembling the cited theorem appears in the reference given. [3, pp. 16 f.]

A more subtle issue is that the level of confidence which the source ascribes to the claim may not be adequately reflected in the citation of that claim. Angluin’s juxtaposition of the next pair of examples nicely illustrates an anxiety about the credibility arbitrage which may lie behind some references:

Proof by eminent authority ‘I saw Karp in the elevator and he said it was probably NP-complete.’

Proof by personal communication ‘Eight-dimensional coloured cycle stripping is NP-complete [Karp, personal communication].’ [3, p. 16]

In other cases, citations may send the diligent on a wild goose chase which is practically (or even logically) impossible to complete:

Proof by reference to inaccessible literature The author cites a simple corollary of a theorem to be found in a privately circulated memoir of the Slovenian Philological Society, 1883.

Proof by mutual reference In reference A, Theorem 5 is said to follow from Theorem 3 in reference B, which is shown to follow from Corollary 6.2 in reference C, which is an easy consequence of Theorem 5 in reference A. [3, p. 17]

Most transparently, the citation may refer to future activity:

Proof by forward reference Reference is usually to a forthcoming paper by the author, which is often not as forthcoming as at first. [3, p. 17]

Proof by deferral ‘We’ll prove this later in the course.’ [17, p. 28]

All of these different cases may be understood as ineffective applications of the same argumentation scheme, that for Appeal to Expert Opinion.

3.1 Appeal to Expert Opinion

Major Premise Source E is an expert in subject domain S containing proposition A .

Minor Premise E asserts that proposition A is true (false).

Conclusion A is true (false).

Critical Questions:

1. Expertise Question: How credible is E as an expert source?
2. Field Question: Is E an expert in the field that A is in?
3. Opinion Question: What did E assert that implies A ?
4. Trustworthiness Question: Is E personally reliable as a source?
5. Consistency Question: Is A consistent with what other experts assert?
6. Backup Evidence Question: Is E 's assertion based on evidence? [21, p. 310]

Citation is indispensable in serious mathematical practice, but must be conducted with care. Walton's list of critical questions, although designed primarily for a forensic context, provides a good starting point for an analysis of what comprises such care. Good citation requires a positive answer to all six questions. All of the above examples block a satisfactory answer to one or more of these questions. For most of them, at least the Opinion Question could not be fully answered. (Mutual reference may be an exception, but it certainly fails the Backup Evidence Question, as do several of the others.)

All of these cases exhibit defective argumentation by the prover, whose use of the scheme is exposed as unwarranted by his inability to answer the critical questions. But the prover must also have a respondent, the audience for his proof. Pseudo-proofs can arise from the argumentational shortcomings of the respondent too:

Proof by vehement assertion It is useful to have some kind of authority relation to the audience. [3, p. 17]

Here the prover leads his audience into employing the scheme for Appeal to Expert Opinion, with him as the expert. In effect, he is appealing to his own authority. Something similar may be happening in the next case:

Proof by vigorous handwaving Works well in a classroom or seminar setting. [3, p. 16]

In each of these cases the audience is browbeaten into acquiescence, something they share with an even more common scenario:

Proof by intimidation 'Trivial' [3, p. 16]

In cases where the result is not trivial, and perhaps not correct, this commonplace usage may be understood as a profoundly compressed enthymematic instance of a different scheme, for Argument from Danger.

3.2 Argument from Danger

Premise If you (the respondent) bring about A , then B will occur.

Premise B is a danger to you.

Conclusion Therefore (on balance) you should not bring about A . [21, p. 334]

The argument induced by the dogmatic declaration 'Trivial!' may be unpacked as an instance of this scheme: 'If you ask for clarification, you will appear to be a fool. Appearing a fool is a danger to you. Therefore (on balance) you should not ask.' (Those in no danger of appearing a fool can ask devastatingly simple questions which might, from a humbler source, provoke mere condescension.) However, where the result seems trivial to the audience too, there can be no intimidation in drawing attention to its triviality; rather, this may be understood as based on a different source: (shared) intuition.

4 INTUITION

Like citation, intuition is ubiquitous, but problematic. As Reuben Hersh observes 'If we look at mathematical practice, the intuitive is everywhere. ... The word intuition, as mathematicians use it, carries a heavy load of mystery and ambiguity. Sometimes it's a dangerous illegitimate substitute for rigorous proof. Sometimes it's a flash of insight that tells the happy few what others learn with great effort' [9, p. 61]. Unsurprisingly, this uneasy ambiguity manifests in jokes:

Proof by appeal to intuition Cloud-shaped drawings frequently help here. [3, p. 17]

Proof by seduction 'Convince yourself that this is true!' [17, p. 28]

The second example may be found, almost verbatim, even in so august an author as G. H. Hardy: 'The reader will convince himself of the truth of the following assertion' [7, p. 4].

Intuition can be private, but if it is intended to convince others it must be shared; that is, it must be obvious. Hardy (crediting J. E. Littlewood) offers a partial resolution of the underlying ambiguity of intuitive obviousness:²

When one says 'such and such a theorem is almost obvious' one may mean one or other of two things. One may mean 'it is difficult to doubt the truth of the theorem,' 'the theorem is such as common sense instinctively accepts,' as it accepts, for example, the truth of the propositions ' $2 + 2 = 4$ ' or 'the base angles of an isosceles triangle are equal.' That a theorem is 'obvious' in this sense does not prove that it is true, since the most confident of the intuitive judgments of common sense are often found to be mistaken; and even if the theorem is true, the fact that it is also 'obvious' is no reason for not also proving it, if a proof can be found. The object of mathematics is to prove that certain premises imply certain conclusions; and the fact that the conclusions may be as 'obvious' as the premises never detracts from the necessity, and often not even from the interest of the proof.

But sometimes (as for example here [If $\phi(n)$ and $\psi(n)$ tend to limits a, b , then $\phi(n) + \psi(n)$ tends to the limit $a + b$]) we mean by 'this is almost obvious' something quite different from this. We mean 'a moment's reflection should not only convince the reader of the truth of what is stated, but should also suggest to him the general lines of a rigorous proof.' And often, when a statement is 'obvious' in this sense, one may well omit the proof, not because the proof is unnecessary, but because it is a waste of time to state in detail what the reader can easily supply for himself [7, p. 130].

Thus a claim of intuitive obviousness is a knowledge claim resting on an equivocal basis. Both of the senses Hardy distinguishes may be characterized by the same scheme:

² Hersh also analyses the ambiguity of intuition, identifying six different senses [9, pp. 61 f.]. Although Hersh ranges wider, Hardy is more incisive: he subdivides the third of Hersh's senses.

4.1 Argument from Position to Know

Major Premise Source a is in position to know about things in a certain subject domain S containing proposition A .

Minor Premise a asserts that A is true (false).

Conclusion A is true (false).

Critical Questions:

1. Is a in position to know whether A is true (false)?
2. Is a an honest (trustworthy, reliable) source?
3. Did a assert that A is true (false)? [21, p. 309]

In the second of Hardy's senses, the prover and (so he assumes) the respondent are in a position not only to know but to prove the 'obvious' result, and the critical questions may be easily answered in the affirmative. In the first sense, an adequate characterization of the basis for the prover's knowledge claim is more fugitive. It may be, as Hersh suggests, 'reliable mathematical belief without the slightest dream of being formalized' [9, p. 61]. Since he views this faculty as reliable, Hersh would still answer the critical questions affirmatively. Sceptics about such reliability would demur from one or both of the first two questions.

5 META-ARGUMENT

A prominent feature of mathematics is the reification of arguments as themselves the objects of mathematical inquiry. However, arguments which bear upon the status of other arguments comprise a variety of source-based argumentation which does occur in non-mathematical contexts. Angluin's list contains two cases which exploit this practice in different ways.

Proof by metaproof A method is given to construct the desired proof. The correctness of the method is proved by any of these techniques. [3, p. 17]

Here the source from which the argument at issue arises is itself an argument, the presumed virtues of which would indeed transfer to the object argument, but of course, only if the method is secure.

Proof by funding How could three different government agencies be wrong? [3, p. 16]

The existence of funding, like the eminence of the author's institution, and other 'esteem indicators', might be taken as indirect evidence for the soundness of his results. While this evidence is pretty weak, similar evidence can be persuasive. One empirical study presented samples of undergraduates and research mathematicians with a heuristic argument for the presence of one million consecutive sevens in the decimal expansion of π . Correctly attributing its authorship to Field's Medallist Tim Gowers significantly increased the positive appraisal of the argument in both groups [10]. As one research subject commented, 'We are told the argument is made by a reputable mathematician, so we implicitly assume that he would tell us if he knew of any evidence or convincing arguments to the contrary' [10, p. 42]. This demonstrates the perceived relevance of the author's ethos to the acceptance of his arguments, an application of the following scheme:

5.1 Ethotic Argument

Premise If x is a person of good (bad) moral character, then what x says should be accepted as more plausible (rejected as less plausible).

Premise a is a person of good (bad) moral character.

Conclusion Therefore, what a says should be accepted as more plausible (rejected as less plausible).

Critical Questions:

1. Is a a person of good (bad) moral character?
2. Is character relevant in the dialogue?
3. Is the weight of presumption claimed strongly enough warranted by the evidence given? [21, p. 336]

Walton's version is not quite suitable for our purposes: the ethos of a must encompass epistemic instead of (or as well as) moral virtue (as Walton explores elsewhere: cf. [21, p. 359]). It is not Tim Gowers's blameless personal life which leads us to trust his arguments, but rather his demonstrably high standards as a working mathematician. Notoriously, academics can have deplorable ethical standards, while still producing work of genuine value. This may be because, unlike politicians, say, whose weak morals can prove more problematic, they back up their claims with independently verifiable argumentation. Does this obviate the need for consideration of their ethos, even restricted to its epistemic aspects? Only if the argument is fully worked out: if we are invited to accept aspects of the reasoning on trust, it is not unreasonable to ask whether the reasoner is trustworthy.

Ethotic Argument is closely related to Appeal to Expert Opinion. Indeed, the research subject's comments quoted above could also be construed as positive answers to some of the critical questions for the latter scheme. As the commenter is inferring the absence of counter-arguments from the absence of evidence for them, he might also be taken as employing a different form of source-based argument, an Argument from Ignorance.

6 EPISTEMIC CLOSURE

Arguments from Ignorance are typically classified as fallacious, but they can be convincing in epistemically closed domains. Mathematics, however, is not such a domain, so legitimate mathematical applications of the following scheme are somewhat scarce.

6.1 Argument from Ignorance

Major Premise If A were true, then A would be known to be true.

Minor Premise It is not the case that A is known to be true.

Conclusion Therefore A is not true.

Critical Questions:

1. How far along has the search for evidence progressed?
2. Which side has the burden of proof in the dialogue as a whole? In other words, what is the ultimate *probandum* and who is supposed to prove it?
3. How strong does the proof need to be in order for this party to be successful in fulfilling the burden? [21, p. 327]

This scheme is an instance of *modus tollens*, so its plausibility just turns on the acceptability of its premises. The major premise states the epistemic closure principle, which can be analyzed as a conjunction of two conditionals: (i) If A were true, then A would be known to be true by someone; (ii) If A were known to be true by someone, then A would be known to be true (by me). In much everyday, social knowledge this is unproblematic: (i) is trivially true and (ii) is a straightforward claim about the speaker's position within an epistemic community.

In mathematical cases, (i) is characteristically false: there are few unknown A of which it can be said that if A were true, then A would be known to be true by someone. A notable exception is the use of computer assistance. If a computer is known to have calculated the logical closure of a set of axioms, or determined which of a wide range of cases have a particular property, then an unsuccessful search through the resultant data set is conclusive: the truth of the closure principle permits the deduction of the conclusion. Even when the closure principle is not known with certainty, it may often be practically assumed. As the science writer George Szpiro remarks of *Mathematica*, ‘You can enter a function that you want to integrate, and out comes the correct expression. And if it does not, you may be reasonably sure that a solution to your question simply does not exist’ [18, p. 198].

However, (ii) is often plausible in mathematics. Experts may confidently assert of many propositions in their field, that if a proof were known, it would be known by them. Even non-experts can be confident about some propositions: if a counter-example to Goldbach’s conjecture is found, we’ll all know about it. So I may safely infer from my ignorance that no such number has been found, but not of course that it does not exist. Hence the *modus tollens* stops at (ii), and the safe conclusion is not that A is not true, but that nobody knows whether A is true.

Angluin offers an interesting variant of this argument:

Proof by cosmology The negation of the proposition is unimaginable or meaningless. Popular for proofs of the existence of God. [3, p. 17]

This ‘proof’ employs a weaker closure principle: ‘If $\neg A$ were true, then $\neg A$ would be imaginable.’ The absence of an objective measure of imaginability makes this principle implausible.

Arguments from Ignorance can also be induced in the respondent:

Proof by cumbersome notation Best done with access to at least four alphabets and special symbols.

Proof by exhaustion An issue or two of a journal devoted to your proof is useful.

Proof by obfuscation A long plotless sequence of true and/or meaningless syntactically related statements. [3, p. 16]

These three cases encourage the respondent to pass the buck, rather than tackle a purported proof of conspicuous rebarbateness. The Argument from Ignorance would run ‘If there was something wrong with this proof, someone else would have noticed. Nobody seems to have noticed anything. So, it must be OK.’

7 GENERALIZATION

Much mathematical practice concerns the correct application of established rules. A notable example is the move from particular to general results. However, as with other mathematical activities, generalization lends itself to abuse when applied carelessly:

Proof by example The author gives only the case $n = 2$ and suggests that it contains most of the ideas of the general proof.

Proof by omission ‘The reader may easily supply the details.’
‘The other 253 cases are analogous.’
‘...’ [3, p. 16]

These patterns of reasoning may be found in both historical and contemporary mathematics. John Wallis, for instance, routinely argued from specific examples to the conclusion that a procedure may be

applied indefinitely. He ‘relied boldly and confidently on his really astounding intuition as to the correlation between the sums of different series. He called [this] “modus inductionis”: later it was termed “incomplete induction.”’ [15, p. 38].

Standards of rigour have changed profoundly since Wallis’s day. But instances of proofs by example or omission still arise in ostensibly rigorous contexts, although not always without reproach. Take the dispute over Wu-Yi Hsiang’s alleged proof of Kepler’s conjecture, that the maximum density of a packing of congruent spheres in three dimensions is $\pi/\sqrt{18}$. In his review, Gábor Fejes Tóth complains that ‘we are given arguments such as “the most critical case is...” followed by a statement that “the same method will imply the general case.” The problem with arguments of this kind is not only that they require the reader to redo some pages of calculations, but, notoriously, that they occur at places where we expect difficulties and most frequently it is impossible to see how the same method works in the general case’ [5]. Both cases exemplify the following scheme.

7.1 Argument from Example

Premise In this particular case, the individual a has property F and also property G .

Conclusion Therefore, generally, if x has property F , then it also has property G .

Critical Questions:

1. Is the proposition claimed in the premise in fact true?
2. Does the example cited support the generalization it is supposed to be an instance of?
3. Is the example typical of the kinds of cases the generalization covers?
4. How strong is the generalization?
5. Do special circumstances of the example impair its generalizability? [21, p. 314]

This scheme underpins incomplete or enumerative induction. This is, of course, formally invalid, but still has a place in informal mathematical reasoning, for example in hypothesis formation: a is F and also G ; I can’t think of anything which is F and not G ; Are all F s G ? The concern with enumerative induction is that it may be mistaken for proof. Some empirical research is worrisome: one study found that 80% of trainee elementary mathematics teachers considered arguments based on specific instances to be mathematical proofs [13].

But enumerative induction is not the only application of this scheme. While Wallis might be described as employing induction, Hsiang is attempting something else, but is still employing Argument from Example. His generalizing moves are contentious because his examples are said to provide insufficient support, and to be poorly chosen, that is to fail the second and third critical questions, not because generalization is in principle non-rigorous. Where the examples are genuinely typical, and the method transparent, application of this scheme is consistent with mathematical rigour. As Jamie Tappenden observes, ‘generalizing is much more than just picking constants and replacing them with variables. Articulating the right structures, which then can be generalized, is an incredibly involved process and it is hard to get it right’ [19, pp. 264 f.].

A special case of generalization arises with diagrammatic proof. Angluin’s list contains one example that articulates the conservative attitude:

Proof by picture A more convincing form of proof by example. Combines well with proof by omission. [3, p. 17]

Defenders of the legitimacy of diagrammatic proof acknowledge the centrality of the ‘Generalization Problem’: ‘After we have proved a theorem using a diagram, how can we legitimately generalize the configuration of this diagram to a wide (usually infinite) class of configurations, and to what class exactly?’ [11, p. 81]. However, they point out that similar issues arise in non-diagrammatic inference, and may be resolved in a similar fashion in the diagrammatic case, by the painstaking assurance that the particular case really does generalize, that is by answering the critical questions of the scheme for Argument from Example.

8 DEFINITION

A less contentious application of a rule is the employment of a definition. Well-chosen definitions can be pivotal to mathematical progress. At a superficial level, this choice can appear arbitrary: mathematicians have more flexibility in the concepts they take as primitive than scientists in any other domain. On the other hand, appeals to the naturalness or otherwise of definitions are frequent, and a change of definitions can revolutionize the understanding of a hitherto intractable problem. Of course, this strategy can be abused.

Proof by semantic shift Some of the standard but inconvenient definitions are changed for the statement of the result. [3, p. 17]

Instead of defining the same problem with different terms, a different (typically much easier) problem is described by the same terms. Such sharp practice is a straightforward misapplication of the scheme for Argument from Definition.

8.1 Argument from Definition to Verbal Classification

Individual Premise a fits definition D .

Classification Premise For all x , if x has property D , then x can be classified as having property G .

Conclusion a has property G .

Critical Questions:

1. What evidence is there that D is an adequate definition, in light of other possible alternative definitions that might exclude a ’s having G ?
2. Is the verbal classification in the classification premise based merely on a stipulative or biased definition that is subject to doubt? [21, p. 319]

If the non-standard definition D is poorly motivated, the example will fail to answer the first question; if the non-standard definition is smuggled past a classification premise only known to hold for the standard definition, the example would (also) fail to answer the second question.

9 CONCLUSION & FUTURE WORK

We have seen how the anxieties of practicing mathematicians, implicit within their folk humour, draw attention to those aspects of their informal practice where their argumentation carries the greatest risk. We have also seen the utility of a diverse selection of argumentation schemes, developed for non-mathematical applications, in the analysis of a wide-ranging collection of spurious varieties of mathematical proof. This may be read as further corroboration that ‘mathematical reasoning is already in accord with principles and techniques

from informal logic—even if this is unnoticed by the practitioners’ [4, p. 150].

The argumentation scheme methodology has ancient roots, but is still a work in progress. In recent years, it has attracted much attention from researchers in artificial intelligence. In particular, schemes have been implemented within systems of defeasible argumentation, and as strategies in dialogue games [21, pp. 370; 383]. The application of schemes to mathematical argumentation is still in its infancy, but the comparative maturity of the methodology in other domains offers an inviting prospect for future research.

REFERENCES

- [1] Andrew Aberdein, ‘Fallacies in mathematics’, *Proceedings of the British Society for Research into Learning Mathematics*, **27**(3), 1–6, (2007).
- [2] Andrew Aberdein, ‘Observations on sick mathematics’, in *Philosophical Perspectives on Mathematical Practice*, eds., Bart Van Kerkhove, Jean Paul Van Bendegem, and Jonas De Vuyst, 269–300, College Publications, London, (2010).
- [3] Dana Angluin, ‘How to prove it’, *ACM SIGACT News*, **15**(1), 16–17, (1983).
- [4] Ian J. Dove, ‘Towards a theory of mathematical argument’, *Foundations of Science*, **14**(1–2), 137–152, (2009).
- [5] Gábor Fejes Tóth, ‘Review of Hsiang, Wu-Yi, ‘On the sphere packing problem and the proof of Kepler’s conjecture.’ *Internat. J. Math.* 4 (1993), no. 5, 739–831’, *Mathematical Reviews*, **95g**(52032), (1995).
- [6] Sigmund Freud, *Jokes and their Relation to the Unconscious*, Penguin, London, [1905] 1976.
- [7] G. H. Hardy, *A Course of Pure Mathematics*, Cambridge University Press, Cambridge, tenth edn., 1967.
- [8] Thomas Heath, *The Thirteen Books of Euclid’s Elements*, Barnes & Noble, New York, NY, 2006.
- [9] Reuben Hersh, *What is Mathematics, Really?*, Jonathan Cape, London, 1997.
- [10] Matthew Inglis and Juan Pablo Mejía-Ramos, ‘The effect of authority on the persuasiveness of mathematical arguments’, *Cognition and Instruction*, **27**(1), 25–50, (2009).
- [11] Zenon Kulpa, ‘Main problems of diagrammatic reasoning. Part I: The generalization problem’, *Foundations of Science*, **14**(1–2), 75–96, (2009).
- [12] Gershon Legman, *Rationale of the Dirty Joke: An Analysis of Sexual Humor, First Series*, Simon & Schuster, New York, NY, 1968.
- [13] W. Gary Martin and Guershon Harel, ‘Proof frames of preservice elementary teachers’, *Journal for Research in Mathematics Education*, **20**(1), 41–51, (1989).
- [14] Robert K. Merton, *Social Theory and Social Structure*, Free Press, New York, NY, 1968.
- [15] Kirsti Møller Pedersen, ‘Techniques of the calculus, 1630–1660’, in *From the Calculus to Set Theory, 1630–1910: An Introductory History*, ed., Ivor Grattan-Guinness, 10–48, Princeton University Press, Princeton, NJ, ([1980] 2000).
- [16] George Pólya, *Mathematics and Plausible Reasoning*, Two Volumes, Princeton University Press, Princeton, N.J., 1954.
- [17] Paul Renteln and Alan Dundes, ‘Foolproof: A sampling of mathematical folk humor’, *Notices of the AMS*, **52**(1), 24–34, (2005).
- [18] George G. Szpiro, *Kepler’s Conjecture: How Some of the Greatest Minds in History Helped Solve One of the Oldest Math Problems in the World*, John Wiley & Sons, Hoboken, NJ, 2003.
- [19] Jamie Tappenden, ‘Mathematical concepts and definitions’, in *The Philosophy of Mathematical Practice*, ed., Paolo Mancosu, 256–275, Oxford University Press, Oxford, (2008).
- [20] Bart Van Kerkhove, ‘Aspects of informal mathematics’, in *Essays on the Foundations of Mathematics and Logic*, ed., G. Sica, 268–351, Polimetrica International Scientific Publisher, Monza, (2005).
- [21] Douglas Walton, Chris Reed, and Fabrizio Macagno, *Argumentation Schemes*, Cambridge University Press, Cambridge, 2008.

Interpreting Naproche – An algorithmic approach to the derivation-indicator view

Merlin Carl¹, Peter Koepke¹

Abstract. In [1], Jody Azzouni proposes a derivation indicator view of mathematical practice and in particular of proofs.

The Naproche Project [5] takes proofs as linguistic entities whose semantics is given by corresponding formal derivations. This view is computationally implemented in the Naproche System where simple natural proof texts are automatically transformed into derivations, using techniques from computational linguistics, formal logic and automatic theorem proving. The development of the system has identified many ways in which parts of proof texts indicate elements of formal derivations. The Naproche Project can thus be seen as a support of the derivation indicator view, and we comment on some of the critique against derivation indication from this standpoint and our experience. Doing so, we extend earlier work by the second author on the relation between natural and formal proofs; since then, further development has shown the relation of DI to Naproche to be an interesting field that can (and should) be elaborated in detail.

1 THE NAPROCHE PROJECT

Naproche is a common project of mathematical logicians from Bonn, formal linguists from Duisburg/Essen and computer scientists from Cologne. Its name is derived from "NAtural language PROOf CHEcking" and it aims at studying actual presentations of mathematical proofs from a logical and linguistic point of view. Hence, we are for example interested in questions concerning the underlying inference rules of natural proving, the role of background knowledge in understanding mathematical discourse and linguistic specialities of actual proof texts from mathematical practice.

The Naproche System is both an application and a way of empirical verification of results of our investigations. In short, it is a computer system where a mathematical text - consisting of axioms, definitions, lemmas, theorems and proofs, in a rich but restricted fragment of English including formulas - can be entered in a TeXMacS environment for automatic proof-check.

The first step in checking e.g. a proof found in a textbook thus consists in a reformulation of the given mathematical text in the Naproche language. A short example of what such a reformulation looks like is found below; the reader interested in more material and a description of the current Naproche language is invited to consider [5]. The Naproche language is (being) developed to resemble the mathematical vernacular as close as possible. Therefore, such reformulations are usually easy to do for the kind of texts we are considering at the moment. In particular, this can be done in a sentence-by-

sentence fashion, without a preliminary complete understanding of the proof to be reformulated.

The processing is then continued by applying adapted techniques from formal linguistics, in particular the discourse representation theory of Kamp and Reyle [8], to generate a semantic representation in a nested box structure that we call "proof representation structures" (see [9] for details). To do so, the text is seen as a series of triggers for mental acts such as introducing or retracting an assumption, drawing a conclusion, introducing and naming an object, applying a proof technique (e.g. induction), recalling an earlier result, making a statement, announcing a proof, declaring it finished etc. Linguistic markers for these 'acts' are automatically recognized and trigger an appropriate processing. The obtained intermediate format is then used to deal with several linguistic and logical phenomena such as anaphorical constructions or the identification of the set of available assumptions at each part of the proof. From here, a rather routine algorithm computes a series of queries of the form 'deduce statement A from the set S of available assumptions and results already proved' in TPTP, a standard input format of automatic theorem provers. These are then used to check whether or not the claimed step can be carried out - resembling a human reader trying to fill in missing steps. As a by-product of this checking, one gets a formalized version of the text by simply concatenating the theorem prover's outputs.

Currently, there are various sample texts in the Naproche syntax available, dealing with group theory, logic, set theory and analysis. The most advanced is [6], an English reformulation of the first chapter of Edmund Landaus elementary textbook "Grundlagen der Analysis" [9].

For an impression of what a text in the current Naproche language looks like, consider the following short excerpt from the cited translation, Landaus 6th definition and 28th theorem, which is accepted by the latest version of the system:

Definition 6: Define $\star$ recursively:

$x \star 1 = x.$

$x \star y' = (x \star y) + x.$

Lemma 28a: For all y , $1 \star y = y.$

Proof: By definition 6, $1 \star 1 = 1$. Suppose $1 \star y = y$. Then by definition 6, $1 \star y' = (1 \star y) + 1 = y + 1 = y'$. Thus by induction, for all y $1 \star y = y$. Qed.

This fragment contains already several features such as the possibility of a recursive definition of a new operator, reference to a definition and the induction technique. Also, it is quite close

¹Mathematical Institute of the University of Bonn

to the German text; it wouldn't look conspicuous if merely presented as an English version of the original.

Our system has several potential applications. People interested in this subject may appreciate it as a natural language interface to formal mathematics. Authors of introductory textbook notes could use it as a tool for authoring and checking. For first-year students, it could serve (and has served) as an automatic tutoring system for learning how to prove. And finally, the generation of proof texts that are readable for computers and humans alike is of course a contribution to the field of men-machine-communication.

To avoid misunderstandings, let us remark on what Naproche is not: It is not – and is not meant to be a tool for understanding research papers in any advanced area of mathematics. The idea of e.g. replacing the process of peer review by automatic checkers is pure science fiction. Also, there are and probably will always be figures in the vernacular that exceed the abilities of the system. Therefore, some preparatory work from the user is required. Checking a completely unadjusted text will usually not work.

The next goal is a check of an English reformulation of the whole of Landau's "Grundlagen der Analysis". To achieve this, the inventory of proof figures, logical features, grammar and vocabulary is constantly being extended. Furthermore, we plan broader text corpus studies as a starting point for further empirically solid research on mathematical reasoning and speech practice.

2 DERIVATION INDICATION

Another interesting aspect is in how far Naproche can help to understand the meaning of "understanding" in the context of proofs. In [11], Jeremy Avigad proposes a notion of understanding that refers to abilities that someone with understanding should have. The points he lists include the ability to formalize, fill in gaps, verify correctness and spot mistakes. In this context, one might state that Naproche is in a weak sense capable of understanding elementary mathematical texts.

In [1], Jody Azzouni looks for an explanation for what he calls the "benign fixation of mathematical practice", i.e. the fact that mathematics, when seen as a social practice, is remarkably stable. This particularly applies to questions concerning what counts as a proof. His claim is that mathematical proofs are "indicators" of formal derivations, symbolic sequences of some logical calculus that can be checked mechanically for correctness and thus cannot be meaningfully disputed any more. Hence, derivations serve as background certificates for the discursive stability of a natural proof.

This view bears an obvious analogy with the way computer scientists deal with algorithms that one might label "Turing program indicator view": Usually, "algorithms" are identified with programs for a Turing machine (or one of many equivalent formal notions, depending on the taste of the author). However, actual Turing programs virtually never appear in any publication containing algorithms. Instead, areas as combinatorial optimization or algorithmic graph theory use natural language descriptions of recipes to indicate Turing programs. The way

from one to the other is far from trivial: Quite often an actual implementation remains a challenging problem. Much more is merely "described" rather than "implemented".

Thus, if Turing programs correspond to formal proofs, and informal process descriptions to natural proofs, the situations resemble each other closely and the well-accepted "Turing program indicator view" becomes the algorithmic counterpart to the DI-view.

In recursion theory, where algorithms become objects of consideration of other algorithms more regularly than mathematical proofs become objects of mathematical proofs, there is a common practice called "proof by Church-Turing-Thesis": Given an informal high-level description of a procedure, the existence of a Turing program for the same purpose is taken for granted. This step is all-important, even though one is rarely interested in carrying it out in detail: Probably nobody wants to see e.g. a Turing program enumerating two incompatible Turing degrees. Nevertheless, a Turing program is a good representation for what underlies our certainty that the algorithm can be carried out in all details, i.e. made unambiguous.

While the DI-view is an inspiring perspective in our opinion, Azzouni does not explain the precise meaning of what he means by "indication", nor does he give a hint how it works. In particular, it remains to explain how mathematical practice could be fixed by derivation indication hundreds of years before formal derivations were introduced in the beginning of the 20th century by Frege, Russell and others.

In order for the derivation indicator explanation to work, there must be a clear sense for the process of "indication" - proof practice would not be fixed by referring to derivations if there was no generally accepted way how to go from a natural proof to a derivation. But if there is such a procedure, however complicated, then by the Church-Turing thesis, it is algorithmically implementable. It is these algorithms that we are looking for.

Consider again the excerpt from our reformulation of Landau cited above. The indicators, or, as we call them, triggers, are easy to identify: "Definition 6", for example, announces the introduction of a new notion which is given the number 6 for further reference. In the proof, "suppose" indicates the introduction of an assumption, "then" prepares that the following is deducible from the available assumptions, "by induction" gives a hint on the proof procedure to be used, while "qed" marks the end of the proof and hence the claim that at this point, the proof goal should follow from what has been said so far.

Our expectation is that the language of more advanced textbooks than [10], though more elaborate and complicated, can be handled in a similar manner. This is supported by the view many mathematicians take on proofs and their relation to derivations: Derivations are seen as representations of discursive limit objects that would arise if every step of a proof was carried out in full detail.

Let's have a look at the following quote by Saunders MacLane [12]:

"As to precision, we have now stated an absolute standard of rigor: A mathematical proof is rigorous when it is (or could be)

written out in the first-order predicate language $L(\epsilon)$ as a sequence of inferences from the axioms ZFC, each inference made according to one of the stated rules. [...] When a proof is in doubt, its repair is usually a partial approximation to the fully formal version."

This gives a hint to view the apparent difference between proofs and derivations as a matter of granularity. The increasing degree of detail can be observed through the stages of mathematical research: From high-level discussions about proof plans to written sketches, research papers, textbook proofs, explanations thereof in lectures, elaborations thereof in tutorial classes over almost gap-free proofs as in Landau to derivations of the kind Naproche produces.

3 CRITICISM

Since its first publication, several points of criticism have been raised against the DI-view, see e.g. [3] and [4]. Based on our experiences with analyzing proofs and the Naproche systems, we feel that we can substantially react to some of them.

In [3], Yehuda Rav claims that proofs contain an "irreducible semantic component", by which he means "intentional objects" that "cannot be formalized but require the active interpretative participation of the reader". Taken for granted that understanding is an active process that needs to be carried out, we don't agree that this can only be done by humans. Even the current Naproche system generates an interpretation of a text through an active process involving language processing and automatic reasoning - one might say that the system questions each proof step, then tries to find reasons for it. Nevertheless, Naproche is an algorithmic system in any sense of the word - one which indeed does "the work of mechanically recognizing the validity of even an informal proof" ([3], p. 304).

One might object here that the Naproche language, in spite of resembling the mathematical vernacular, is still by definition itself a formal language and that the actual formalization is therefore carried out by the reader who reformulates the text, not by the system.

We think that even the current stage of Naproche gives a strong hint that the frontier between formal and natural languages is rather a question of degree. If the 'formal languages' reach a degree of resemblance to the vernacular that makes formalization a question of basic changes of words and formulations, much more of the work of generating an interpretation is done by the machine than by the human reader involved – in this sense, the human reader can be eliminated from the process of understanding.

Later on, Rav refers to the formalization of Landau in the Automath language. He claims, citing Feferman, that the fact the checking of the resulting derivation was successful is no argument in favour of the correctness of the original text, since a human reader was needed to understand the text prior to its formalization, and "understanding subsumes checking". But where the reading process can be replaced by a mere routine reformulation – which, based on our experience, appears to be

the case for considerable parts of basic mathematical literature - this problem is at least heavily reduced.

While a complete formalization of a proof may indeed require a deep understanding of the original material, possibly even a deeper one than a usual reader will obtain, a reformulation of e.g. Landau's proofs in the Naproche syntax requires much less. It is in fact closer to a translation of a proof into a different language than to a formalization in, say, predicate calculus.

This does not presuppose understanding to a degree that 'subsumes checking': It is well possible (and has probably already happened several times) that A, a native German speaker without knowledge of foreign languages, writes a text that he considers to be a proof. B, who knows German and English reads it, considers it correct and, translates it to English, and C, who knows English, but not German, spots a mistake in the translated text. The translation done by B cannot have subsumed checking, otherwise B would have spotted the mistake. So translation in another natural language does not necessarily subsume checking.

The question is then whether semi-formal languages can serve the same purpose at least for some tasks. Based on the first author's personal experience in writing Naproche texts, even the basic language of Naproche is already good enough for this. The translation of, say, the first chapter of Landau, was (on purpose) mostly done without an actual thought about the content of the original text, let alone a checking. If there was a mistake in this chapter, it could well have been discovered by Naproche though it evaded the person adapting the text.

It is also said that formalizations, even where possible, do not contain new information, nor do they enhance the rigor or security of an argument that can easily be seen to be correct by understanding it. Let us, however, consider again our main sample text [10]. It contains nothing but fairly elementary theorems in a fairly elementary area, and its author, Edmund Landau, was an accomplished mathematician of international rank and fame. Yet, in the introduction, he mentions that a colleague spotted a flaw in his supposedly clear and complete presentation that lead him to reorganize some of the material. It therefore appears that there is space for an 'epistemic gain' even in the most basic examples of informal standard mathematics. Attempts to formalize an argument are a possible way to obtain a higher degree of certitude. Reformulating texts in a semi-formal language like that of Naproche is one way to make this gain cheaper.

Other points referring to the impreciseness of terms like "algorithmic system", "indication" etc. are, in our view, no counterarguments but open research fields. Furthermore, it should be noted that these arguments could also be raised against "algorithm indication", which nevertheless does not seem to cause any problems.

4 FUTURE PROSPECTS AND RELATED WORK

There are also related projects going on at the moment.

While, to the authors' knowledge, there is no other group that focuses on the connection between informal and formal proofs, there are already several proof checking systems that also emphasize the readability of their respective input language.

In the Mizar system, started in 1973 by Trybulec, theorems of advanced mathematics are formulated and proved by an active community, and results are published in the journal *Formalized Mathematics*.

The developing team of Isabelle is working on a 'human-readable structured proof language' named Isar.

We are collaborating with the VeriMathDoc project, which includes the PLATO program. Their goal is to develop a mathematical assistant system that naturally supports mathematicians in the development, authoring and publication of mathematical work.

In his dissertation, Claus Zinn developed a system for processing mathematical texts [17]. In his work, he also introduces extended DRSSs, called Proof Representation Structures. Despite the equality of their names, the PRS notions of Naproche and the work of Zinn are quite different.

The Naproche project itself is still at an early stage of development. Even though it already reaches its goal to mediate between easy natural proofs such as the excerpt from Landau cited above and fully formalized, mechanically checkable derivations, a lot is still to be done before one can approach more sophisticated textbooks. On the linguistic side, grammar and vocabulary are permanently expanded while more and more linguistic phenomena like resolution of anaphora or presuppositions are taken into account. Corpus studies are planned and carried out to get information about common proof strategies, argumentative figures and how they are indicated. Also, we attempt to make the current CNL closer to the actual natural language; as an example, the current way to trigger the retraction of an assumption by 'thus' is quite unsatisfying. In reading mathematical texts, one encounters virtually never a doubt on what the open assumptions are without such artificial rules.

How this works is one big question for the time being.

On the logical side, one easily notices that implicit background knowledge such as spatial imagination or intuitive type systems play a large role in understanding mathematical discourses. For example, human readers easily deal with objects such as sets, sequences, lists etc. without using explicit knowledge on concatenations, permutations or set theory. Our goal is to model this by providing the system with a weak set theory. Here, it is quite a challenge to find the right balance between premises that are on the one hand strong enough to allow usual references to sets to be processed, and harmless enough to keep the search space in control. (E.g., one might try to use ZFC as the background theory, which is certainly strong enough. But then the prover would have to deal with infinitely many axioms in each step, which does not look very promising.)

Talking about complexity issues, these already are a permanent threat to incautious expansions of the system. When gaps in natural proofs become too big to be filled in by an ATP in an

acceptable amount of time, the proof check is unsuccessful. One way to approach this difficulty is to exploit implicit information and hints on how to proceed given in the proof more efficiently. Still, if the search space consists of all accessible theorems, checking of longer texts might soon become unfeasible. This can be avoided by carefully selecting the premises that are likely to be useful for the current step. The selecting algorithms are still at an experimental stage, yet even these result in a enormous reduction of running times.

We are optimistic that these developments will allow us to check much more and more advanced texts. For the near future, we plan a mechanical check of further chapters of Landau's 'Grundlagen der Analysis', including those where set theory becomes unavoidable, and of Euklids 'Elements'.

5 CONCLUSIONS

We hope to have made plausible that the DI-view on mathematical practice can lead to interesting research and concrete, useful technical applications. In our opinion, this view is rather an inspiring credo, or a "regulative imperative", than a complete explanation of the stunning stability of mathematical practice. In the long run, the Naproche project could be a way of empirically sustaining this perspective. It is our conviction that enriched formalisms in the spirit of the Naproche language can provide a bridge between what are called the "two streams in the philosophy of mathematics" and lead to a notion of proof that incorporates the advantages of both.

REFERENCES

- [1] Azzouni, Jody (2006): *Tracking Reason* (Oxford University Press)
- [2] Azzouni, Jody (2004): The derivation-indicator view of mathematical practice (*Philosophia Mathematica* (3) 12, 81-105)
- [3] Rav, Yehuda (2007): A Critique of a Formalist-Mechanist Version of the Justification of Arguments in Mathematician's Proof Practices (*Philosophia Mathematica* (III) 15, 291-320)
- [4] Carlo Cellucci (2008): *Why proof? What is a proof? (Deduction, Computation, Experiment. Exploring the Effectiveness of Proof,* Springer Berlin)
- [5] The Naproche homepage: www.naproche.net
- [6] Carl M., Cramer M., Kühlwein D.: Landau in Naproche-Syntax (available at [5])
- [7] Koepke, Peter: *Mathematical Proofs as Derivation-Indicators: Theory and Implementation* (Talk given in Utrecht 2009)
- [8] Kamp/Reyle (1993): *From discourse to logic* (Dordrecht: Kluwer Academic)
- [9] Nikolay Kolev (2008): *Generating Proof Representation Structures in the Project NAPROCHE* (available at [5])
- [10] Landau, Edmund (1965): *Die Grundlagen der Analysis* (Wissenschaftliche Buchgesellschaft Darmstadt)
- [11] Avigad, Jeremy (2008): *Understanding Proofs* (in: *The Philosophy of Mathematical Practice*, Oxford University Press)
- [12] MacLane, Saunders (1986): *Mathematics, Form and Function* (Springer)

Applying the GC Combined Reasoning Framework to Mathematical Discovery

John Charnley¹

Abstract.

In previous work, we have demonstrated how the GC combined reasoning framework has been used for mathematical discovery. However, the current approach has several limitations, for example a single basic representation scheme. By reference to a mathematics text book, we highlight some challenging aspects of mathematical discovery and comment on how characteristics of the GC approach provide useful solutions. In particular, the distributed nature of the framework would allow multiple representation schemes to be considered, which is useful in human studies of mathematics and allows us to introduce more advanced reasoning systems. We believe that, by implementing these solutions, we can create a highly effective mathematical discovery package.

1 INTRODUCTION

Automatic mathematical discovery has long been a goal of AI researchers and several systems have been developed, for example Graffiti [5] and HR [3]. In general, these systems aim to create theories that would be both interesting to, and verifiable by, human mathematicians. The creation of a mathematical theory involves developing new mathematical concepts, identifying relationships between them and discovering and proving theorems. Consequently, it encompasses numerous specific skills such as symbolic manipulation, theorem proving, example generation etc. As a result, several previous systems, such as HR, employ a combination of automated reasoning techniques. In previous work, we developed the GC framework for combined reasoning, which we describe in §2. The GC framework, which takes its name from the initial letters of Global Workspace Theory and combined reasoning, allows users to create effective systems which employ multiple forms of automated reasoning in performing a particular task [2].

We have already used GC to create a discovery system, GCATF, which takes HR as its inspiration, and we have applied it to a number of mathematical domains. However, GCATF has several limitations and we believe there is a great deal of scope to extend its capabilities. In particular, the GC approach has characteristics which we believe will be of benefit in mathematical discovery tasks. The most obvious limitation is the system's representation scheme. Mathematicians use a variety of visualisations and representations for mathematical notions. Like the underlying ideas they represent, they vary in their complexity. Furthermore, the representations are dynamic, having developed hand-in-hand with the subject of mathematics itself. Most discovery systems use a limited representation scheme. In the case of GCATF, this is Prolog-style first-order logic. Whilst useful for many

domains, this single static representation places an artificial ceiling on the complexity of mathematical notions that can be considered. It is also a barrier to leveraging more advanced automated reasoning systems.

In order to better understand this issue and other deficiencies of GCATF, we found it useful to go back to basics. So, we reviewed a typical mathematical textbook in the area of mathematical analysis [9]. In only the first few pages, we see several instances where a more complex approach is needed. In §3, we describe some of these and comment on how GC offers potential solutions.

2 The GC Framework

In cognitive science, Global Workspace Theory (GWT) has been proposed to describe the operation of mammalian consciousness [1]. At its heart lies the Global Workspace Architecture (GWA), which describes an arrangement where a number of parallel specialist processes are attached to a global workspace. In each round of processing, some information is broadcast on the workspace. In response, the attached processes can, if they wish, propose some new information for broadcast. The information proposed by the process making the 'strongest' proposal is broadcast in the next round of processing. The GC framework allows developers to create distributed systems according to GWT. In the GC setting, each of the processes performs some form of reasoning and the contents of the workspace are reasoning artefacts. In addition, the GWT is modified slightly to allow processes to spawn and attach new processes during processing. The GC toolkit provides software allowing users to develop processes, bring them together into combined systems and run them in either a serial or distributed manner. To do this, a user defines the processes to be included and the reasoning artefacts that may be broadcast. Overall control is implicit in how the processes weight their suggestions, so the user must specify an appropriate weighting system. The framework is extremely light-weight and GC provides only the skeleton with which to develop and run systems. An effective system requires the user to implement specialist reasoning processes, for example by creating wrappers for external stand-alone reasoning systems. Alternatively, they could re-use or adapt existing processes from previous GC systems.

GC offers several benefits for creating combined reasoning systems. The use of specialist processes and a simple communication scheme means that GC systems are modular, clear and extensible. In addition, the architecture addresses issues of relevancy by neatly distributing the choice of which action or tool to apply. Importantly, the architecture imposes no specific semantics or representation scheme to be followed. The GC framework has been shown as an effective basis for many combined reasoning applications [2]. In particular it

¹ Imperial College, London, email: john.chnley@imperial.ac.uk

lends itself to free-form discovery tasks, where specific sub-systems contribute to a task when they see themselves to be relevant.

Several other systems that have been developed with similar goals to those of the GC project are described in [2]. For instance, Sorge's Ω -ANTS framework [10] and the MATHWEB architecture [7]. These systems also aim to facilitate the inter-operation of heterogeneous distributed reasoning systems in mathematics. However, these systems have mainly been designed for a particular purpose, such as automated theorem proving, and their suitability for more free-form discovery tasks has not been established.

3 Textbook Mathematics

Mathematics text books offer one way of understanding the discovery process from a human perspective. We read the opening chapters of an elementary mathematics text book [9], which is a first year degree-level introduction to analysis. It is interesting in that it aims to build the infinitesimal calculus from the basic axioms of real numbers. As such, its structure is similar to how we would want our automated discovery system to create theorems from first principles. Although this is a very basic starting point, the first few pages of this book highlight a number of key considerations, including:

- **Representation:** the book begins with some initial mathematical concepts which are considered known to the reader. For example, the concept of a real number is not formally defined but taken from the reader's knowledge of continuous measurements, such as mass. A common human understanding of real numbers is a mental visualisation as a line, together with certain allowed manipulations and expected behaviours. Using analogies in this way is key aspect of human mathematics [8]. In addition, there are also symbolic representations for concepts. Indeed, there are often several visualisations or symbolic representations. For example, $sign(x)$ can be defined as $-1 \leftrightarrow x < 0$ and $1 \leftrightarrow x \geq 0$, alternatively it can be seen as $x/|x|$, both of which are useful. The ability to move between different representations of the same concept is fundamental in mathematical discovery, for instance in applying varied proving techniques. Consequently, an effective mathematical discovery system should have access to multiple representations.

GC provides one answer to this problem by allowing different conceptual interpretations to be distributed. We refer to this as *distributed representation*. At the simple level, this involves using sub-systems with different formalisms and making the interpretation task local to those sub-systems. At a more complex and challenging level, we envisage leveraging the parallel nature of GC to consider all possible representations for a given concept in a given context as required. The GC framework has its origin in cognitive science, which may provide insights into how humans employ different schemes. For instance, we may be able to mimic the way people identify loose similarities between different ideas, leading to the development of concrete relationships.

- **Concept Formation:** the book introduces new concepts in several ways. Many of the mechanisms used are similar to those we have considered in previous studies. For example, concepts are introduced as being particular properties of existing concepts e.g. $|x|$. However, some methods go beyond the scope of our existing methods, for example concepts which involve infinity, e.g. infinite sequences and series. In addition, different representations require their own specific concept formation methodologies, e.g. symbolic transformation or diagrammatic manipulation. Consider,

for instance, the development of the concept of a plane as two perpendicular real lines versus its symbolic analogy of bringing two real numbers into a tuple, (x, y) . Again, this is challenging, especially when trying to automatically ensure consistency between representations. It is unlikely that analogues of all concepts could be maintained in every representation scheme at all stages. This would require the system to create, for each new concept generated in one representation, a semantically equivalent concept in all other representations, which may not be possible. Perhaps a more achievable approach would be to allow concept formation to continue freely in multiple representation schemes. The identification of analogues would then become a separate task, which might sporadically identify correspondences.

One benefit of the GC approach is the ease with which it is possible to add new functionality. This feature should allow us to incrementally boost our concept formation capabilities as we introduce new formation methods or develop approaches to maintaining consistency of representation.

- **Investigative Direction:** the book progresses by introducing new interesting concepts and investigating them in some depth, with new theorems and exercises. In our system, concept formation drives the direction of the investigation. However, in a free-form discovery task, often a large body of uninteresting information obscures the important discoveries and focus can be lost. Some work has already been done on assessing interestingness of generated concepts, for example by using feature analysis to rank them [4].

GC's weighting system provides a mechanism by which we might introduce interestingness and control the direction of investigation. It would also be useful to make this dynamic by introducing a retrospective review of how important a particular concept becomes as the discovery process progresses.

- **Symbolic Manipulation and Proving:** this key aspect of mathematical study is, fortunately, a very successful area of AI and has spawned many different systems and approaches, including computer algebra systems and provers for many types of logics. To date we have only used a tiny fraction of the available systems and there is vast scope to leverage advances in the field. This is particularly relevant if we are to use multiple representations. For instance, the first exercises in the textbook ask for a proofs of the irrationality of $\sqrt{2}$ and $\sqrt{3}$. We were able to prove both of these, the first using purely symbolic manipulation but we required an element of diagrammatic reasoning to convince ourselves of the second. There are existing systems for each of these approaches which we would seek to use, for example the DIAMOND diagrammatic reasoning program [6].

The GC approach engenders extensibility, meaning that introducing new tools is straightforward. However, one problem is that it is not always possible to leverage the full power of external reasoning systems as they often rely upon some form of human intervention to achieve their full potential. Adding new systems hopefully gives us an opportunity to feed back into the development of external reasoning tools by giving an assessment of their usability by non-expert users and providing a common test platform.

4 SUMMARY

GC has already been shown to be a useful tool for mathematical discovery but there is a great potential for improvement. Several aspects of the GC approach suggest it would be able to deal with specific challenges that mathematical discovery presents. In particular, we believe the way in which GC distributes relevancy allows us to introduce multiple representation schemes. The extensibility of GC systems means it is relatively straightforward to introduce new functionality, such as advanced concept formation methods or more powerful reasoning tools. Moreover, the weighting scheme provides a way in which we might introduce notions of interestingness by which to guide the discovery process. We believe that GC represents a promising basis for a mathematical discovery package which would provide an opportunity to bring together disparate reasoning systems and contribute to their development.

ACKNOWLEDGEMENTS

REFERENCES

- [1] B Baars, *In the theater of consciousness: The workspace of the mind*, Oxford University Press, 1997.
- [2] J Charnley, *A Global Workspace Framework for Combined Reasoning*, Ph.D. dissertation, Imperial College, London, 2010.
- [3] S Colton, *Automated Theory Formation in Pure Mathematics*, Springer-Verlag, 2002.
- [4] S Colton, A Bundy, and T Walsh, ‘On the notion of interestingness in automated mathematical discovery’, *International Journal of Human Computer Studies*, **53**(3), 351–375, (2000).
- [5] S Fajtlowicz, ‘On conjectures of graffiti’, *Discrete Mathematics*, **72**, 23,113–118, (1988).
- [6] Mateja Jamnik, *On Automating Diagrammatic Proofs of Arithmetic Arguments*, Ph.D. dissertation, Division of Informatics, University of Edinburgh, 1999.
- [7] M Kohlhase and J Zimmer, ‘System description: The MathWeb software bus for distributed mathematical reasoning’, in *Proceedings of the Conference on Automated Deduction (CADE)*, (2002).
- [8] George Lakoff and Rafael E. Núñez, *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*, Basic Books, 2000.
- [9] John B. Reade, *An Introduction to Mathematical Analysis*, Oxford Science Publications, 1986.
- [10] V Sorge, *A Blackboard Architecture for the Integration of Reasoning Techniques into Proof Planning*, Ph.D. dissertation, Universität des Saarlandes, Saarbrücken, 2001.

Origins of Spatial-Numerical Bias

Martin H. Fischer¹ and Peter Brugger²

Abstract. Spatial-numerical associations are prevalent in numerical cognition yet their origin is poorly understood. We discuss several candidate accounts and conclude that reading is a secondary factor while finger counting, natural number statistics, and embodied magnitude processing are likely to be of primary importance for the origin SNARC.

1 INTRODUCTION

Healthy Western adults tend to associate small numbers with left space and larger numbers with right space. This association affects their entire behavioral repertoire, from response choices to key pressing, arm and eye movements, hand apertures, and attention allocation. Over 100 experiments have now studied this spatial-numerical association of response codes or SNARC [1]. Yet there is no consensus as to how SNARC originates. This omission delays the understanding of SNARC in real-life numerical cognition, such as arithmetic [2, 3]. We first reject the frequently proposed dominant notion of SNARC as emerging from reading habits and then discuss alternative candidates.

2 THE ROLE OF READING

Early studies [4, 5] proposed that a left-to-right directional scanning habit is initially acquired with reading and might “spill over” into the numerical domain, thus causing the SNARC. But three arguments suggest that left-right reading habits alone cannot account for the entire SNARC.

First, some cultures have multiple reading habits that all contribute to SNARC: In Chinese-English readers the presentation format of a number (Chinese or Arabic symbol) determines whether it is mapped vertically or horizontally [6]. Hebrew readers have no SNARC - presumably because word and number reading habits cancel each other [7].

Second, effects of reading habits are much more short-lived than originally thought. The original idea that it requires years of exposure to a different reading culture to change a person’s SNARC [4] runs against the observation that Russian-Hebrew bilinguals modify their SNARC after reading only a few minutes of Cyrillic or Hebrew text [8]. Even reading a single word changes the SNARC from one second to the next [9]. And even during reading, SNARC reflects the positioning of small digits on the left or right side of a text [10]. Thus, directional reading habits provide only a small contribution to the overall SNARC.

Developmental data also argue against a strong contribution from reading to SNARC. 4.5-year-old children already explore objects more efficiently when they are numbered in left-to-right

ascending order [11, 12], establishes the small-left association as a default SNARC in Western cultures that needs to be explained.

3 OTHER CULTURAL CONTRIBUTIONS

First, SNARC might reflect culturally acquired finger counting habits. Most children acquire number concepts through finger counting when observing their parents. A cross-cultural comparison showed that adults in English-speaking countries prefer to start counting on the fingers of their left hand [13]. Thus, they associate small numbers with left space, and might have done so as children. Left-starters as a group have a stronger and more consistent SNARC [14] but studies of finger counting preferences in pre-school children are, surprisingly, still missing.

Second, language use offers an explanation for the default SNARC. The adjectives “left” and “small” are both linguistically marked (as are “odd” and “bad”). These concepts are more complex and less frequently used than their opposites (“right”, “large”, “even”, “good”). Linguistic markedness can account for the odd-left association (the MARC effect; [15]) and this might extend to small-left associations. Limited familiarity [16] or complete lack of larger number words [17] is associated with logarithmically compressed spatial-numerical mappings but nothing is known about SNARC in animals although they show size and distance effects [18].

Comparing SNARC across cultures and species can inform this debate. Finger counting habits are clearly culture-dependent [13] and different languages vary in how they mark spatial and numerical concepts.

4 EMBODIED MAGNITUDE PROCESSING

Our brains have developed to control our bodies, which in turn constrain our cognition and action. This elementary truth has recently regained attention as “embodied cognition”, often also as “grounded” or “situated cognition”. We propose a hierarchical relationship between these terms before discussing implications of this view for the origin of SNARC.

The most fundamental aspect of cognitive representations is their *grounding* which reflects properties of the world. One example is the large numbers-up and small numbers-down association from accumulating objects [19]. Interestingly, small digits occur more frequently on the left side of multi-digit numbers. According to the Newcomb-Benford Law, the probability of the digit “1” occurring in the first, second, third or fourth position of multi-digit numbers drawn randomly from real life, is usually .37, .19, .12, and .09, respectively [20]. SNARC might be grounded in our internalization of this statistic [10].

On top of conceptual grounding, *embodied* representations are shaped by sensori-motor constraints of the body, as in finger counting [14] and grasp aperture modulation [21]. Finally, range and context dependence of SNARC [4,22] reflect *situated*

¹ School of Psychology, University of Dundee, DD1 4HN Scotland UK
Email: m.h.fischer@dundee.ac.uk

² Neuropsychology Unit, University Hospital, 8091 Zurich, Switzerland.
Email: peter.brugger@usz.ch

number concepts. The embodiment stance makes three predictions about mathematical practice that can inform our search for the origins of SNARC.

First, different body postures, and their effects on spatial reference frames, should influence SNARC. Supporting this prediction, healthy adults generate smaller random numbers while turning their head left and larger random numbers while turning their head right [23]. Crossing hands over lateralized response keys can eliminate SNARC [24]. For right-hand responses, there is an index over ring finger advantage for small numbers, which reverses when the hand is held palm up [25]. Also, small digits facilitate left-hand recognition from an egocentric perspective (fingers pointing up), but right-hand recognition in an allocentric view (fingers pointing down) [26]. However, when situated in conflict, visual space dominates over hand space [27].

Secondly, embodiment predicts that special bodies should impact on our spatial minds and thus affect SNARC. There is clinical evidence that this is in fact the case. Patients with right hemisphere lesions, who fail to attend to the left side of their body and who neglect left hemispace, err towards large numbers when asked to bisect number intervals [28]. The error in explicit number interval bisection is not correlated, however, to rightward displacements in the bisection of real lines [29], nor does left-sided neglect manifest itself as small-number bias in random generation [30]. This indicates that, while hemispatial neglect affects orientation along a number line, the laws that govern attention in physical and number space are not identical.

A third prediction from the view of embodied magnitude processing is that functional hemispheric asymmetries should interact with the magnitude of SNARC in a given situation. Evidence for this comes from dual task paradigms, where left-hemisphere verbal memory load abolished SNARC in parity decisions but not in magnitude comparisons, whereas the opposite interference pattern occurred for right-hemisphere visual-spatial load [31]. These findings point to opposite hemispheric contributions to SNARC in different situations. Hemispheric activation paradigms can also bias healthy subjects' preference for small or large numbers during digit randomization [32]. Handedness, one of the most conspicuous signs of hemispheric specialization, does not influence SNARC per se [4] but could well do when space-number interactions are further modulated by finger use [14] or emotional valence [33].

5 OUTLOOK

Our brief review illustrates the bewildering number of potential sources of the association between number and space. We believe that this reflects the human capacity to quickly learn to associate any symbol or abstract relation with a spatial position or relationship [34, 35]. Thus, SNARC is the expression of some general cognitive rule that reflects the "placement" of an image in space (the spatialisation of ideas) and the relative (in)compatibility that emerges from using lateral effectors to respond to these ideas. Studies with children and amputees are likely to further advance our understanding of this intriguing phenomenon. The embodied cognition approach is a most suitable framework for further progress along this path.

REFERENCES

- [1] Wood G. et al. (2008). *Psychol. Sci. Quart.*, 50, 489-525.
- [2] Pinhas M. & Fischer MH. (2008). *Cognition*, 109, 408-415.
- [3] Knops A. et al. (2009). *Science*, 324, 1583-1585.
- [4] Dehaene S. et al. (1993). *J. Exp. Psychol. Gen.*, 122, 371-396.
- [5] Berch D. et al. (1999). *J. Exp. Child Psychol.*, 74, 286-308.
- [6] Hung et al. 2008. *Cognition*, 106, 1441-1450.
- [7] Shaki S. et al. (2009). *Psychon. Bull. Rev.*, 16, 328-331.
- [8] Shaki S. & Fischer MH. (2008). *Cognition*, 108(2), 590-599.
- [9] Fischer MH. et al. (2009) *Exp. Psychol.*, 56(5), 361-366.
- [10] Fischer MH. et al. (2010). *Brain & Cogn.* (in press).
- [11] Opfer J. & Furlong X. (2010). *J. Cross-cult. Psychol.* (in press).
- [12] Tversky B. et al., (1991). *Cog. Psychol.*, 23, 515-557.
- [13] Lindemann O. et al. (2010). *J. Cross-cult. Psychol.* (in press).
- [14] Fischer MH. (2008). *Cortex*, 44, 386-392.
- [15] Nuerk HC. et al. (2004). *Quart. J. Exp. Psychol.*, 57A, 835-863.
- [16] Siegler R. & Opfer J. (2003). *Psychological Science*, 14, 237-243.
- [17] Dehaene S. et al. (2008). *Science*, 320, 1217-1220.
- [18] Dehaene S. (1997). *The number sense*. London: Macmillan.
- [19] Lakoff G. & Nunez R. (2000). *Where mathematics comes from*. New York: Basic Books.
- [20] Burns, R. (2009). *Proc. Cog. Sci. Soc.*, paper 637
- [21] Andres M. et al. (2008). *Cortex*, 44, 414-419.
- [22] Bächtold D. et al. (1998). *Neuropsychologia*, 36, 731-735.
- [23] Loetscher T. et al. (2008). *Curr. Biol.*, 18, R60-R61.
- [24] Wood G. et al. (2006). *Cortex*, 42, 1069-1079.
- [25] Leuthard J. et al. (2005). *Cog. Behav. Neurol.*, 18, 173-178.
- [26] Conson M. et al. (2009). *Neurosci. Lett.*, 452, 176-180.
- [27] Brozzoli C. et al. (2008). *Proc. Nat. Acad. Sci.*, 105, 5644-5648.
- [28] Zorzi M. et al. (2002). *Nature*, 417, 138-139.
- [29] Doricchi F. et al. (2005). *Nature Neurosci.*, 8, 1663-1665.
- [30] Loetscher T. & Brugger P. (2009). *Neuropsychologia*, 47, 276-279.
- [31] Van Dijck J-P. et al. (2009). *Cognition*, 113, 248-253.
- [32] Loetscher T. & Brugger P. (2007). *Exp. Brain Res.*, 180, 655-665.
- [33] Casasanto D. (2009). *J. Exp. Psychol. Gen.*, 138, 351-367.
- [34] Bächtold D. et al. (2000). *Neuropsychologia*, 38, 864-872.
- [35] Gattis M. (2002). *Cog. Dev.*, 17, 1157-1183.

Acknowledgment: MHF was supported by a Travel Award from the Carnegie Trust.

Singing Numbers... in Cognitive Space

Martin H. Fischer¹, Marianna Riello², Bruno L. Giordano³ and Elena Rusconi^{2,4}

Abstract. We assessed the automaticity of spatial-numerical and spatial-musical associations with a dual-task paradigm. Participants performed magnitude and pitch comparisons on sung numbers with variable pitch. Concomitant tasks required retention of either colour or location information. Results suggest that spatial associations of both magnitude and pitch are resource-dependent and that the spatial association for pitch is more powerful than that for magnitude.

1 INTRODUCTION

When we make lateralized responses to sort stimuli into two categories (e.g., small or large numbers, low or high tones) we show a preferential association between left or low response keys and small numbers or low tones, and between right or high keys and large numbers or high tones. These two response biases are called Spatial-Numerical Association of Response Codes [SNARC, ref. 1] and Spatial-Musical Association of Response Codes [SMARC, ref. 2], respectively. Both findings suggest that we use spatial representations for problem solving. Initial results suggested that SNARC requires no cognitive resources as it emerges automatically even when number magnitude is task-irrelevant [3] but cultural differences and flexible results raised the possibility of a non-automatized SNARC strategy [4]. Supporting the latter, SMARC is stronger in trained musicians than novices [2], suggesting learned spatial associations. This argument assumes that SNARC and SMARC use similar processing mechanisms.

A diagnostic tool to help understand processing mechanisms in a given task is the dual-task paradigm: If task A performance deteriorates when performed together with task B then task A is not automatized. Different task combinations can reveal the underlying processes. This method recently showed that SNARC disappears under verbal load in parity judgment but under spatial load in magnitude comparison, suggesting that spatial codes can be either visuo-spatially or verbally mediated [5]. Here we apply the dual-task logic to investigate the relationship between SNARC and SMARC and whether these are automatized effects.

2 METHOD

16 musically untrained students (9F, 7M, mean age: 22 years, 3 left-handers, all native English) participated. They had no visual or hearing impairments. In the two main tasks (magnitude and pitch comparison tasks, MCT and PCT) stimuli consisted of 5 English number words (*two, three, four, five and six*) sung by 4

different professional singers (2M, 2F) at 5 equidistant pitches (*B2, E3, A3, D4 and G4*). In MCT participants decided whether a sung target number was smaller or larger than the preceding reference; in PCT they decided whether the target was lower or higher in pitch than the preceding reference. Number *four* sung in pitch *A3* (the central value on each dimension) served as fixed reference for both tasks. Responses were given by pressing the lower left *V* and upper right *U* keys on a QWERTY keyboard to capture both horizontal SNARC and vertical SMARC.

Participants fixated a cross which appeared centred on a laptop display. The reference stimulus (1000 ms) was played through headphones, immediately followed by the target (1000 ms). Responses were recorded in ms from target onset and followed by a 300 ms central fixation cross (for correct responses), exclamation mark (for errors) or question mark (for omissions).

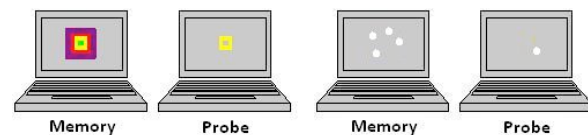


Figure 1. Sample memory and probe displays in the colour task (left, “yes” trial) and the location task (right side, “no” trial).

Two concomitant short-term memory tasks required maintenance of either colour or location information presented before each trial of the main task. In the *colour* task, squared concentric frames appeared centred on the screen; participants had to remember their colours. At the end of a trial, a single squared frame appeared and participants said aloud whether it had been present in the previous memory display (Fig. 1, left panels). In the *location* task, participants remembered the position of four dots appearing simultaneously on the screen at different positions of an imaginary central circle (each of them in a different quadrant). At the end of a trial, a single dot appeared at a position on the same imaginary central circle, and participants said whether it was part of the previous memory display (Fig. 1, right panels). The experimenter entered answers by clicking one of two mouse buttons. In either concomitant task, the memory stimulus remained visible for 150 ms, whereas the probe remained on screen until a response was entered.

The experiment was divided in two sessions to avoid fatigue, and completed on two different days. Each session comprised an initial training with either the colour and location tasks alone. Then baseline data were collected for one of the main tasks (single-task condition, first half), followed by the same task with both concomitant tasks, and finally by another baseline (single-task condition, second half). On each session a different main task was tested. Order of main and concomitant tasks was fully counterbalanced between participants, as were response rules.

1 School of Psychology, University of Dundee, Scotland UK.

2 Center for Mind/Brain Sciences, University of Trento, Italy

3 Schulich School of Music, McGill University, Montreal, Canada

4 Dept. of Security and Crime Science, UCL, London, UK

3 RESULTS

Main tasks were both performed with mean accuracy of 93% in the single-task condition, and 93% and 92% in the colour and location dual-task conditions, respectively. ANOVAs evaluated effects of *Task* (MCT, PCT), *Condition* (single-task, colour dual-task, location dual-task), *SMARC* (compatible, incompatible), and *SNARC* (compatible, incompatible) on correct RT and arcsine-transformed accuracy values. We report effects with $p < .05$.

Correct RT showed faster decisions in magnitude than pitch comparison. Single-tasks were performed faster than either dual-task. *SMARC* was reliable but *SNARC* was only marginal. Figure 2 shows a *Task* $\times$ *SMARC* $\times$ *SNARC* interaction.

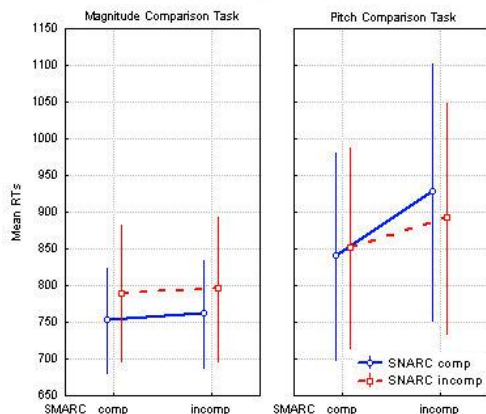


Figure 2. MCT showed a 35 ms SNARC but no SMARC. PCT showed a 63 ms SMARC but a 36 ms reverse SNARC in *SMARC* incompatible trials. Bars: 95% confidence intervals.

The accuracy analysis revealed main effects of *Task*, *SMARC* and *SNARC*. In MCT there was reliable *SNARC* without *SMARC*. In PCT there was reliable *SMARC* and a reverse *SNARC* in *SMARC* compatible trials, thus complementing RT results. Of interest for the present study, *Condition* revealed two interactions: First, *SNARC* was only reliable in the colour dual-task. Second, *SMARC* was not present during MCT but significant in both dual-task conditions during PCT (Figure 3).

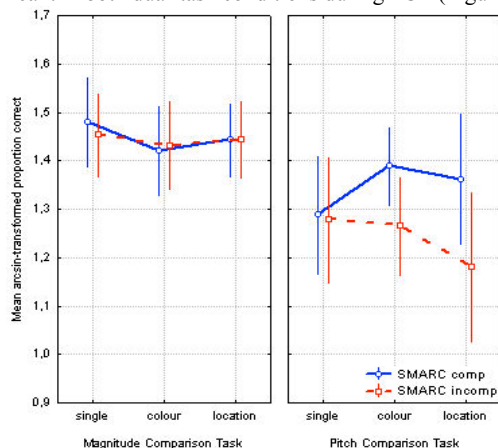


Figure 3. Relative to single task, SMARC suffered from dual-tasks in MCT whereas it improved in dual tasks during PCT.

4 DISCUSSION & FUTURE DEVELOPMENTS

Single tasks were easier than dual tasks (as required by the dual-task logic) and MCT was easier than PCT for our musically untrained participants. Both speed and accuracy data indicate that attending to magnitude induced SNARC without SMARC and attending to pitch induced SMARC without SNARC. This clear top-down modulation of both effects argues against automatic spatial associations of either stimulus dimension.

Only pitch but not numbers showed an *overall* spatial mapping in RT, and only SMARC was relatively immune to the different dual task demands (in terms of accuracy). This suggests that SMARC is overall stronger (more prevalent and consistent) than SNARC, which agrees with the considerable flexibility and cultural variability of SNARC effects [4].

The spatial memory task abolished SNARC in accuracy. This was not due to increased difficulty (which was comparable in the colour task) and suggests that both spatial mapping of numbers and memorizing locations use a common process. A candidate process is rapid successive deployment of spatial attention to all four probe locations to assign and maintain their coordinates during a trial [6]. Such attentional rehearsal would reduce resources available for associating numbers to spatial locations but would not be required to maintain verbal labels (colour names) in working memory.

This proposal might also account for the reversed SNARC while participants focused on pitch under incongruent SMARC mappings. For example, pressing a low button for a high pitch generated an incongruent spatial code (associating small height with large frequency) but also yielded faster responses for high magnitudes. Given the relevance of pitch, this result probably reflects the flexibility of mapping the irrelevant magnitude dimension of our compound stimuli: the assigned response coordinates linked high magnitude via high pitch.

The stronger SMARC under both dual task conditions might reflect the cost of incompatible more than the benefit of compatible SMARC mappings. Thus, even under load it remained easy to associate high pitch with high keys but it became harder to impose a reversal of this natural mapping.

Overall, this study documented resource-dependent spatial mapping of both magnitude and pitch. It suggests that the spatial mapping of pitch is more powerful than that of magnitude, consistent with its ontogenetically earlier acquisition and universal cultural occurrence. Future work should measure SNARC and SMARC throughout musical skill acquisition.

REFERENCES

- [1] S. Dehaene, S. Bossini and P. Giraux. *J. Exp. Psychol.: Gen.* 122: 371-396 (1993).
- [2] E. Rusconi, B. Kwan, B. L. Giordano, C. Umiltà and B. Butterworth. *Cognition* 99: 113-129 (2006).
- [3] W. Fias, M. Brysbaert, F. Geypens, and G. d'Ydewalle. *Mathem. Cognition*, 2: 95-110 (1996).
- [4] M.H. Fischer. *Cortex*, 42(8): 1066-1068 (2006).
- [5] J.P. Van Dijk, W. Gevers, and W. Fias. *Cognition*, 113: 248-253 (2009).
- [6] E. Awh and J. Jonides. *Trends Cognit. Sci.*, 5: 119-126 (2001).

Acknowledgments: MHF was supported by the Carnegie Trust. ER was sponsored by an EC Reintegration Grant (MERG-CT-046511) and a Research Fellowship from the University of Trento.

Interpretation is an action: understanding diagrams by manipulating them

Valeria Giardino¹

Abstract. In this paper, I will defend the idea that *interpretation is an action*. In particular, I will argue that mathematical diagrams and figures are understood only when the user knows how to manipulate them, to the aim of making an inference and obtaining a new conclusion.

1 SEEING AS AND REPRESENTATIONAL SYSTEMS

The motto in the title, ‘interpretation is an action’, is taken from Wittgenstein. To introduce the main idea of this paper, I will present one of his numerous examples of reasoning with visual patterns [12]. I am well aware that, as Wittgenstein himself suggested, the philosopher should resist the use of a ‘one-sided diet of examples’. Yet, I believe that this example alludes in a striking way at the link I am interested in between interpretation and action in diagrammatic reasoning.

Imagine two shapes. Shape A is a shape that may or may not have symmetrical properties of some kind, but most importantly it does not look familiar to the observer. That means for example that the observer cannot find a word to name it - she *does not have any name* for it. Moreover, the observer is not sure to what it alludes to, she can only guess: a lamp on a table maybe? Shape B instead is recognizable as a sequence of letters arranged in a meaningful word.

Now, imagine reversing both shapes as in Fig.1 and Fig.2²:

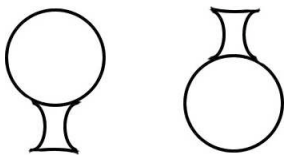


Figure 1. Shape A and its reversed version



Figure 2. Shape B and its reversed version

The difference between the visual impressions of shape B and of its reversed version is perceived as bigger than the difference between

the visual impressions of shape A and of its reversed version. Why that? One reason could simply be that shape B looks more salient to the observer than shape A. Of course the observer could speculate that shape A is a picture of a lamp on a table and its reversed version a picture on how the lamp looks like if it is hanging from the ceiling. Yet, this reading would not be necessarily shared by other observers. By contrast, if we consider another salient shape such as shape C and its reversed version in Fig.3, we will see that also in this case looking at the reversed version disturbs us more than in the case of shape A.

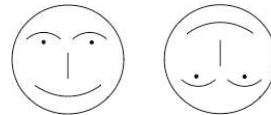


Figure 3. Shape C and its reversed version

Nevertheless, I want to suggest something more that what is happening for shape C. In fact, salience here is not enough. Or better, salience needs to be explained. My claim is that shapes are perceived as salient when the observer knows how to *use* them. To clarify, the observer who recognizes shape B or shape C is ready to *reproduce* them, in a meaningful way and without damages, far better that she could in the case their reversed version. Moreover, in the case of shape B, the observer has been *trained* to do that - that is, she has been trained to write. By contrast, in the case of shape A and its reversed version, they are both arbitrary shapes equally easy (or difficult) to be drawn[6].

The hypothesis here is that the understanding of a picture, that means its *correct interpretation*, has consequences on our *behavior*, and triggers some particular reaction in us, in what we can *do of* the picture, reproducing or manipulating it, in a meaningful way. The observer who has learnt how to write and who masters the writing system, is ready to recognize something, a word, in shape B. In the same way, shape C is meaningful, since the observer would know in which context to use it. Moreover, she would also know in which way to change it in order to change its meaning as well (as with emoticons). Shape A and its reversed version instead, do not recall any system of representation to the observer and as a consequence any kind of possible meaningful manipulation.

Can we assume this kind of approach also for *mathematical* diagrams and figures? To explore this possibility, I will introduce in the following sections the notion of *constructional* drawings and consider its relationship with mathematical practice.

¹ Institut Nicod (EHESS-ENS-CNRS), Paris, email: Valeria.Giardino@ens.fr

² Actually, Wittgenstein reverses shape A and reflects shape B as in mirror, but I will reverse both to avoid incongruities.

2 THE NOTION OF CONSTRUCTIONAL DRAWINGS

The hypothesis behind this paper is that what is relevant in discussing diagrammatic reasoning in mathematics is the *use* of diagrams and figures more than the individuation of their propositional content. As Avigad suggests ascriptions of understanding are best understood in terms of the possession of certain *abilities*. It is an important philosophical task to try to characterize the relevant abilities in sufficiently restricted contexts in which such ascriptions are made. This is a central issue also for automated formal verification, where, as he suggests, when it comes to verifying an inference, different sorts of agents will have different strengths and weaknesses. As humans, we spontaneously refer to diagrams because we are good at recognizing symmetries and relationships in information so represented; machines, in contrast, can keep track of gigabytes of information and carry out exhaustive computation where we are forced to rely on little tricks. If we are interested in differentiating good teaching practice from good programming practice, we have no choice but to relativize our theories of understanding to the particularities of the relevant class of agents [1].

Let us consider a useful analogy and take into account drawings for technical uses, e.g. drawings that are behind the design of an artifact. In spite of their appearance, these drawings are not depictive but serve as *inferential* or *prescriptive tools*. An effective technical drawing has the property of ‘good legibility’, which is a property that does not depend upon the respect of some mathematical rule, but upon the specifics of *biologically and socially characterized* visual systems of the human users, even when these are assisted by artificial processing systems[7]. These constraints long precede the regularization of representational systems by the definition of explicit rules, such as, for example, in projections. Furthermore, thanks to these constraints, drawings can be *used* and *reproduced*. Learning to use them is learning to correctly manipulate them in order to acquire new knowledge.

The same happens for other cases of diagrams. For example, the user demonstrates to know how to use a graph when, once asked to reproduce it, in doing that she gets rid of ‘chartjunk’, that is of the set of all visual elements that are not necessary to comprehend the information represented and that in some cases can even be distractors [8]³. Analogously, Euler or Venn circles surely exploit some ‘visual schemas’ such as *shape containment* [5]; nevertheless, they can be manipulated and reproduced correctly only when the user knows which of their spatial properties are invariant in all their possible meaningful manipulations. For example, circles can be replaced by squares, provided intersections are preserved; no color can be inserted in a Euler circle, while on the contrary the location of the color should be preserved in a Venn circle.

I will define this kind of drawings that are given as *tools* to calculate and measure, as well as to plan and compare different solutions, as *constructional drawings*. They can - and to some extent, they must - be adjusted, reconsidered, prepared in advance. In the study of drawing there has been the tendency to neglect much of what today we would call *detail drawing*, which in the past was marked on the actual stone or wood in order to produce an artifact [2]. The shapes of constructional drawings may be understood as *operations* that are not visible in the final product: it is because we are able to make the drawing that we can understand the result it designs. Until the second half of the 20th century, engineering schools have illus-

trated the understanding of drawings by teaching *how to make them*: training included apprenticeships in which students learnt to appreciate the nature of materials and machines through laboratory experience. The knowledge was thus based on sensory observations, and masters guided the apprentices showing them what to look for [3].

3 FROM CONSTRUCTIONAL DRAWINGS TO MATHEMATICAL DIAGRAMS

Diagrams and figures in mathematics as well can be defined as *constructional*, in analogy with the cases considered in the previous section.

First of all, they are subject to constraints both as two-dimensional *physical* objects and as objects that require a user to *interpret* them. Topological relations, for example, are very basic spatial relations such as proximity or enclosure that would not change in a diagram if the diagram were printed on a rubber sheet and the sheet were stretched or twisted [10]. Nevertheless, the recognition of such spatial relationships must be accompanied by interpretation, so that the diagram can be used to the aim of obtaining new conclusions within a specific theory.

To integrate these two kinds of constraints, it is possible to refer to the notion of *manipulation practice*. A diagram is interpreted *dynamically*: informal inferences take the form of transformations. In the range of all possible transformations only some of them are legitimate within the theory considered: the rules of diagrammatic representations are therefore normally externalized as procedures, and, as a consequence, what must be learned is not abstract rules, but *instructions* on how to act on the diagram and to read and interpret it correctly. To sum up, the *correct interpretation* of a diagram gets intimately connected to the *systematic actions that are performed on it* [4].

Some seminal studies in psychology have discussed *constructive* perception, i.e. the coordination of two processes: reorganizing perception and associating ideas [9]. These processes are correlated independently with a *perceptual* ability, reorganizing parts of figures, that means grouping them in such a way that they can be reinterpreted, and with a *conceptual* ability, associative fluency. The perceptual component facilitates perceptually reorganizing what one sees in the external environment, enabling detection of subtle features and relations that novices might not discern; the conceptual component enables fluency in generating new and related thoughts. To keep generating new ideas in a domain these two components must be tightly coordinated. It would be interesting in this respect to analyze the educational software actually available to teach mathematics, logic and geometry. Does this kind of constructive perception get nurtured by using them?

Expertise is thus attained by coming to differentiate and reorganize parts, wholes, perspectives, and reference frames. Perceptual expertise must then be coordinated with behavior. The implication is that expertise is not acquired straightforwardly by teaching, but rather by a coordinated act of perceiving new features and relations previously unheeded, conceiving of new and associated thoughts previously unattended to and acting on the surrounding environment in the domain.

Some work has been done in computer science as well, to the aim of understanding reasoning and problem solving activities that combine representation and inferences in both diagrammatic and symbolic modes in such a way that different subproblems are solved by the mode that is best suited for them.

One attempt is the implementation by Chandrasekaran and col-

³ Examples of chartjunk include heavy or dark lines in the graph, unnecessary text or inappropriately complex fonts, pictures or icons to represent data graphs, or ornamental shading.

leagues, based on three dimensions and on different perceptual routines [11]. The first and most basic dimension is that of seeing the world as composed of *objects*, which are discriminated from ground and perceived as shapes. The second dimension is that of *seeing as*, e.g. recognizing an object as a telephone, or a figure in a diagram as a triangle. The recognition vocabulary comes in a spectrum of domain-specificity. The third dimension is *relational* - seeing that an object is to the left of another object, taller than another, is part of another, etc.. In this implementation, the only perceptual routines that are immediately applied as a diagram is created or changed are represented by the set of emergent object routines. By contrast, all the other routines apply only in response to a problem solving need. Furthermore, even in the case of emergent objects, the number of these objects, when linked to problem solving, can grow exponentially.

Even if their motivation is not to propose a psychological theory, but rather to expand the options available to designers of AI and human-machine systems, I think that their strategy goes in the direction I discuss here about how to deal with diagrammatic reasoning.

Let us assume then that in mathematics, the most interesting diagrams are all dynamic and constructional, that means inference-promoting and correctly understood only when reproduced and correctly manipulated. In fact, diagrams and figures would invite the user to perform some operation on them, in order to solve a problem. Going back to Wittgenstein, we could claim that one of the consequences of this view would be that mathematics is a set of techniques and not a series of true and false propositions. To clarify, the point is that mathematical understanding would not concern the knowledge of a list of propositions of which we want to know the truth value, but rather of a web of interconnected facts that results in the learning and in the application of a mathematical practice. One consequence is seeing proofs in a new light: it is the very process of giving the proof that teaches the mathematician what the proof is about.

4 A POSSIBLY MORE GENERAL CONCLUSION (work in progress)

If the use of mathematical diagrams involves the coordination of several processes (and abilities), we could go even further. In fact, three cognitive systems appear to play a role here: the *conceptual* system, the *visuo-spatial* system and the *motor* system. The hypothesis is that diagrams have been invented (or discovered) with the function of facilitating and organizing our reasoning and are used dynamically. Their use would engage all of these three systems: the same manipulation of a diagram is seen by the conceptual systems as the compliance to some visual invariance, by the visual-spatial system as a transformation in time, by the motor system as a movement or a movement plan. Dynamically interpreted diagrams work at the interface between these three systems to churn out an inference that puts forward a new conclusion.

ACKNOWLEDGEMENTS

I would like to thank the referees for their comments which helped improve this long abstract.

REFERENCES

- [1] J. Avigad, 'Understanding proofs', in *The Philosophy of Mathematical Practice*, ed., Paolo Mancosu, Oxford University Press, (2008).
- [2] P. J. Booker, *A History of Engineering Drawing*, Chatto and Windus, 1963.
- [3] E. S. Ferguson, *Engineering and the Mind's Eye*, MIT Press, 1992.
- [4] V. Giardino, 'Intuition and visualization in mathematical problem solving', *Topoi*, (forthcoming, accepted on October 26, 2009).
- [5] G. Lakoff and R. Nuñez, *Where mathematics comes from: How the Embodied Mind Brings Mathematics into Being*, Basic Books, 2001.
- [6] W. G. Lycan, 'Gombrich, wittgenstein, and the duck-rabbit', *The Journal of Aesthetics and Art Criticism*, **30**(2), 229 – 237, (1971).
- [7] P. Maynard, *Drawing distinctions: the variety of graphic expression*, Cornell University Press, 2005.
- [8] E. Tufte, *The Visual Display of Quantitative Information*, Graphics Press, 2nd edition edn., 1983/2001.
- [9] M. Suwa B. Tversky, 'Constructive perception: A skill for coordinating perception and conception', *Proceedings of the Cognitive Science Society Meetings*, (2003).
- [10] J. Willats, *Art and Representation: New Principles in the Analysis of Pictures*, Princeton University Press, 1997.
- [11] B. Chandrasekaran U. Kurup B. Banerjee J.R. Josephson R. Winkler, 'An architecture for problem solving with diagrams', *Lectures Notes in Computer Science*, **2980**, 235–256, (2004).
- [12] L. Wittgenstein, *Philosophische Untersuchungen, Philosophical Investigations*, Basil Blackwell, 1958.

The body in Renaissance arithmetic: from mnemonics to embodied cognition

Albrecht Heeffer¹

Abstract. In Medieval and Renaissance arithmetic we find several instances of references to body parts or actions involving body parts. In this paper we will address the question on the historical functions of body parts in mathematics and discuss its relation to the currently prevailing practice of symbolic mathematics.

1 INTRODUCTION

There can be no doubt that the earliest human practices of counting and measuring involved body parts, most notably the fingers and hands. The bijective relation between the fingers on our hands and counted objects allows for counting and ordering of ten objects. Our current use of the decimal system is most likely related to this fact. Other systems, such as base-20, practiced during the first millennium among the Maya of Central America refer to other combinations of fingers or body parts. But the art of dactylonomy improves on such basic relations and includes bending, crossing and touching of fingers to represent all possible numbers up to 10,000. We have strong indications that the first symbols used in our script draw their analogy from finger counting practices. It is generally accepted that Chinese numerals are a direct representation of hand gestures [4]. The first three digits 一, 二, 三, obviously represent fingers; 四, four: the thumb is held in palm and four fingers are extended, 五, five: the fingers are extended with the thumb facing upwards, 六, six: the little finger and thumb are extended and other fingers are closed, and so on. For Egyptian or Babylonian mathematics we have no such clues but references exist to the practice of finger reckoning in these cultures. The Latin term *digit* both denotes a cipher as well as a finger and the verb *computare* originally referred to Roman number arithmetic. The first chapter of Bede's *de temporum ratione liber* (c. 730) is entitled "de computo et loquela digitorum" (On computing and speaking with the fingers). Finger reckoning was also practiced in the Arab world and known as *isāb al-'aqd* (literally: arithmetic of the joints). By the time Fibonacci wrote his *Liber abbaci* on reckoning with Hindu-Arabic numerals (1202), the European abacus and finger reckoning were two competing systems, both using a tangible form of practicing mathematics.

2 FUNCTIONS OF BODY PARTS

In this section we discuss the possible functions that body parts had in Renaissance arithmetic. We discern four different functions:

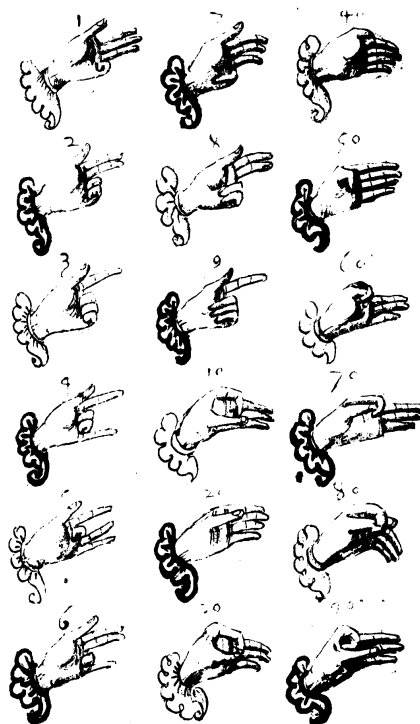


Figure 1. Number representations in abbaco culture (from ms. Add.8784 © The British Library)

- **Memorization:** though finger reckoning leaves no permanent record as does the abbaco method of calculating with the pen, short-term memorization is the primary and most essential function of body parts in arithmetic. Fingers are used to memorize intermediate results during calculation such as the carry in multiplication procedures. Body parts can also have long-term mnemonic functions. An example is the use of knuckles on the fist to memorize the months of a year which have 31 days.
- **Communication:** the representation of all numbers up to 10,000 on a single hand was a skill which was taught at the abbaco schools of fourteenth and fifteenth century Italy. As extant witness of this practice we have about 250 abbaco manuscripts of that period. Figure 1 shows the gestures for numbers up to 90, from an anonymous manuscript of 1443. We find the same representation in Pacioli's *Summa* of 1494 (pp. 36-7). The chapter in which the technique is described is called "to set out numbers on the hand" (*porre alle mani*). The purpose thus is not to calculate but to represent and display numbers. Within the context of abbaco practices such skill functioned as a way of

¹ Fellow of the Flemish Research Foundation (FWO). Center for Logic and Philosophy of Science, Ghent University, Belgium. Email: albrecht.heeffer@ugent.be.

communicating and negotiating prices or quantities in merchant transactions. Gestures assured reliable communication across language barriers, functioned in noisy environments and could span short distances as between a ship and shore.

- Visualization: our current practice of symbolic mathematics allows us to represent a problem or a mathematical structure using a single formula or expression (as the systems of equations (1), discussed below). The importance of visualization in mathematics becomes painfully evident when one is confronted with loss of sight. Nicholas Saunderson, Lucasian professor of mathematics at Cambridge from 1711, was blinded by smallpox. For his own purpose of visualizing arithmetical operations he devised a wooden board and a method called “palpable arithmetic” which is described as a preface to his *Elements of algebra*. A large wooden board was divided into 100 little squares with each nine holes. Digits were represented by combinations of two pens in these holes. While the contrivance could be compared with the abacus, its purpose was basically to envisage large numbers. The tactile interaction of hands and fingers with the wooden board and pegs thus functioned as a way to visualize and arrange mathematical objects and operations.
- Representation of procedures: this last function is the most advanced one and will be discussed in the following section. In Lakoff’s terminology body parts here use the metonymic mechanism to represent both an unknown quantity of a problem as well as the operations on that quantity prescribed by a proto-algebraic recipe [2, 74].

3 PROBLEM SOLVING PROCEDURES

In previous section we discussed the basic functions of body parts in counting, representation of numbers and calculations in Renaissance arithmetic. But we found references to body parts in more advanced problem-solving procedures. Renaissance mathematics is very much based on the use of what I have coined as proto-algebraic rules [1]. These are procedures which are not algebraic themselves but are based on previous algebraic solutions or procedures. Such rules are usually named and are specific to standard problem types. The book on arithmetic by Johannes Widmann [6] includes no less than 32 such rules.

One of the rules Widman discusses is called *regula coecis* and demonstrates how proto-algebraic rules can be represented by tangible operations and actions. The rule is intended to solve a type of problem best known as “the hundred fowls problem”. In European arithmetic it appears more frequently as a problem about a number of men, woman and children who have to pay the bill at a tavern. This formulation, which explains the alternative name *Regula virginum*, first appears in the Parmiers manuscript written around 1430 [6]. The second arithmetic book by Adam Riese [5, 135], lists two problems which can be represented in modern symbolism as follows:

$$\begin{cases} x + y = 12 \\ 3x + 5y = 81 \end{cases} \quad (1)$$

$$\begin{cases} x + y + z = 20 \\ 3x + 2y + z = 40 \end{cases}$$

The first problem involves 12 men and woman who consume a total of 81. Men pay 5, woman 3. The question is how many of each is present in the tavern. In a later edition of 1574, the operation is described as *Zech rechnen* (f. 69^v):

Schreib vor dich gegen der lincken handt die anzahl de Personen. Gegen der rechten handt / wie viel sie vertroncken / und in die mitte / wie vil einjegliche Person / jeglichs geschlechts in sonderheit gibt. Darnach mach das gelt dem menigsten uberall gleich / als dann multiplicir das kleinst an der bezahlung mit den Personen / und nimb von dem das sie vertroncken haben / Das da bleibt ist die zahl / welche getheilt sol werden.

The procedure prescribes to hold the number of people at the left hand and the consumed total at the right hand; the contribution of each, 5 and 3 in the middle. Then 3 is subtracted from 5 and the difference is called the divisor. The amount of the left hand is multiplied by 3 and subtracted from the amount on the right hand, thus 18. Divided by the divisor 2 the result gives the number of men, thus 9. From this the number of woman can be established.

While the procedure described by the rule can be matched with the one from the *Ganitasārasamgraha* by Mahāvīra (850, Sūtra 143 ½; [3], 303), the embodiment of the proto-algebraic rule is apparently a European invention. The same tangible procedure is discussed in the Flemish arithmetic of Vander Gucht (1569, ff. 80^v-81^v, 1594, ff. 70^f-71^f) with small differences in formulation.

7 CONCLUSION

We have shown that the use of body parts in Renaissance mathematics goes back to a tradition of finger reckoning which is probably as old as mathematics itself. While basic finger arithmetic functions as an aid for memorization, communication and visualization we find in Renaissance arithmetic examples of embodiment of proto-algebraic procedures. We believe that this form of embodiment compensated a lack of symbolism which developed by the end of the sixteenth century. Most of the functions of tangible procedures have now been replaced by our current practice of symbolic mathematics.

REFERENCES

- [1] A. Heeffer, “The Tacit Appropriation of Hindu Algebra in Renaissance Practical Arithmetic”, *Ganita Bhārati*, vol. 29, 1-2, pp. 1-60 (2007).
- [2] G. Lakoff and Rafael Núñez, *Where Mathematics Comes From*. Basic Books, New York (2000).
- [3] J. Needham, *Science and Civilization in China*. Cambridge University Press, vol. IV, (1959).
- [4] Padmavathamma and Rao Bahadur M. Rangācārya (eds.) *The Ganitasārasamgraha of Sri Mahāvīrācārya with English transliteration, Kannada translation and notes*, Sri Siddhāntakirithi Granthamāla, Hombuja (2000).
- [5] A. Ries, *Rechenung auff der linihen und federn in zal, mass und gewicht auff alleley handierung*, Mathes Maler, Erfurt (1522).
- [6] J. Sesiano, “Une Arithmétique médiévale en langue provençale”, *Centaureus*, 27, pp. 26-75 (1984).
- [7] J. Widmann, *Behende vnd hubsche Rechenung auff allen Kaufmanschafft*, Leipzig (1489).

Dynamic construals, static formalisms: Evidence from co-speech gesture during mathematical proving

Tyler Marghetis¹ and Rafael Núñez²

Abstract. This paper presents results from ongoing research using co-speech gesture to investigate the nature of mathematical concepts. Traditional accounts of the content of mathematical concepts have focused on formal language and definitions. In analysis, key concepts such as continuity are defined by means of static existential and universal quantifiers ranging over static numbers. Recent evidence from Cognitive Science, however, suggests that mathematical concepts are construed through various mechanisms of everyday cognition — such as conceptual metaphor and fictive motion — which are at odds with static conceptions. We analyse the co-speech gesture produced by graduate students while collaborating on a proof in analysis. The results support the claim that many mathematical concepts, which formally make use of static entities and relations, are, cognitively, inherently dynamic.

1 INTRODUCTION

Advanced mathematics is characterized by rigorous methods and symbolic manipulation. Traditionally, this has been taken as evidence for the abstract, formal nature of mathematical concepts. Contrary to this view, recent research from Cognitive Science suggests that nature of such concepts is largely embodied, and so mathematical concepts — such as infinity, continuity, number, and the rest — are not abstract and generated from formal definitions, but metaphorical and grounded in experience [17]. The evidence for the embodiment of mathematics has been drawn largely from mathematical language using techniques in cognitive linguistics, although recent research has added some qualitative studies of gesture [21].

This paper reports on the results of an ongoing research project on the nature of mathematical proof. In particular, we focus on the co-speech gesture produced by graduate mathematics students while collaborating on a proof and argue that the character of these gestures supports the claim that mathematical concepts are largely metaphorical and embodied and that their nature cannot be reduced to pure formalisms. This research extends earlier work that used gesture to adjudicate the cognitive reality of conceptual metaphor in mathematics [21, 22], moving beyond the pedagogical setting to consider advanced mathematical practice in a naturalistic setting.

The paper is organized as follows. First, we briefly review the evidence for metaphor and fictive motion in mathematical cognition. Second, we introduce the study of gesture as a tool for investigating the cognitive reality of these phenomena during mathematical practice. Then we present some quantitative results on co-speech gesture production from a semi-structured case study of collaborative proof

by mathematics graduate students at a large research university. Finally, we discuss implications for the nature of mathematical concepts.

2 DYNAMISM IN MATHEMATICAL DISCOURSE

How can we, as limited beings, understand abstract concepts that transcend our human experience? According to some approaches in Embodied Cognition, the answer lies largely in our shared embodiment, in the way our bodies — modulated by culture — structure thought and experience. Research in Cognitive Science suggests that abstract concepts are created and understood via a number of fundamental cognitive processes that extend the inferential structure of our bodily experiences. These processes include conceptual metaphor and metonymy [16], conceptual blending [4], and fictive motion [29]. Using the tools of Cognitive Linguistics, Lakoff and Núñez [17] argued that even mathematical concepts rely on these cognitive processes. While the details of the various construals underlying mathematical thought are beyond the scope of this article, in this section we will quickly review a few salient instances in which fictive motion lends dynamism to putatively static mathematical concepts.

Consider the concept of a limit, a central notion in calculus. Formally, the limit of a function is defined by a chain of inequalities:

Let a function f be defined on an open interval containing a , except possibly at a itself, and let L be a real number. Then $\lim_{x \rightarrow a} f(x) = L$ means that, for all $\epsilon > 0$, there exists $\delta > 0$, such that whenever $0 < |x - a| < \delta$, then $|f(x) - L| < \epsilon$.

This formal definition of a limit refers to static universal and existential quantifiers, static numbers, motion-less arithmetic difference and static inequalities. Nowhere in this definition is there any mention of movement. Mathematicians, on the other hand, speak of a function “tending to,” “moving toward” or “reaching” a limit — all of which, contra the formal definition, invoke a sense of motion [21]. These expressions are a form of fictive motion [29], the process by which we unconsciously conceptualize static entities in dynamic terms.

Fictive motion construals always involve the motion of a *trajector* across a *landscape*. When we say, for instance, that “the Equator passes through Brazil,” the Equator — a purely imaginary static entity — is construed as a moving agent (trajector) dynamically crossing a country (landscape). Similarly, we can say that a fence *stops* at a tree or that a road *runs* along the coast — even though both fences and roads are completely stationary and thus incapable of stopping or running.

¹ University of California, San Diego, USA, email: tmarghet@cogsci.ucsd.edu

² University of California, San Diego, USA

This same cognitive mechanism of fictive motion injects dynamism into a wide range of statically defined mathematical entities. A function, for example, is formally defined as a static relation between two sets, the domain and the range, but mathematicians nevertheless describe functions dynamically as “reaching a limit,” “going down towards a minimum,” or “oscillating,” in each case evoking a construal in which an imaginary trajectory travels along the path of the function [21]. Fictive motion is similarly at work when we say that sequences are “approaching,” “decreasing,” or “converging,” and when arithmetic is construed as motion along a number-line³. Dynamism, therefore, is present throughout the language of mathematics — showing up in the discourse surrounding continuity, functions, and even arithmetic — and lends credence to the claim that mathematical thought itself is dynamic, metaphorical, and embodied.

3 MATHEMATICAL GESTURE

One objection to this line of reasoning is that these metaphorical construals are mandatory in mathematical discourse — and that, therefore, they are conventionalized, dead, stripped of any cognitive reality. Consider the geometrical procedure of “reflection,” such as when a point in the Cartesian plane is *reflected across the origin*. The word “reflection” has roots in the Latin verb *reflectere*, meaning “to bend back.” There is no other way of describing the geometrical procedure, and therefore the same lexical item, “reflection,” is always used to describe the procedure — yet there is also no evidence suggesting that the concept of reflection involves an automatic construal of *bending backwards*. Certain aspects of mathematical discourse, therefore, are codified and devoid of cognitive significance. Could the dynamic discourse surrounding limits, functions, and sequences similarly involve conventionalized discourse? To address this objection, we must supplement corpus studies of mathematical discourse with additional lines of converging evidence.

The cognitive reality of mathematical construals is supported by the study of gesture, that is, motor action co-produced with speech and thought in real time. Gesture is universal, unconscious, and essential to communication. Most importantly, gesture offers a “window into the mind” [8]. When co-produced with abstract thinking, gestures parallel the metaphorical mappings exhibited linguistically [19, 2, 26], and give us insight into the representation of mathematical concepts and solution strategies [1, 5]. In particular, Núñez [21, 22] demonstrated that mathematicians’ gestures in pedagogical contexts supply converging evidence for the metaphorical and embodied nature of mathematical concepts.

The previous research on mathematical gesture, however, has dealt primarily with gesture production during pedagogy or in the context of elementary mathematical problem solving. In these settings, gestures were found to be dynamic — in line with the predictions of cognitive linguistics, and suggesting that mathematical concepts are metaphorical *in those settings* [22, 3]. But would we expect any other behavior? It is standard pedagogical practice to use “real world” examples of abstract concepts, to ground the abstruse in the everyday. Physics teachers might describe electricity, for instance, as “water running through a pipe,” effectively mapping intuitions about water volume and pressure onto the more abstract concepts of electrical current and voltage [6]. The use of this pedagogical scaffold, however, does not imply that electrical current is in reality the flow of water particles. Certainly, the expert physicist may call on such

metaphors while instructing a naïve student, or may subtly deploy these evocative images during heuristic reasoning, but all the while they might recognize that electrical current is *fundamentally different* from water flow. Expert practice requires the careful amendment of pedagogical metaphors. The learning of physics — one story goes — is marked by the gradual abandonment of these metaphorical construals, replacing such didactic scaffolds with genuine intuitions about basic physical phenomena.

Thus, while the evidence for dynamic gesture in mathematical pedagogy and communication is suggestive, it does not directly address the nature of mathematical practice — or of mathematics itself. Indeed, to date there is little research on co-speech gesture during the activities that are central to research mathematics, such as proving and communicating non-trivial results. When mathematicians are generating a proof or communicating with other expert mathematicians, do they deploy the same conceptual metaphors that are evidenced in corpus studies of mathematical discourse and in gesture studies of mathematical pedagogy? Or is the metaphorical content of these utterances an artifact of the pedagogical context?

The current study uses the tools of gesture studies, cognitive linguistics, and embodied cognition to empirically investigate with quantitative methods the cognitive reality of fictive motion in mathematical practice. In a semi-controlled situation, we looked at the co-speech gesture of graduate mathematics students as they collaborated in pairs on a mathematical proof involving key concepts in analysis. If the meaning of such essential mathematical concepts is determined by their formal definition, then we should expect static co-speech gesture. If mathematical concepts are truly metaphorical and dynamic, on the other hand, we should expect the co-speech gesture of the graduate students to reflect this dynamism.

4 METHOD

4.1 Participants and Procedure

Twelve graduate mathematics students from a large American research university were paid to participate. Students worked in pairs to prove a fixed point theorem:

Theorem 1 *Let f be an strictly increasing function from $[0, 1]$ to $[0, 1]$. Then there exists a number a in the interval $[0, 1]$ such that $f(a) = a$.*

We selected this problem for our study because the mathematician who proposed it reported that when proving the theorem, he experienced a palpable sense of motion⁴. Further, the problem involves crucial concepts such as limits, continuity, and increasing functions. The proving of this theorem, therefore, presents an excellent opportunity to investigate the cognitive reality of metaphor and fictive motion.

Participants had up to 40 minutes to solve the problem, working alone in a room with a blackboard. Once they were satisfied with their proof or the 40 minutes had passed, the participants explained their proof to the experimenter. The entire session was video-recorded.⁵

4.2 Coding

Based on previous research in Cognitive Linguistics [29, 25, 21], we generated a list of lexical items thought to elicit fictive motion

⁴ We thank Guershon Harel for proposing this theorem

⁵ The experimenter also conducted pre- and post-proof interviews with the participants, but these were not used in the present study.

³ The number-line is itself a *conceptual blend*, the result of combining the mental spaces for number and space.

or metaphorical construals. These included mathematical terms (e.g. function, continuity, limit, contain), verbs of motion (e.g. to cross, to move, to jump), and spatial terms (e.g. up, between, left). The gesture production of each participant was examined for representational gestures that were co-produced with lexical items on this list.

The production of gesture, as co-speech motor *action*, necessarily involves dynamic movement. To differentiate between gestures for which dynamism is an artifact of gesture production, and those gestures for which dynamism is truly expressive, we devised a coding scheme that attended to details of the motion and timing of the co-speech gesture. A gesture was coded as dynamic if it used smooth, unbroken motions; as static if it consisted of beats and segmented motions, or of a smooth motion bookended by beat gestures; and as ambiguous if it was difficult to fit into either of these categories. This coding scheme is very similar to one used in the literature with good reliability to classify “continuous” and “discrete” representations [1].

5 RESULTS AND ANALYSIS

5.1 General results and statistics

Every dyad arrived at a reasonable solution that included most of the ingredients of a complete proof. In all cases, participants’ proofs followed the same outline as the proof supplied by the mathematician who proposed the problem. Half of the dyads finished before the end of the allotted time. All six dyads make extensive use of diagrams — primarily graphs of functions — in combination with extensive symbolic inscriptions, usually involving $\epsilon - \delta$ and set notation.

Participants produced a large number of representational gestures. Every participant but one produced representational gestures that were co-timed with the specified lexical items. A total of 166 of these co-timed representational gestures were coded, for a mean of 13.8 coded gestures per participant. Of these, the majority were coded as dynamic (50.6%); slightly less were coded as static (41.6%).⁶ Figure 1 summarizes gesture production by participant.

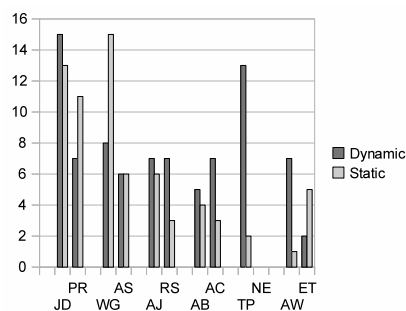


Figure 1. Breakdown of gesture production by participant. Paired participants are listed in adjoining columns.

It bears noting that the majority of gestures produced by the participants during the proof session were deictic, anchored to inscriptions on the blackboard. Deictic gestures appeared to play a number of roles, including maintaining attention during particularly complex deductions and directing the attention of a collaborator to a salient inscription. These indexical gestures, however, were ignored by the present study.

⁶ The remaining gestures were ambiguous.

Gesture production varied according to the co-occurring concept. Certain concepts were associated with a prevalence of dynamic or static gestures. Gestures co-produced with talk of “increase,” “continuity,” and “intersection” were more often dynamic; those co-produced with talk of “containment” and “closeness” were more often static. See Figure 2.

To test the statistical significance of this result, we examined those participants who produced gestures that co-occurred with these concepts and calculated the proportion of dynamic gestures. Among those participants who gestured while speaking of “increase,” co-produced gestures were significantly more often *dynamic* ($F(1,10) = 28.90, p = .0003$). Participants produced a significantly higher proportion of *static* gestures, on the other hand, while discussing “containment” ($F(1,12) = 6.75, p = .0232$) and “closeness” ($F(1,10) = 76.73, p < .0001$).

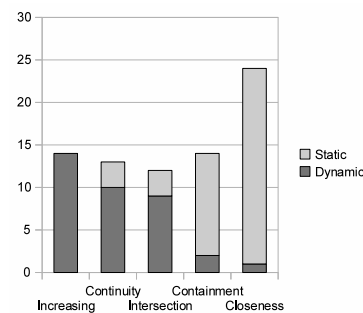


Figure 2. Number of dynamic and static gestures by co-occurring concept

5.2 Examples of dynamic and static co-speech gesture

The following are examples of dynamic and static representational gestures co-produced with coded lexical items.

5.2.1 Dynamic gesture

The concept of “increase” seemed to demand an exclusively dynamic treatment. Participants produced fourteen representational gestures that co-occurred with the lexical items “increasing” or “increase,” all of which were dynamic.

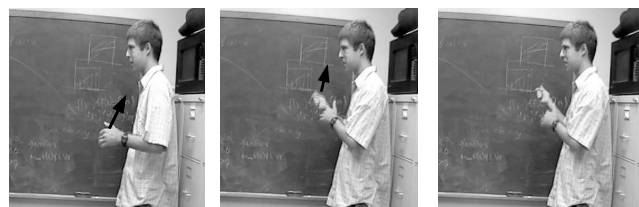


Figure 3. Example 1. Dynamic gesture produced while describing an increasing sequence. Both hands evoke the sequence’s fictive motion.

In Example 1 (see Figure 3), the participant is discussing “increasing sequences” and produces a smooth, unbroken motion, co-timed with his speech. At the onset of the word “increasing,” he begins to

fluidly move his left hand upwards and toward the right, with his thumb pointing in the direction of motion (Figure 3, first frame). As he reaches the end of the word, the motion of his left hand slows slightly while his right hand begins to accelerate, once again moving upwards and to the right in the direction of his extended thumb. As he begins to say “sequence,” his right hand reaches its top speed (Figure 3, second frame). Both hands begin to slow to a stop (Figure 3, third frame), and their retraction is co-timed with the end of “sequences.” Neither his gaze nor his thumbs are directed toward a blackboard inscription.

In Example 2 (Figure 4), the student is at the blackboard, writing a series of inequalities that contradict the assumption that the function is increasing. As he finishes writing, he steps back from the blackboard, drops his hands to his side, and states, “So that contradicts uhhh increasing” (Figure 4, left frame). Precisely co-timed with the onset of “increasing,” his right hand flies upwards and to the right, with his index finger extended in a prototypical pointing handshape (Figure 4, middle frame). As he finishes saying “increasing,” his right hand slows, pointing to the right of the blackboard for a moment (Figure 4, right frame), before smoothly dropping to his side. This is a quintessential example of dynamic co-speech gesture, consisting of a continuous motion co-timed with a dynamic lexical item.

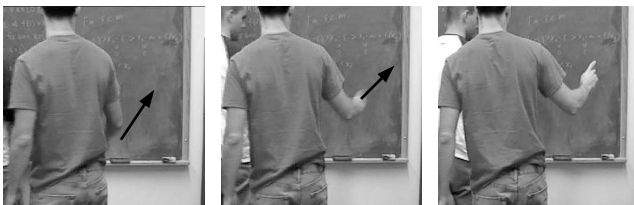


Figure 4. Example 2. Dynamic gesture during discussion of an increasing function. A fluid rightward hand movement is co-timed with speech.

5.2.2 Static gesture

Static gestures often accompanied discussions of closeness and containment. In Example 3 (Figure 5), the student is discussing the values of a function, and wants to consider only those values contained within a restricted region. At first, all her gestures are deictic, anchoring her discussion to the graph she has drawn on the blackboard. She introduces the region of interest by saying, “Well, if you look at a, sort of, small enough [region].” When she begins to describe the region of interest (“Well...”), she retracts her hands from the blackboard and positions them in front of her chest, pointing towards each other, and pauses as she says “sort of” (Figure 5A). Co-timed with the onset of “small,” she quickly moves her hands toward each other, stopping when they are 10 cm apart, then quickly retracting her hands to their original distance (Figure 5A-C). Co-timed with the word “enough,” she repeats the same inward staccato stroke, stopping abruptly when her hands are at the same distance, and retracting once again (Figure 5D-E). By indexing two particular points in space with a repeated beat, she evokes the endpoints of a delimited region containing the function’s domain. This is a static representational gesture since it consists of distinct beats, indexing exact points in space.

5.2.3 Variable gesture

Often, a single utterance was amenable to very different representations in gesture. The utterance “to the left,” for instance, received

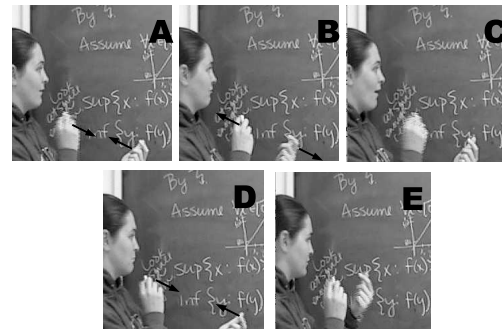


Figure 5. Example 3. A static gesture co-timed with the utterance, “small enough.”

both dynamic and static treatments in gesture. See Figure 6.

In the first frame of Figure 6, the participant is producing a static gesture co-timed with the utterance “to the left.” One participant suggests to his collaborator that they should look for “implicit continuity to the left of what [they]’re talking about.” As he says “to the left,” he forms his thumb and forefinger into a U-shape, representing the interval to the left, performs one forward beat with this handshape, and then holds his hand still for nearly a full second. This gestural representation of “to the left” captures the static notion of containment.

The second frame of Figure 6 shows the same lexical affiliate (“to the left”) co-timed with a *dynamic* gesture. While discussing the intersection of a function and the line $y = x$, a participant asks if the function was “going a little bit to the left.” As he begins to say “to the left,” he points his forefinger to the left and fluidly moves his hand in that direction, retracting his hand as he says “left.” This gesture seems to represent a dynamic sense of the fictive motion for the function.

These two cases illustrate that gestures may share a lexical affiliate but differ in their dynamism. Although highly synchronous with the utterance “to the left,” these two gestures exhibit different kinetics and handshapes.

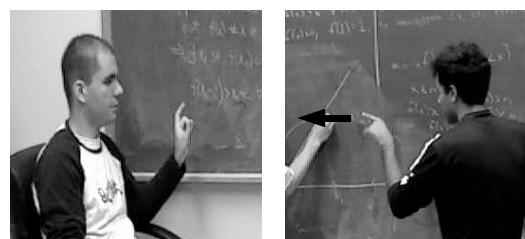


Figure 6. Example 4. Two gestures co-produced with the utterance, “to the left.” While the gesture on the left is static, the one on the right is dynamic.

Finally, gesture also varied by mode of reasoning, as documented in Figure 7. While engaged in diagrammatic reasoning — either considering or producing a diagram on the blackboard — most participants were more likely to produce a dynamic gesture. Formal symbolic reasoning, on the other hand, was associated with slightly higher rates of static gesture. Verbal reasoning was associated with a mix of dynamic and static gestures.

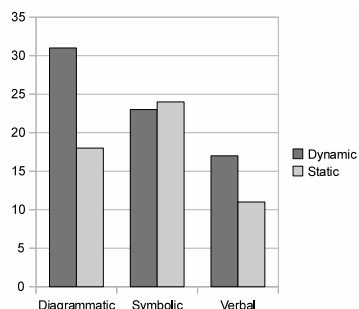


Figure 7. Number of dynamic and static gestures by mode of reasoning

5.3 Discussion

The results presented here support our hypothesis that dynamic representational gesture is widespread and recurrent in mathematical practice. Moreover, dynamism in gesture is associated with fictive motion in speech. Certain mathematical lexical items were particularly amenable to co-production with dynamic gesture, while others were most often treated statically. Overall, these results demonstrate that dynamic, metaphorical gesture is not restricted to pedagogical or elementary mathematical contexts, but is an essential ingredient in expert mathematical practice.

Investigating the nature of the cognitive processes undergirding mathematical cognition and gesture production requires the deployment of quantitative methods. Goldin-Meadow and her associates have successfully employed statistical methods, but these studies have been limited to simple arithmetic reasoning [1, 7, 10, 9]. Contrary to the results of the present study, they concluded that gestures have *propositional* mental representations. However, they elicited gesture during the explanation of a simple arithmetic task, and the elicited gesture consisted almost entirely of deictic gestures anchored to algebraic blackboard inscriptions. Our own previous research on dynamic gestures in advanced mathematics has shown that the explanation of advanced mathematics is often accompanied by dynamic, metaphorical gestures, but these studies were largely qualitative and opportunistic [21, 22]. This study is, to the best of our knowledge, the first to investigate quantitatively co-speech gesture during expert mathematical practice.

The present study only coded representational gestures, purposefully disregarding environmentally-coupled deictic gestures. Did this unnecessarily ignore an essential component of mathematical gestural behaviour? Pointing at the graph of a function, for instance, was not coded, even if the hand dynamically swept along the curve of the graph. This aspect of the coding scheme represented a conscious decision to sacrifice breadth of analysis for experimental traction and focus. The coding scheme successfully discriminated between mathematical gestures that expressed mathematical content (e.g., “The function is increasing”), and meta-mathematical gestures that conveyed something about the local practice of mathematics (e.g. “Let’s move the diagram a bit higher on the blackboard”). While deictic gestures are undeniably an important part of mathematical practice, particularly when that practice involves collaboration at a blackboard [28, 7], such gestures are referentially ambiguous, anchored either to the inscription itself or to the mathematical entity represented by that inscription [12]. By restricting our attention to representational gestures, we avoided this complication. Future research will explore the

ways in which environmentally-coupled gestures elaborate meaning and direct local practices.

How are we to understand the various ways in which participants used gesture to represent mathematical concepts? Kendon [13] distinguishes three varieties of gestural representation: enactment, depiction, and modeling. In *enactment*, the motor action is meant to reproduce some features of the activity being represented. For example, gesturing in the air as if you were drawing a graph would be an enactment of the act of drawing. None of the twelve participants produced enactments that were not directly coupled to the environment (e.g. re-tracing a graph with the chalk barely off the blackboard), and thus no enactments were coded in this study. In *depiction*, the gesture “creates an object in the air” [13, p.160], often using the index finger to trace an object’s shape. Example 3 above, where the participant uses a sequence of staccato beat gestures to delimit a region in space, and the first frame of Example 4, in which the handshape stamps out a region “to the left,” are both depictions of numerical intervals that exploit the metaphorical mapping between number and space [23]. In *modeling*, the gesturing body part stands in for another object, as when a fist represents a stone. Both specimens of dynamic gesture above, Examples 1 and 2, involve the modelling of a mathematical concept — and, in particular, of a fictive trajectory. In Example 1, the motion of the hands is modeling the dynamic increase of the sequence, and the hands themselves are standing in for the trajectory of that fictive motion, perhaps understood as some element of the sequence. In Example 2, the hand is modeling the trajectory that travels along the path of a function. For both these examples, the hands’ index fingers indicate the direction of the trajectories, constituting, in a way, a vector representation of the trajectory’s fictive motion.

Critics of Cognitive Linguistics often object that metaphorical expressions in mathematical discourse may be “dead metaphors,” expressions that once reflected underlying psychological processes but that are now entirely conventional [20, 21]. Certainly, mathematical co-speech gesture exhibits certain signs of conventionalization. In the two examples of gestures co-produced with the lexical affiliate *increasing*, both participants exhibited a marked rightward hand trajectory, mirroring the contingent fact that, historically, graphs of functions have been drawn from left to right. Is this evidence for conventionalization? Hardly. Although the direction of a numerical-spatial association may have a historical or cultural origin [23], this does not deny its cognitive reality. Indeed, the very process of associating number and space may require the cognitive mechanisms of conceptual blending and metaphor [23], and a similar conceptual association between number and space has been shown to have cognitive reality in the context of arithmetic reasoning [18, 27, 15]. This shows that even a near-universal construal requires individual cognitive elaboration and has measurable implications for reasoning. Furthermore, as we saw above, the phrase “to the left” was treated dynamically in one context, and statically in another. If this phrase were merely a conventional way of referring to certain intervals — as when we say the negative numbers are “to the left” of zero — then we would not expect a co-produced dynamic gesture under any circumstances. To the contrary, gesture in this case was dynamic, as it was in many others. The presence of dynamism in gesture, therefore, is evidence that these metaphorical expressions are not “dead,” but have a cognitive reality.

Since the construction of mathematical diagrams is necessarily a dynamic process, involving the tracing of chalk across a blackboard, one might wonder if the dynamism of gesture is a mere echo of the dynamism of the inscriptive motion. Certainly, the motion that inevitably accompanies the creation of a mathematical diagram is

probably a factor in the historical and developmental origin of fictive motion for such notions as continuity, function, and limit. The contemporary dynamism of mathematical thought, however, is incredibly robust. Mathematical discourse is rife with fictive motion in the absence of diagrams altogether, as demonstrated by corpus studies of purely formal analysis textbooks [25, 24]. In Example 2 above, the participant produces a dynamic gesture after writing a series of static inequalities; the blackboard is covered with static notations, with not a diagram in sight, and yet his gesture reveals the dynamism of his understanding of increasing sequences. While a graph of a function is visible in Example 1, the participants have not recently attended to the diagram; they are discussing and symbolically representing the limit of the function. Moreover, in Example 1, the participant's gesture moves orthogonally to the orientation of the blackboard diagram. The dynamism evinced in gesture is not merely an echo of inscriptive dynamism, therefore, but evidence for the participants' dynamic construals.

Although certain dynamic lexical items were more often co-produced with dynamic gesture, this gestural dynamism was not mandatory. Indeed, dynamic lexical items were sometimes paired with static gesture, and in those cases where dynamic gesture was absent, the cognitive reality of the fictive motion remains an open question.

6 CONCLUSION

Contrary to the traditional view of mathematics — in which concepts acquire their meaning from formal definitions — the results of this study provide quantitative evidence for the claim that certain mathematical concepts are inherently metaphorical and dynamic. Participants' gesture production corroborated earlier results derived using the methods of Cognitive Linguistics [21, 22]. While previous studies of mathematical gesture production have been largely descriptive and opportunistic, this study attempted to quantify the prevalence of dynamic gesture during mathematical practice. Furthermore, this is one of the first studies to investigate gesture during advanced mathematical reasoning; previous studies focused on pedagogy or very simple mathematical problems.

Goldin-Meadow and her colleagues have shown that, in the context of elementary arithmetic, gesture can help us infer a speaker's solution strategy [5]. We found that certain modes of reasoning were associated with a preponderance of dynamic gestures, suggesting that, even in advanced mathematics, gesture may provide insight into a speaker's reasoning strategy. A speaker's topic and mode of reasoning, however, did not perfectly predict their gesture production. When speaking of a concept that exhibits fictive motion, when exactly is a speaker apt to gesture dynamically? Further research is required to identify the circumstances under which mathematical practice exploits various multimodal resources, including but not limited to co-speech gesture.

Traditional, foundationalist accounts of mathematics have focused exclusively on the *products* of mathematical activity, while ignoring the rich local practices that contribute to and, arguably, constitute the real content of mathematics. Kendon [14, p. 357] argues that a similar myopia leads to a purely formalist account of *language*, where "what is transferred to paper is abstracted away from what is actually done within an enacted utterance," and as a result essential aspects of communicative practice such as "intonation, voice quality [...] not to mention kinesics," or gesture, are ignored. In attending exclusively to formal, disembodied theorems and proofs — products of the sedimentation of local mathematical practices — the study of mathemat-

ics has ignored the rich meaning-making practices of flesh-and-blood mathematicians, and collapsed multi-agent and multimodal practice [11] into a single idealized agent working within a single written modality. In order to account for the exceptional traits of mathematics — objectivity, necessity, precision, stability — we must remember that, "Of course, in one sense, mathematics is a body of knowledge, but still it is also an *activity*" [30, §349].

The present study is a direct response to this ahistorical tradition and a contribution to an emergent study of mathematics that does justice to mathematicians and mathematical activity. This recent turn towards practice has multiplied the modalities through which we can study the nature of mathematics. By including gesture in that analysis, we take the mathematician's body seriously as a semiotic resource in the creation and communication of mathematical content.

ACKNOWLEDGEMENTS

The current results are drawn from video data collected during a larger collaborative research project on the conceptual basis of proof funded by the Spencer Foundation. The principal investigator on this project was Laurie Edwards, Professor of Education at Saint Mary's College; the co-principal investigators were Guershon Harel, Professor of Mathematics at the University of California, San Diego, and the second author of this paper. We would like to thank Molly Kelton, Kensy Cooperrider, D Doan and Anastasia Nikoulina for their insightful comments on earlier versions of these results. We thank two anonymous referees for constructive and encouraging feedback.

REFERENCES

- [1] Martha Alibali, Miriam Bassok, Karen Olseth Solomon, Sharon Syc, and Susan Goldin-Meadow, 'Illuminating mental representations through speech and gesture', *Psychological Science*, **10**, 327–333, (1999).
- [2] A. Cienki, 'Metaphoric gestures and some of their relations to verbal metaphoric counterparts', in *Discourse and Cognition: Bridging the Gap*, ed., J. P. Koenig, Stanford, CA, (1998). CSLI.
- [3] L. D. Edwards, 'Gestures and conceptual integration in mathematical talk', *Educational Studies in Mathematics*, **70**(2), 127–142, (2009).
- [4] Gilles Fauconnier and Mark Turner, *The Way We Think*, Basic Books, 2002.
- [5] Philip Garber and Susan Goldin-Meadow, 'Gesture offers insight into problem-solving in adults and children', *Cognitive Science*, **26**, 817–831, (2002).
- [6] D. Gentner and Dr. R. Gentner, 'Flowing waters or teeming crowds: Mental models of electricity', in *Cognitive functions: Classic readings in representation and reasoning*, eds., D. Gentner and A.L. Stevens, 99–129, Greenwich University Press, (1983).
- [7] Susan Goldin-Meadow, Susan Wagner Cook, and Zachary Mitchell, 'Gesturing gives children new ideas about math', *Psychological Science*, **20**, (2009).
- [8] Susan Goldin-Meadow, San Kim, and Melissa Singer, 'What the teacher's hands tell the student's mind about math', *Journal of Educational Psychology*, **91**, 720–730, (1999).
- [9] Susan Goldin-Meadow, Howard Nusbaum, Spencer D. Kelly, and Susan Wagner, 'Explaining math: Gesturing lightens the load', *Psychological Science*, **12**, 516–522, (2001).
- [10] Susan Goldin-Meadow and Susan Wagner, 'How our hands help us learn', *TRENDS in Cognitive Sciences*, **9**, 234–241, (2005).
- [11] Charles Goodwin, 'Action and embodiment within situated human interaction', *Journal of Pragmatics*, **32**, 1489–1522, (2000).
- [12] Edwin Hutchins and Leysia Palen, 'Constructing meaning from space, gesture, and speech', in *Discourse, Tools, and Reasoning: Situated Cognition and Technologically Supported Environments*, eds., L.B. Resnick, R. Saljo, C. Pontecorvo, and B. Burge, chapter 1, 22–40, Springer-Verlag, Germany, (1998).
- [13] Adam Kendon, *Gesture: Visible Action as Utterance*, Cambridge University Press, 2004.

- [14] Adam Kendon, 'Some reflections on the relationship between 'gesture' and 'sign'', *Gesture*, (2008).
- [15] André Knops, Bertrand Thirion, Edward M. Hubbard, Vincent Michel, and Stanislas Dehaene, 'Recruitment of an area involved in eye movements during mental arithmetic', *Science*, **324**, 1583–1585, (2009).
- [16] George Lakoff, 'The contemporary theory of metaphor', in *Metaphor and Thought*, ed., Andrew Ortony, Cambridge University Press, 2 edn., (1992).
- [17] George Lakoff and Rafael Núñez, *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*, Basic Books, 2000.
- [18] Koleen McCrink, Stanislas Dehaene, and Ghislaine Dehaene-Lambertz, 'Moving along the number line: Operational momentum in nonsymbolic arithmetic', *Perception and Psychophysics*, **69**, 1324–1333, (2007).
- [19] David McNeill, *Hand and mind*, University of Chicago Press, 1992.
- [20] Gregory Murphy, 'On metaphoric representation', *Cognition*, **60**, 173–204, (1996).
- [21] Rafael Núñez, 'Do real numbers really move? Language, thought, and gesture: The embodied cognitive foundations of mathematics', in *18 Unconventional Essays on the Nature of Mathematics*, ed., Reuben Hersh, chapter 9, Springer, (2006).
- [22] Rafael Núñez, 'Mathematics, the ultimate challenge to embodiment: Truth and the grounding of axiomatic systems', in *Handbook of Cognitive Science*, eds., O. Calvo and T. Gomilla, chapter 17, 333–353, Elsevier, (2008).
- [23] Rafael Núñez, 'No innate number line in the human brain', *Journal of Cross-Cultural Psychology*, (in press).
- [24] Rafael Núñez, L. D. Edwards, and J. F. Matos, 'Embodied cognition as grounding for situatedness and context in mathematics education', *Educational Studies in Mathematics*, **39**, 45–65, (1999).
- [25] Rafael Núñez and George Lakoff, 'What did Weierstrass really define? The cognitive structure of natural and epsilon-delta continuity', *Mathematical Cognition*, **4**, 85–101, (1998).
- [26] Rafael Núñez and Eve Sweetser, 'With the future behind them: Convergent evidence from aymara language and gesture in the crosslinguistic comparison of spatial construals of time', *Cognitive Science*, **30**, 401–450, (2006).
- [27] Michal Pinhas and Martin Fischer, 'Mental movements without magnitude? A study of spatial biases in symbolic arithmetic', *Cognition*, **109**, 408–415, (2008).
- [28] Nathaniel Smith, *Gesture and beyond*, Unpublished manuscript, University of California, Berkeley, 2003.
- [29] Leonard Talmy, *Towards a Cognitive Semantics*, volume 1, chapter 3: How Language Structures Space, The MIT Press, 2000.
- [30] Ludwig Wittgenstein, *Philosophical Investigations*, Wiley-Blackwell, 4th edn., 2009.

If Learning Maths Requires a Teacher, Where did the First Teachers Come From?

Aaron Sloman¹

<http://www.cs.bham.ac.uk/~axs>

Abstract. This is the latest progress report on a long term quest to defend Kant's philosophy of mathematics. In humans, and other species with competences that evolved to support interactions with a complex, varied and changing 3-D world, some competences go beyond discovered correlations linking sensory and motor signals. Dealing with novel situations or problems requires abilities to *work out* what can, cannot, or must happen in the environment, under certain conditions. I conjecture that in humans these products of evolution form the basis of mathematical competences. Mathematics grows out of the ability to use, reflect on, characterise, and systematise both the discoveries that arise from such competences and the competences themselves. So a "baby" human-like robot, with similar initial competences and meta-competences, could also develop mathematical knowledge and understanding, acquiring what Kant called synthetic, non-empirical knowledge. I attempt to characterise the design task and some ways of making progress, in part by analysing transitions in child or animal intelligence from empirical learning to being able to "work things out". This may turn out to include a very general phenomenon involved in so-called "U-shaped" learning, including the language learning that evolved later. Current techniques in AI/Robotics are nowhere near this. A long term collaborative project investigating the evolution and development of such competences may contribute to robot design, to developmental psychology, to mathematics education and to philosophy of mathematics. There is still much to do.

1 Mathematics as a Biological Phenomenon

1.1 Evolved competences and meta-competences.

I once met a software engineer who poured scorn on evolution as a designer because it had produced no animals with wheels. I have not met one who poured scorn on evolution for designing human neonates to be helpless while many other species have self-sufficient neonates: e.g. chicks get themselves out of the egg and soon peck for food, while a new foal or calf soon gets up on four legs and finds its way to the mother's nipple. It is commonly supposed that humans are born so immature because larger skulls would kill their mothers, but evolution produced some mammals with larger heads, e.g. elephants. Another common suggestion is that in an environment that is continually changing, learning capabilities are more important than innate knowledge: but that does not explain why evolution does not produce both in the same species, namely a collection of innate modules that control various aspects of behaviour along with a general purpose learning module (one version of the so-called "modularity" thesis).

An alternative answer is that in addition to those competences required from birth (e.g. breathing, sucking, digesting food, doing various things that attract attention of parents) there are special-purpose, more abstract, *meta-competences* for acquiring competences (e.g. for acquiring competences involved in use of independently movable manipulators, such as hands, claws, or trunk), where the perceptual and motor processes are so complex and varied that they cannot be encoded in the genome (possibly for reasons suggested in [33]). On this alternative, some behaviours are produced by sophisticated run-time control competences that have to be learnt, using specialised meta-competences (for learning) that also develop through the influence of the environment on meta-meta-competences: throughout life, humans go on learning how to learn many different things.

Sloman and Chappell [5] conjectured that humans are born with, among other things, sophisticated, evolved, *meta-knowledge about how to learn in a complex 3-D world containing more or less complex structures and processes, opportunities and obstacles, and also other intelligent individuals*. This may include learning how to grow layers of meta-competence suited to increasingly complex and/or abstract features of the environment, through play and exploration. The innate/learned dichotomy does not do justice to such diversity.

This paper extends those ideas by providing some (informal) empirical observations and theoretical conjectures about layered development and learning, that may lead to further progress in understanding how mathematics fits in with the rest of animal cognition, including precursors in human infants and toddlers, and in other animals, unlike more widely studied, allegedly innate, precursors concerned with recognition or estimation of objecthood, numerosity, size, etc. [41].

1.2 Philosophy of mathematics.

This proposal is related to the philosophy of mathematics of Immanuel Kant [15], according to whom mathematics is neither empirical (as suggested by Mill [22]), nor merely about "relations between ideas" (as suggested by Hume). But that does not mean mathematical knowledge is innate, or that mathematical competences are infallible. Neither is the nature of mathematical knowledge usefully described in terms of "feelings of compulsion" [1], e.g. since mathematical discoveries may be tentative conjectures at first. Rather the structure, kinematics and dynamics of the 3-D world generally requires generalisations and predictions to be tested *empirically*, but includes features that allow powerful *short-cuts* (using reasoning) to be discovered and used in creative problem-solving [6]. This makes possible biologically useful *productive laziness* which reduces or eliminates the need for empirical trial and error, or symbolic searching, and sometimes provides relatively quick solutions for

¹ University of Birmingham, UK, email: a.sloman@cs.bham.ac.uk

novel practical problems, avoiding empirical experiments that could be dangerous, or even fatal. However, the mechanisms are not infallible: testing and debugging may be required, as Lakatos [17] demonstrated.

In humans, these products of evolution develop into what we understand as mathematical competences. (Examples are given below.) Many of species-specific details are of psychological interest, but that does not mean that the subject matter of mathematics is psychological: it is primarily concerned with the structure of the world, and also how that structure can be discovered, and used, leading to the study of alternative possible structures.

1.3 Russell vs Feynman on Mathematics.

Russell and Whitehead [43] attempted to demonstrate that all of mathematics could be reduced to logic (Frege had attempted this only for Arithmetic). Despite the logical paradoxes, Russell thought the goals of the project could be or had been achieved, and concluded that mathematics was just *the investigation of implications that are valid in virtue of their logical form, independently of any non-logical subject matter*. He wrote: “Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true” (*Mysticism and Logic* [25]).

In contrast, Feynman described mathematics as “the language nature speaks in”. He wrote: “If you want to learn about nature, to appreciate nature, it is necessary to understand the language that she speaks in” [7]. I believe that Feynman’s description is closely related to generally unnoticed facts about how a child (and perhaps future robots) can develop powerful, reusable concepts and techniques related to patterns of perception and patterns of thinking that are learnt through interacting with a complex environment, part of which is the information-processing system within the learner [28, Ch 6]. Despite the role of experience in such learning, the *results* of such learning are not empirical generalisations. Kant wrote, in [15], “though all our knowledge begins with experience, it by no means follows that all arises out of experience.” Feynman seems to agree.

1.4 Learning Mathematics and Learning Language.

I suspect that some of the re-conceptualisation involved in developing mathematical insight uses mechanisms produced by evolution for unifying disparate empirical discoveries in a deeper, more general, explanatory theory. A special case of this, which evolved later, and had to be modified to allow exceptions to explanatory theories, seems to be the mechanism for language learning that allows empirical learning of useful patterns of verbal communication after a time to trigger the development of a rule-based syntactic theory that is far more general in its predictive and explanatory power than the previously learnt patterns. In children this leads to “U-shaped” learning² because at first the new system cannot handle exceptions, so incorrect utterances are produced (e.g. “she spoke”, “he hit me”) by a child who earlier used the correct forms (“she spoke”, “he hit me”). Later the architecture is extended to cope with exceptions (a non-trivial alteration to an information-processing system) and both previously known and newly learnt exceptions to the rules are dealt with correctly. Further reorganisation may occur as the language changes.

This ability to develop a powerful (implicit) generative theory of the language in use, replacing previously acquired empirical generalisations, may be a specialised version of mechanisms that evolved earlier because they support productive laziness in coping with a physical environment. Language is not the only aspect of

the world that has general laws with exceptions that have to be accommodated. But some aspects are exceptionless, including things learnt about topological features of space and time and some of the metrical features. There are also exceptionless generalisations about numbers and about logical forms, which can be learnt later and embedded in a powerful theory.

The possibility of empirical learning being supplemented by more powerful mechanisms supporting “productive laziness” does not seem to be widely recognised. For example, many examples of learning about affordances in young children are described in [8]. The authors, like many others, assume there are only three ways for a child to have knowledge about what can and cannot occur, namely either it is innate, or it is acquired by learning associations by observation and trial and error, or the information is acquired from someone else, by imitation or being instructed. They, like many others, ignore a fourth way: learning by *working things out*. Having an appropriate theory about some aspect of reality allows a child (or animal, or physicist, or engineer) to *work out* what must happen even in some situations that have never previously been encountered (e.g. dealing with new spatial configurations) whereas if the learning were purely empirical, every novel situation would have to be tested before predictions could be relied on. Examples are given in the next section. Systematic collection and testing of examples in young children and other species could form a major interdisciplinary research project. Finding explanations, producing working models of the processes, and explaining how neural mechanisms can support both the processes of development and the processes of discovery is a bigger challenge. Meeting that challenge is a requirement for explicating and defending Kant’s philosophy of mathematics – though that will not be obvious to most philosophers.

From this standpoint, someone composing a poem or story is working out what can be thought or said, using mechanisms that are partly similar to mechanisms used by someone working out what can be done in the environment, or working out what structures and processes can and cannot exist in the environment, or what relationships can and cannot hold between numbers. All of these require learning through interaction between meta-meta-(meta-...) competences produced during evolution of the species, and information provided by the individual’s environment.

2 Examples of “Working Things Out”

The ability to learn by creatively working something out is being actively studied in animal behaviour research, a spectacular example being Betty the hook-making crow (search for “betty+hook+crow” for videos). Paradoxically the phenomenon has escaped the notice of many researchers investigating human children, mainly because of the wide-spread assumption that their learning is mainly Bayesian. My examples are mainly from human learners – young and old.

2.1 Examples of “Productive Laziness”

(1) A deceptively obvious, though quite sophisticated, example concerns invariant results of sequential matching operations. A child may discover empirically (and with surprise and delight) that counting the same set in different orders (e.g. counting fingers on a hand from left to right and from right to left) always produces the same result – if no items are added/removed or counted twice, omitted from the count etc. This can start off as an empirical discovery and later become obvious (at least for some learners). How? Is it merely that after finding much evidence to support the generalisation learners associate a very high *probability* with it?

² <http://www.cis.udel.edu/~case/slides/nugget-ushape.pdf>

That would be rash, since the evidence would be limited to only a tiny subset of possible cases: what justifies the extrapolation to all the very many sets of objects much larger than any the learner has encountered? Can you give a proof? Some philosophers believe that the result is not proved – merely stipulated as a criterion for doing the counting properly. (They have presumably never experienced the process of making a mathematical discovery.) Few mathematical results have only one proof. This result could be understood on the basis of transitivity of one-to-one correspondences, shifting the problem – an exercise for the reader.

(2) Another example can occur after the child has learnt that there are different kinds of matter, some rigid, some not. It is possible to discover empirically that if a part of a rigid object is grasped and rotated all the other parts of the object will move, and their movements will be greater the further they are from the rotated portion. Later on comes understanding that rigidity (which is a local property of resistance to deformation) can have global effects, so that movement of part of object can cause all other parts to move, and that rotary movements have amplified remote effects. The reasoning can be applied to objects of many shapes and sizes in many contexts, and makes it possible to predict consequences of rotation and other movements, and also constraints on such movements (assuming rigidity and impenetrability of the materials). This sort of discovery can, of course, have practical consequences that are important.

(3) A child who can lift a cut-out picture from its recess may know which recess it should go back into but fail to insert it without a lot of random movement. Later, the child discovers the need to *align the boundary of the picture with the boundary of the recess*. That may require a non-trivial extension of the child's ontology to include concepts of boundary and alignment. If the boundaries are aligned and the picture piece is slightly smaller than the recess then all parts of the piece can simultaneously avoid obstruction from the surrounding material. Any slightly different position causes overlap that prevents insertion. Unless the recess and picture piece are both circular, most *rotations* will cause obstruction. However if the piece is symmetrical there may be several rotations that allow insertion, but most rotations will cause obstruction: why?

A child who can reliably get such pictures back into their recesses, may at first merely attach a very low probability of success to insertion without boundary alignment then later come to see that the insertion *must* fail, but without being able to explain that this impossibility depends on the materials being rigid and impenetrable. What changes between noticing an empirical regularity and seeing that it cannot have exceptions?

(4) There is much to learn about strings. If an inelastic but flexible string is attached to a remote movable object, then if the end is pulled away from the object a process can result with two distinct phases: (1) curves in the string are gradually eliminated (as long as there are no knots), and (2) when the string is fully straightened the remote object will start moving in the direction of the pulled end. However, if the string is looped round a fixed pillar, the first sub-process does not produce a single straight string but two straight portions and a portion going round the pillar, and in the second phase the attached object moves toward the pillar, not toward the pulled end; and so on. Many more complex variants can be reasoned about.

Many more examples including string, elastic, pins, stones, mazes, and marks on paper are described in [26]. Some of their experiments can be performed on *simulated* physical objects, e.g. simulated rubber bands, using a graphical package like 'dia'. E.g. if a *star* is defined to be a polygon with alternating concave and convex vertices, what's the minimum number of pins required to hold a

stretched rubber band in a star shape? This and other problems are discussed in my slides on "toddler theorems" [36]. As discussed in [27] proofs generally work on *representations* of the subject matter rather than *instances* of the subject matter, which would not be possible for empirical knowledge.³ The paper also refutes the common assumption that analogical/pictorial representations are always *isomorphic* with what they represent.

(5) J.J. Gibson [9] drew attention to the biological importance of perceiving not only *structures*, but also *processes* (e.g. optical flow patterns), and being able to detect possibilities for action, and constraints on actions, in the environment: positive and negative affordances for the perceiver. That requires the ability to perceive the *possibility* of some processes and configurations and the *impossibility* of others [30]. Gibson considered only possibilities involving the perceiver's actions. However those are only a subset of a broader range of possibilities and constraints that can be perceived: they can be called "proto-affordances", since they involve possibilities and constraints on processes that need not involve actions, e.g. seeing that an object is too large to fit through a gap, whereas another could get through, as long as it is rotated during the process (e.g. if it is L-shaped). These are *potential* affordances, not *actual* affordances if you don't care about moving the object.



Figure 1. How many different action sequences can do the transformation? There is a huge variety of ways in which using your right hand, left hand, both hands, teeth, a pair of pliers, tweezers, etc., you could rearrange the objects on the left to the configuration on the right. To do this a robot would need forms of representation that: (a) can be derived from sensory input (even low-resolution noisy input, as here); (b) have sufficient definiteness to allow goals to be expressed; (c) allow representation of effective multi-step plans to achieve goals; (d) are sufficiently abstract to allow effective reasoning about future possibilities and impossibilities (e.g. grasping the saucer in the middle, using finger and thumb is impossible), without considering all the continuously variable configurations and processes; and (e) allow plans to be used to control actions (using visual servoing, not necessarily ballistically). An infant cannot do all this but a normal child will eventually develop the ability to do it and later the ability to describe ways to do it. Byrne & Russon [4] showed that some other species seem to have partly similar abilities.

Figure 1 illustrates a vast array of possible processes and constraints on processes involving a small number of objects which can be moved from one configuration to another via a very large collection of different trajectories, using many possible modes of manipulation. Such transformations of 3-D configurations are achievable by children before they can talk, and also by other animals, as discussed in [37]. So the ability to perform them cannot depend on use of human language, though there must be an *internal* form of information processing (a "generalised language") supporting not only perception of what exists but also reasoning about what is and is not *possible* including representation of goals (non-existent states to be achieved) so that action sequences can be thought about and selected before they are performed, and their consequences *worked out*, using general capabilities, even if the task is a new one [29]. Physical trial-and error is not always required, for toddlers, or other animals solving novel problems: they seem to be able to *work things out*: they are productively lazy. (That's not easy!)

(6) A familiar example is stacking nested cups (facing up). At first, as

³ Instances can also function as representations, however!

Piagetians found, children succeed only by a procedure that appears to involve randomly selecting the next cup, and later, when a cup is found too big to insert, removing some of the previously inserted cups and trying again with the omitted cup. Eventually a child may notice that all the back-tracking involves wasted effort and time and go on to “debug” the procedure, until she works out that wasted moves can be avoided if the cups are selected in order of decreasing size: inserting the largest remaining cup at each stage.

A child who cannot *describe* that successful procedure, may nevertheless discover it. Perhaps she has *understood* but is unable to *say*: (a) ‘smaller than’ is transitive; (b) a cup cannot be inserted into a smaller cup; (c) if at some stage a cup X is inserted that is smaller than some other available cup Y, then (d) all the cups inserted after X will be smaller than X (from (a)) and therefore (e) all the cups inserted after X will be smaller than Y, and (f) Y cannot therefore be inserted after X – so that back-tracking will be required. A child who has somehow understood all that may be able to *work out* that always inserting the largest remaining cup will prevent the blockage occurring. A further step is understanding that preventing that blockage will suffice to get all the cups stacked. Yet another step is generalising the result to nesting the cups upside down – requiring a partly similar, but importantly different, strategy.

I doubt that most adults who understand the strategy can *explain why* the strategy works, unless they have had training in computer science or mathematical reasoning. *How* such understanding is implemented is another matter, requiring further research. It seems unlikely that a young child, or even most non-mathematical adults, can use reasoning in predicate calculus, or a computational process-algebra. Rather, I suspect the reasoning uses a form of representation that is more pictorial than logic, but more abstract than normal pictures – features required also for planning multi-step future actions. Compare [27, 10, 14, 45, 36]. Spelling out details remains a hard task for the future.

(7) There are very many examples involving counting and numbers. Much philosophy of mathematics assumes numbers are abstract entities named by numerals, but for a young child, as pointed out (independently) in [44] and [28, Ch 8], number words are primarily used for *doing* things, including reciting them, matching them and using them to perform tasks involving collections of objects (which later can include collections of number names). In addition to example (1) there are many more numerical discoveries that can be made empirically, then later understood as necessarily exceptionless, including the discovery that adding 3 more objects to a set of 4 objects gives the same “How many?” result as adding 4 more objects to a set of 3 objects; then later generalising this so as to understand that adding N objects to a set of M objects gives the same result as adding M objects to a set of N objects. One way to see this is to make the (non-trivial) discovery that if the objects are indistinguishable, and addition is done one object at a time, then adding the larger set to the smaller set necessarily goes through a stage which is the initial state of adding the smaller set to the larger set.⁴ (Contrast the analysis in [12].) Moving from the empirical view to the non-empirical view of these results requires a fairly sophisticated ability to attend to abstract features of *processes* as well as structures, namely parametrisable patterns that can be instantiated in different ways, while (necessarily) preserving some of their structure. This mathematical use of *abstraction* to bring different cases under the same specification, is often confused with use of *metaphor*. Metaphors cannot explain why certain things are possible and others

impossible, though they may suggest explanations.

(8) Links between counting abilities and abilities to manipulate and rearrange objects can be discovered when playing with blocks, e.g. the discovery that some sets of blocks can be arranged in rectangles, or in cubes, while others cannot. E.g. 5 or 7 blocks cannot be arranged as a rectangle, except a rectangle with only one row, whereas 8 or 9 blocks can be arranged in a rectangle (the latter a square). In contrast, 8, but not 9 blocks can be arranged to form a cube, and so on. In very unusual children this exploration might lead to invention of the concept of a prime number or even conjecturing the unique-factorisation theorem for natural numbers.

Additional examples were investigated by Piaget, though as far as I know Piagetians did not ask the questions posed here from an AI standpoint, including: What forms of representation does the learner need? What perceptual and reasoning mechanisms are required? What is the information processing architecture within which such processes function? What meta-cognitive mechanisms are needed? What ontology is required, and how are ontologies extended? Are all the components, and the architecture containing them, present from birth (which seems unlikely) or do they grow? If so, what mechanisms (forms of representation, architectures, etc.) are required to produce that growth and what are the roles of the genome and the environment in making such growth possible? The work of Karmiloff-Smith on representational redescription, e.g. [16], seems to be related, but as far as I can see it does not propose mechanisms.

2.2 Toddler Theorems

The preceding examples are a small subset from a potentially much larger collection of examples, based on informal observations of infants, toddlers and older children, as well as observations reported in developmental psychology and animal behaviour literature. One way to make progress in this field is to go on collecting, analysing, and organising, examples of what could be called “toddler theorems”, supported, where possible, by videos. Analysing the behaviours, from the viewpoint of a designer of robots that need to function and learn in 3-D environments of the sorts that humans and other animals can cope with, would include partially ordering the examples in terms of the information processing mechanisms that are required. That could be tested by building and demonstrating such mechanisms producing similar achievements, and analysing dependency relationships.

I suspect there are very many examples of the transition, based on “productive laziness”, from empirical discovery to knowledge about necessity of structural relations (in spatial configurations, in processes, etc.). Many of the examples will be about proto-affordances (possibilities for and constraints on change), vicarious affordances (possibilities and constraints for *other* agents) and epistemic affordances (how changes in the environment alter the information available to particular perceivers). Building a shared repository of examples may require overcoming the misguided prejudice among many psychologists that if your observations cannot be fed into a statistics package you are not doing science.

A large and varied collection of examples could be the basis of systematic investigation into requirements for robots able to go through similar (partially ordered) transitions – including requirements that have not been noticed by most robot designers (for instance researchers who over-emphasise the importance of embodiment, morphology, and dynamic coupling with the environment, criticised in [39]). Since animals with different morphologies seem to learn complex abilities to manipulate 3-D objects, and some seem to be able to solve non-trivial planning

⁴ An anonymous referee helpfully noted a flaw in an earlier formulation.

problems (as illustrated in [4]) it may be that what is *common* to animals with *different* morphologies, namely information-processing capabilities evolved in a 3-D environment, is more important for their intelligence (productive laziness!) than their morphology. Compare humans born limb-less, or blind, etc.

Much of human (and animal) intelligence is concerned not with real-time dynamic interaction with the environment (which is obviously important some of the time) but with *avoiding* unnecessary, time-wasting, dangerous, or inefficient interaction. It is no accident that much mathematical thinking about physical events and processes is done without actually producing those events and processes. As Kant noticed, development of those competences is triggered by experience, e.g. producing and observing physical processes in play and exploration, but what results is not a conclusion *derived* from or based on those observations but a new theoretical understanding, providing *explanations* of those observations. (Compare Popper's view of science, in [23].)

3 The Core Conjecture

A core conjecture, partially expounded in [5] is that the individual information processing architecture builds itself by adding new "layers" of competence, partly as a scientific community does.

3.1 Growing layers of competence

Diverse mathematical competences seem to be based on expanding layers of meta-cognitive competences that grow out of more common biological competences that originally evolved to support interactions with a complex, varied and changing 3-D world. These competences include: (a) the ability to detect the need to extend the current ontology substantively (as happens in deep science, e.g. adding "gene", and "valence" without defining them in terms of previous concepts [34]), (b) the ability to create new explanatory and predictive theories and to extend or modify those theories (e.g. by modifying the ontology, basic postulates and reasoning methods they use), (c) the ability to deploy theories, including noticing new opportunities for deployment, (d) the ability to detect and debug flaws in those theories, and (e) the ability to create new forms of representation, including new ways of manipulating representations that are usable in reasoning, explaining, predicting, and inventing.

So far, robots are nowhere near matching these competences. We don't yet understand the requirements well enough to specify the resources needed to support them: e.g. the forms of representation, information-processing mechanisms, and self-modifying multi-functional architectures. There has been much AI work on attempting to model aspects of mathematical competence, e.g. by Seely-Brown, Lenat, Bundy, Jamnik, Winterstein, Colton, Pease, and very recently [12], but that work starts with systems that manipulate abstract symbolic structures (e.g. logical formulae) rather than objects in the environment. Instead, we need to investigate how empirical knowledge is acquired through interactions with the environment, and later transformed to a non-empirical status, as suggested by Kant.

A good way to test Kant's view, is to try to design and build a robot that develops mathematical understanding as humans seem to do, initially by building on their understanding of structures and processes in 2-D and 3-D space using more general biological competences, and also reflecting on their own thinking about structures and processes, and then later reflecting on and talking about the results of their earlier reflection. Formal work in AI on theorem proving does not address these issues, because it starts at a much later stage of sophistication. Contrast [36].

3.2 Psychological Theories About Number Concepts

It is often supposed that the visual or auditory ability to distinguish groups with different numbers of elements (subitizing) displays an understanding of number. However this is simply a perceptual capability. A deeper understanding of numbers requires a much wider range of abilities. Rips *et al.* [24] rightly criticise theories that treat number concepts as abstracted from perception of groups of objects, and discuss alternative requirements for a child to have a concept of number, concluding that having a concept of number involves grasping (not necessarily consciously) a logical schema something like Peano's five axioms. They claim that that is what enables a child to work out various properties of numbers, e.g. the commutativity of addition, and the existence of indefinitely larger numbers. This requires logical capabilities to draw conclusions from the axioms, possibly unconsciously. How can children acquire such competences? They conclude that somehow the Peano schema and the logical competences are innately *built into* the child's "background architecture" (but do not specify how that could work).

Instead, I conjecture such competences are "meta-configured", i.e. not wholly determined in the genome, but produced through interactions with the environment that generate layers of meta-competences, competences that enable new competences to be acquired ([5, 40]. Some hints about how that might occur are presented below.

Many psychologists and researchers in animal cognition misguidedly search for experimental tests for whether a child or animal does or does not understand what numbers are.⁵ Rips *et al.* are not so committed to specifying an experimental test, but they do require a definition that makes a clear distinction between understanding and not understanding what numbers are.

3.3 An alternative approach.

As far as I know, no developmental psychologists have considered the alternative view, presented over 30 years ago in [28], chapter 8, that there is no single distinction between having and not having a concept of number, because learning about numbers involves a never-ending process that starts from relatively primitive and general competences that are not specifically mathematical and gradually adds more and more sophistication, in parallel with the development of other competences. Independently developed ideas in [44] also suggest that learning about numbers involves developing capabilities of the following sorts:

1. memorising an ordered sequence of arbitrary names;
2. performing a repetitive action;
3. performing two repetitive actions together, in synchrony;
4. initiating such processes, with different stopping conditions, depending on the task;
5. doing so when one process is uttering a learnt sequence of names;
6. learning rules for extending the sequence of names indefinitely;
7. observing various patterns in such processes and storing information about them, e.g. information about successors and predecessors of numerals, or results of counting onwards various amounts from particular numerals ([28], Chap 8);
8. noticing commonalities between static mappings and process mappings (e.g. paired objects vs paired events);
9. finding mappings between components of static structures as well as the temporal mappings between process-elements;

⁵ Compare the mistake of striving for a definitive test for whether animals of some species understand causation, criticised here in presentation 3: <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac>

10. noticing that such mappings have features that are independent of their order (e.g. counting a set of objects in two different orders must give the same result);
11. noticing that numbers themselves can be counted, e.g. the numbers between two specified numbers;
12. noticing possibilities of and constraints on rearrangements of groups of objects – e.g. some can be arranged as rectangular arrays, but not all;
13. learning to compare continuous quantities by dividing them into small components of a standard size and counting.
14. grasping (or inventing) the notion of a *limiting case*.

Such competences and knowledge can be extended indefinitely. Some can be internalised, e.g. counting silently. Documenting all the things that can be discovered about such structures and processes in the first few years of life could fill many pages. (Compare [26, 18, 36, 32].) The sub-abilities are useful in achieving practical goals by manipulating objects in the environment and learning good ways to plan and control such achievements, for example: fetching enough cups to give one each to a group of people, or matching heights of two columns made of bricks, to support a horizontal beam, or ensuring that enough water is in a big jug to fill all the glasses on the table.

Gifted teachers understand that any deep mathematical domain has to be explored from multiple directions, gaining structural insights and developing a variety of perceptual and thinking skills of ever increasing power: learning new constructs, new reasoning procedures, learning to detect deficiencies in understanding, and finding out how to remedy them. Our work and Wiese's [44] provide conjectures about some of the information-processing mechanisms involved. As far as I know, nobody has tried giving a robot such capabilities. Simplified versions should be attempted. It seems that no other known species has an information-processing architecture able to support all this – and neither does a newborn human infant. We need to understand not “core concepts” [41] so much as *core architecture-building mechanisms*.

3.4 Internal Construction Competences

The processes described above require the ability to create (a) new internal information structures, including, for example, structures recording predecessors of numbers, so that it is not necessary always to count up to N to find the predecessor of N , and (b) new algorithms for operating on those structures. As these internal information-structures grow, and algorithms for manipulating them are developed, there are increasingly many opportunities to discover more properties of numbers. The more you know, the more you can learn.

Moreover those constructions do not happen instantaneously or in an error-free process. Many steps are required including much self-debugging, as illustrated by Sussman's HACKER [42]. This depends on self-observation during performance of various tasks, including observations of both external actions and thinking. One form of debugging is what Sussman called detecting the need to create new “critics” that run in parallel with other activities and interrupt if some pattern is matched, for instance if disguised division by zero occurs.

The ongoing discovery of new invariant patterns in structures and processes produced when counting, arranging, sorting, or aligning sets of objects, leads to successive extensions of the learner's understanding of numbers. Initially this is just empirical exploration, but later a child may realise that the result of counting a fixed set of objects cannot depend on the order of counting. (Why not?) That invariance (a kind of transitivity) is intrinsic to the nature of one-to-one mappings and does not depend on properties of the things

being counted, or on how fast or how loud one counts, etc. However, some learners may never notice this non-empirical character of mathematical discoveries until they take a philosophy class!

3.5 Approaches to infinity

One of the non-empirical discoveries is that the sequence of natural numbers is never ending. Kant suggested that this requires grasping that a rule can go on being applied indefinitely, which fits the outline theory summarised above in 3.3. A child able to produce repetitive processes and simultaneously monitor and reflect on what is happening, could learn that there are repetitive processes of two kinds: those that start off with a determinate stopping condition that limits the number of repetitions and those that do not, though they can be stopped by an external process. Tapping a surface, walking, making the same noise repeatedly, swaying from side to side, repeatedly lifting an object and dropping it, are all examples of the latter type. (The idea of approaching a limiting value might grow out of such of meta-cognition.) This contrasts with the suggestion by Rips *et al.* [24] that a child somehow acquires logical axioms which state that every natural number has exactly one successor and at most one predecessor, and that the first number has no predecessor, from which it follows logically that there is no final number and the sequence of numbers never loops.

The general notion of something not occurring is clearly required for intelligent action in an environment. E.g. failure of an action to achieve its goal needs to be detectable. So if the learner has the concept of a repetitive process leading to an event that terminates the process, then the general notion of something not happening can be applied to that to generate the notion of something going on indefinitely. From there, depending on the information processing architecture and the forms of representation available, it may be a small step to the representation of two synchronised processes going on indefinitely, one of which is a counting process.

What is more sophisticated is acquiring a notion of a sequence of sounds or marks that can be generated indefinitely without ever repeating a previous mark. An obvious way to do that is to think of marks made up of one or more dots or strokes. Then the sequence could start with a single stroke, followed by two strokes, followed by three strokes, etc., e.g. | || ||| |||| ||||| etc. Such patterns grow large very quickly. That can motivate far more compact notations, like arabic numerals, though any infinitely generative notation will ultimately become physically unmanageable.

3.6 Backward to zero, and beyond

A different sort of extension is involved in adding *zero* to the natural numbers, which introduces “anomalies”, such as that there is no difference between *adding* zero apples and *subtracting* zero apples from a set of apples. A special meaning has to be invented for “multiplying by zero” – or it could be deemed meaningless.

Negative integers add further confusions. This extension is rarely taught properly, and as a result most people cannot give a coherent explanation of why multiplying two negative numbers should give a positive number. It cannot be *proved* on the basis of previous knowledge because what multiplying by a negative number means is undefined initially. For a mathematician, the *simplest* way to extend multiplication rules to include negative numbers without disruption of previous generalisations, is to stipulate that multiplying two negatives produces a positive. (Compare defining what 3^{-1} and 3^0 should mean.)

Some teachers use demonstrations based on the so-called “number line” to introduce notions of negative integers, but this can lead to serious muddles (e.g. about multiplication). Pamela Liebeck [19]

developed a game called “scores and forfeits” where players have two sets of tokens: addition of a red token is treated as equivalent to removal of a black token, and vice versa. Playing and discussing the game seemed to give children a deeper understanding of negative numbers and subtraction than standard ways of teaching, presumably because the set of pairs of natural numbers can be used to model accurately the set of positive and negative integers. Productive laziness can stimulate a search for good notations and procedures for manipulating them: a difficult task with high payoff.

3.7 Beyond discrete sets

Cardinality and orderings are properties of *discrete* sets. Extending the notion of number to include *measures* that are continuously variable, e.g. lengths, areas, volumes and time intervals, requires sophisticated extensions to the learner’s ontology and forms of representation – leading to deep mathematical and philosophical problems (e.g. Zeno’s paradoxes). I suspect that the development of such notions in children requires more intermediate stages than most investigators have noticed, including stages where there are only *partial orderings* of “amount” or “quantity” rather than global measures. For example, if P1 and P2 are polygons, and P2 can fit inside P1, then P2 is clearly smaller than P1. If P1 can fit inside P2, then clearly P2 is larger than P1. But if P1 is a circle and P2 is a square, and the diagonal of the square is slightly more than the diameter of the circle, then it can turn out that neither can fit inside the other. So deciding which is bigger (or has more area) requires invention of a new mode of comparison. Thinking of doing that by dividing both into many small squares of the same size and counting is both a major achievement and a source of problems because areas with curved boundaries, or with sides that have irrational ratios cannot be exactly tiled using squares. I expect few primary school teachers understand the intellectual challenges facing the child. Psychologists who investigate whether children have or have not grasped some notion of area, or volume, or conservation of amount, do not always probe for the intermediate stages. There is no requirement for all learning trajectories to be the same, when the terrain is so rich and complex: making rigidly sequenced curricula deeply flawed.

4 Problems Faced by Biological Evolution

4.1 Affordances, Visual Servoing, and Beyond

Analysis of biological requirements for vision (including human vision) enlarges our view of the functions of vision, requiring goals of AI vision researchers to be substantially expanded. An example is the role of vision in servo-control, including control of continuous motion as well as discrete condition-checking.

Gibson’s work on affordances in his [9] showed that animal vision provides information not merely about geometrical and physical features of entities in the environment, as in [2, 21], nor merely about recognising or categorising objects (the focus of much recent AI ‘vision’ research), but about what the perceiver can and cannot do, given its physical capabilities and its goals. I.e. visual information includes what processes *can and cannot occur in the environment* [30]. This requires an “exosomatic” ontology that is more concerned with what exists in the environment than with contents of retinal arrays. Gibson did not go far enough, as we have seen in discussing proto-affordances, vicarious affordances, etc., in 2.1 and 2.2.

4.2 Perception of Actual and Possible Processes

An EU FP6 cognitive robotics project CoSy,⁶ included analysis of

⁶ Described in <http://www.cs.bham.ac.uk/research/projects/cosy/>

requirements for a robot capable of perceiving and manipulating 3-D objects, including: (a) representing 3-D objects with parts and relationships; (b) multiple ontological layers in scene structures (as in chapter 9 of [28]); (c) representing “multi-strand relationships”, since both objects and parts of objects are related within and across object boundaries (e.g. faces, edges and vertices of a block); (d) representing “multi-strand processes” when multi-strand relationships change, e.g. with metrical, topological, causal, functional, continuous, and discrete changes occurring concurrently; (e) representing new *possible* multi-strand processes and constraints on such possibilities – the positive and negative “proto-affordances” that underlie affordances; (f) predicting outcomes of possible processes that are not currently occurring (e.g. when planning); (g) constructing possible explanations of how a perceived situation came about. Some representations of processes may use a partial simulation of the processes [6, 11]. Those biological mechanisms, not yet matched in AI, appear to be the basis of mechanisms capable of use in geometrical/topological reasoning.

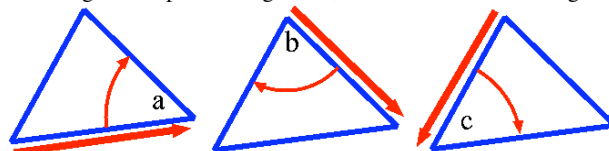
As explained in [38, 31], that requires an *amodal, exosomatic* form of representation of processes; one that is not tied to the agent’s sensorimotor processes. That possibility is ignored by researchers who focus only on sensorimotor learning and representation, and assume all semantic content requires “symbol-grounding” [34].

4.3 The Importance of Kinds of Matter

A child, or robot, learning about kinds of process that can occur in the environment, needs to be able to extend the ontology she uses indefinitely, and not merely by defining new concepts in terms of old ones: there are also *substantive* ontology extensions (as in the history of physics and other sciences). For example, concepts of different kinds of matter are required to explain different sorts of deviations from rigidity: string and wire are flexible, but wire retains its shape after being deformed; an elastic band that returns to its original length after being stretched, need not restore its shape after bending. Some kinds of stuff easily separate into chunks in various ways, if pulled, e.g. mud, porridge, plasticine and paper. A subset of those allow restoration to a single object if separated parts are pressed together. There are also objects that are marks on other objects, like lines on paper, and there are some objects that can be used to produce such marks, like pencils and crayons. Marks produced in different ways and on different materials can have similar structures. Compare [26]. However, it is important to remember ([27]) that 2-D pictures/diagrams of 3-D structures/processes *cannot* be isomorphic with what they represent, as commonly assumed.

4.4 Perception and Mathematical Discovery

Many mathematical discoveries involve noticing an invariant in a class of processes. Mary Ensor, a mathematics teacher, once told me she had found a good way to teach children that the internal angles of a triangle add up to a straight line, as demonstrated in the figure.



An arrow starting at one corner, pointing along one side, can be rotated three times as indicated, going through positions shown in the three successive figures. The successive rotations *a*, *b* and *c* go through the interior angles of the triangle, and reverse the arrow, so they must add up to a straight line. This discovery may initially be made through empirical exploration with physical objects, but the

pattern involved does not depend on what the objects are made of, the colours used, lengths of lines, particular angles in the triangle, temperature, strength magnetic field, atmospheric pressure, etc. How such invariants are discovered, represented, and used, and which animals are capable of this is still unknown. But that capability lies at the heart of Kant's claims about the nature of mathematical knowledge.

4.6 Non-empirical Does Not Imply Infallible or Innate.

The processes are not infallible: mathematical discoveries can have "bugs" as Lakatos [17] demonstrated using the history of Euler's theorem about polyhedra. Further examples are in [36]. This does not imply that mathematical knowledge is empirical in the same way as knowledge about the physical properties of matter. Lakatos used the label "quasi-empirical", but it needs to be explicated in terms of mechanisms required. Even logic is quasi empirical, insofar as logicians can produce buggy proofs and generate paradoxes requiring "debugging". The kind of self-debugging demonstrated in [42] is very different from finding out more about what is in the physical environment. The discovery of bugs in proofs and good ways to deal with them is an important feature of mathematical learning as well as learning to program. Ensor's proof (in 4.4) would break down on a spherical surface, for example, since the triangle could have three 90 degree angles (among other possibilities inconsistent with the original result). Noticing this might lead a learner to investigate properties of triangles that distinguish planar and non-planar surfaces. But that exploration does not *require* experiments in a physical laboratory, though it may benefit from them. Kant claimed that such discoveries are about the perceiver's forms of perception, but they are not restricted to any particular perceivers. Making Kant's claim more precise, in the form of a working design for a robot learner remains a challenge.

4.7 Humean and Kantian Causation

The forms of representation required for counting and matching groups of entities or parts of processes, are different from those required for understanding topology and Euclidean geometry. The latter seem to build on reasoning/planning competences based on visual competences, illustrated in [26]. Some of these competences are apparently shared with some other animals – those that are capable of planning and executing novel spatial actions [4]. Adding properties of matter, such as rigidity and impenetrability, to representations of shape and topology allows additional reasoning about and prediction of results of processes. An example is the ability to use the fact that two meshed gear wheels are rigid and impenetrable to *work out*, in advance of experimenting, how rotating one will cause the other to rotate. That kind of reasoning about processes in the environment is not always available.

If the wheels are not meshed, but there are inaccessible, hidden mechanisms, then the only basis for predicting the consequence of rotating the wheels is to use a Humean (statistical, Bayesian) notion of causation: basing predictions of results of actions or events solely on observed correlations. In contrast, where the relevant structure and constraints are visible, or already known, then mathematical reasoning (using geometry and topology) may be used. I believe this underlies Kant's anti-Humean conception of structure-based, deterministic, causation.⁷ A significant subset of the causal understanding of the environment that a child acquires (though not all) is Kantian because it allows the consequences of novel processes to be *worked out* on the basis of geometric and

topological relationships, and kinds of matter involved. But for this, much engineering design would be impossible.

4.5 Doing Philosophy by Doing AI

I suspect that many philosophical disputes about the nature of mathematics will be transformed if we can build not a standard theorem proving machine but a robot that works out how to achieve "productive laziness" by doing spatial/geometrical reasoning. Many researchers (e.g. those in [10]) have pointed out the need to provide intelligent machines with spatial forms of representation and reasoning, but progress in replicating animal visual/spatial abilities discussed above has been very slow. In part, this is because the requirements for human-like (or animal-like) visual systems have not been analysed in sufficient depth (as illustrated in [31, 38]). There is a vast amount of research on "object recognition" that contributes nothing to our understanding of how 3-D spatial structures and processes are seen or how information about spatial structures and processes is used, for instance in reasoning and acting. Differences between perception of 3-D structures and proto-affordances, and mere recognition are discussed in [35]. The deeper competences appear to be required to support Kant's philosophy of mathematics.

5 Individual and social processes

5.1 Is Human Mathematical Competence Unique?

I have tried to present a view of products of biological evolution used for interacting with a complex spatio-temporal environment whose structure allows some things to be *worked out* by an individual instead of having to be learnt empirically, even though the working out is sometimes preceded by empirical discovery of the result, followed by a re-interpretation at a later stage. Further work is required to understand both the evolutionary pressures and what exactly those pressures have produced in humans and other animals (e.g. corvids, elephants, dolphins, octopuses, dogs, and others).

The evidence considered here suggests that the *long term potential* to learn about even fairly advanced mathematics is in most newborn human infants (though there may also be genetic individual differences which can only be understood in the light of future, more detailed theories of the mechanisms involved). Moreover, the existence of the potential does not mean that that potential can be realised in a normal lifetime without additional help. One product of evolution that is not unique to humans, but is supported by especially powerful mechanisms in humans, is the ability of individuals to speed up their learning through guidance and scaffolding based on learning done by predecessors. This makes cultural evolution, and cultural learning possible. It is facilitated by toys, games, development of tools, shelters, clothing, collaborative experiences with acquiring and consuming food, etc. (The explanatory power of imitation can be over-rated if the mechanisms required for imitation are ignored.)

Many animals, e.g. dogs, parrots, various primates, elephants, and others, are able to acquire unusual cognitive competences as a result of domestication and training by humans, but are not able to pass on what they have learnt to others, although there are other things that they teach their young. This means that humans in groups can share the process of acquiring new knowledge and competences and passing them on with much faster progress between generations than other species. A well-known spectacular example was the spontaneous development of a novel sign language by a group of deaf children in Nicaragua, but human history is full of examples of collaborative advance, usually spread out over longer time periods and with a large age spread in the communication from experts to learners. That applies also to mathematical advances. So although the

⁷ See also <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac>

products of biological evolution discussed in previous sections are necessary for development of mathematics, they do not suffice for the development of advanced and systematised mathematical knowledge in a human lifetime: that requires cultural development. It also seems to be needed for the growth of meta-mathematical and philosophical reflection on what has been learnt by doing mathematics. Perhaps future robots, or a future animal species, will not need that cultural support for rapid mathematical or philosophical development – if they can internalise the feedback loops.

6 Conclusion

I have sketched a research programme motivated (for me, though not necessarily for others) by the aim of explaining the nature of mathematical discovery, and supporting major features of Kant's philosophy of mathematics. We need further investigation of architectures and forms of representation that allow playful exploration by a robot, leading to development of "framework theories" that can support the discovery of patterns and invariants in structures and processes that are needed for development of deep mathematical concepts and mathematical forms of reasoning. The robot should be able to go through the following stages:

1. Acquiring familiarity with some domain, e.g. through playful exploration;
2. Noticing (empirically) certain generalisations;
3. Discovering a way of thinking about them that shows they are not empirical;
4. Generalising, diversifying, debugging, deploying, that knowledge;
5. Formalising the knowledge, possibly in more than one way.

This will be based on mechanisms for building powerful explanatory theories about the environment that seem to have evolved in other species, and are used (unwittingly) by infants and toddlers before they start discovering what I have called "toddler theorems".

Although this work emphasises embodiment, because spatial embodiment presents both the problems that need to be solved (including problems of manipulating, route-finding, building things, eating food, avoiding being eaten, finding shelter, etc.), it does not emphasise embodiment in the way that much current research in neuroscience and AI/Robotics does e.g. [20, 3], namely by focusing on the detailed morphology of a species or the *real-time dynamic interaction with the environment*. That is certainly a problem for all organisms. What is special about the organisms we regard as most intelligent (including humans, primates, some birds, octopuses) seems to be their relatively unusual ability to use an understanding of the structures and potential in the environment for purposes of deliberation rather than action, e.g. selecting goals, and creating plans, as well as carrying out plans. They can control their behaviour in a productively lazy way: not driven simply by the physics of the environment and their own morphology (like a ball going down a helter-skelter), but in part by using information acquired by perception of structures to reason (mathematically) about structures and future possible and impossible processes [39].

I have so far described only a subset of the requirements for working designs. Some more detailed requirements are in cited papers. But I don't claim that we understand all the requirements yet. It is clear that AI still has a long way to go before the visual and cognitive mechanisms of robots can match the development of a typical human child going through infant and toddler learning, but it is not clear exactly what the gaps in our understanding are.

A possible (fanciful?) outcome of this symposium might be creation of an informal multidisciplinary network for collaboration

on refining the requirements, and collecting more evidence regarding the fine structure of individual development of the kind discussed here. This could feed into a collaborative project to produce a working prototype system as a proof of concept, using a simulated robot, perhaps one that manipulates 2-D shapes in a plane surface, discovering properties of various kinds of interactions, involving objects with different shapes made of substances with various properties that determine the consequences of the interactions, e.g. impenetrability, rigidity, elasticity, etc. A later stage would add reflective capabilities enabling the robot to discover the differences between what it can work out and what it merely discovers about its world empirically. Much later it might become a philosopher of mathematics able, unlike us, to reflect on known features of its own design to explain its capabilities, fulfilling the dream (nightmare for some) expressed in my 1978 book [28] of basing Kant's epistemology and philosophy of mathematics on a working model.

Acknowledgements

I have benefited from discussions with too many people to list here, including colleagues in philosophy, psychology, biology, AI and computer science. Referees for the symposium kindly made several useful comments. Some of these ideas were developed while working on two EU-funded cognitive robotics projects, CoSy and its successor CogX.

REFERENCES

- [1] Jody Azzouni, 'The compulsion to believe: Logical inference and normativity', *ProtoSociology*, **25**, 69–88, (2008).
- [2] H.G. Barrow and J.M. Tenenbaum, 'Recovering intrinsic scene characteristics from images', in *Computer Vision Systems*, eds., A. Hanson and E. Riseman, 3–26, Academic Press, New York, (1978).
- [3] Alain Berthoz, *The Brain's sense of movement*, Perspectives in Cognitive Science, Harvard University Press, London, UK, 2000.
- [4] R W Byrne and A Russon, 'Learning by imitation: a hierarchical approach', *Behavioral and Brain Sciences*, **21**, 667–721, (1998).
- [5] Jackie Chappell and Aaron Sloman, 'Natural and artificial meta-configured altricial information-processing systems', *International Journal of Unconventional Computing*, **3**(3), 211–239, (2007). <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609>.
- [6] K. Craik, *The Nature of Explanation*, Cambridge University Press, London, New York, 1943.
- [7] R.P. Feynman, *The Character of Physical Law*, MIT Press, Cambridge MA, 1965. The 1964 Messenger Lectures.
- [8] Eleanor J. Gibson and Anne D. Pick, *An Ecological Approach to Perceptual Learning and Development*, Oxford University Press, New York, 2000.
- [9] J. J. Gibson, *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, MA, 1979.
- [10] *Diagrammatic Reasoning: Computational and Cognitive Perspectives*, eds., J. Glasgow, H. Narayanan, and B. Chandrasekaran, MIT Press, Cambridge, MA, 1995.
- [11] R. Grush, 'The emulation theory of representation: Motor control, imagery, and perception', *Behavioral and Brain Sciences*, **27**, 377–442, (2004).
- [12] Markus Guhe, Alison Pease, and Alan Smaill, 'A Cognitive Model of Discovering Commutativity', in *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, eds., N. A. Taatgen and H. van Rijn, pp. 727–732, Austin, TX, (2009). Cognitive Science Society.
- [13] Eva Jablonka and Marion J. Lamb, *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*, MIT Press, Cambridge MA, 2005.
- [14] M. Jamnik, A. Bundy, and I. Green, 'On automating diagrammatic proofs of arithmetic arguments', *Journal of Logic, Language and Information*, **8**(3), 297–321, (1999).
- [15] I. Kant, *Critique of Pure Reason*, Macmillan, London, 1781. Translated (1929) by Norman Kemp Smith.

- [16] A. Karmiloff-Smith, 'Precis of: Beyond Modularity: a Developmental Perspective on Cognitive Science', *Behavioral and Brain Sciences*, **17**(4), 693–706, (1994). <http://www.bbsonline.org/documents/a/00/00/05/33/index.html>.
- [17] I. Lakatos, *Proofs and Refutations*, Cambridge University Press, Cambridge, UK, 1976.
- [18] Pamela Liebeck, *How Children Learn Mathematics: A Guide for Parents and Teachers*, Penguin Books, Harmondsworth, 1984.
- [19] Pamela Liebeck, 'Scores and Forfeits: An Intuitive Model for Integer Arithmetic', *Educational Studies in Mathematics*, **21**(3), 221–239, (Jun 1990).
- [20] M. Lungarella and O. Sporns, 'Mapping information flow in sensorimotor networks', *PLoS Computational Biology*, **2**(10:e144), (2006). DOI: 10.1371/journal.pcbi.0020144.
- [21] D. Marr, *Vision*, W.H. Freeman, San Francisco, 1982.
- [22] J. S. Mill, *A System of Logic, Ratiocinative and Inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*, John W. Parker, London, 1843.
- [23] K.R. Popper, *Objective Knowledge*, Oxford University Press, Oxford, 1972.
- [24] L. J. Rips, A. Bloomfield, and J. Asmuth, 'From Numerical Concepts to Concepts of Number', *The Behavioral and Brain Sciences*, **31**(6), 623–642, (2008). doi:10.1017/S0140525X08005566.
- [25] B. Russell, *Mysticism and Logic and Other Essays*, Allen & Unwin, London, 1917.
- [26] J. Sauvy and S. Suavy, *The Child's Discovery of Space: From hopscotch to mazes – an introduction to intuitive topology*, Penguin Education, Harmondsworth, 1974. Translated from the French by Pam Wells.
- [27] A. Sloman, 'Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence', in *Proc 2nd IJCAI*, pp. 209–226, London, (1971). William Kaufmann. <http://www.cs.bham.ac.uk/research/cogaff/04.html#200407>.
- [28] A. Sloman, *The Computer Revolution in Philosophy*, Harvester Press (and Humanities Press), Hassocks, Sussex, 1978. <http://www.cs.bham.ac.uk/research/cogaff/crp>.
- [29] A. Sloman, 'The primacy of non-communicative language', in *The analysis of Meaning: Informatics 5 Proceedings ASLIB/BCS Conference, Oxford, March 1979*, eds., M. MacCafferty and K. Gray, pp. 1–15, London, (1979). Aslib. <http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#43>.
- [30] A. Sloman, 'Actual possibilities', in *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifth International Conference (KR '96)*, eds., L.C. Aiello and S.C. Shapiro, pp. 627–638, Boston, MA, (1996). Morgan Kaufmann Publishers. <http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15>.
- [31] A. Sloman, 'Putting the Pieces Together Again', in *Cambridge Handbook on Computational Psychology*, ed., Ron Sun, chapter 26, 684–709, Cambridge University Press, New York, (2008). <http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#710>.
- [32] A. Sloman. Ontologies for baby animals and robots. From "baby stuff" to the world of adult science: Developmental AI from a Kantian viewpoint., 2009. Online tutorial presentation: <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#brown>.
- [33] A. Sloman and J. Chappell, 'The Altricial-Precocial Spectrum for Robots', in *Proceedings IJCAI'05*, pp. 1187–1192, Edinburgh, (2005). IJCAI. <http://www.cs.bham.ac.uk/research/cogaff/05.html#200502>.
- [34] Aaron Sloman. Why symbol-grounding is both impossible and unnecessary, and why theory-tethering is more powerful anyway., 2007. <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models>.
- [35] Aaron Sloman. A Multi-picture Challenge for Theories of Vision, 2008. <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/multipic-challenge.pdf>.
- [36] Aaron Sloman. A New Approach to Philosophy of Mathematics: Design a young explorer, able to discover "toddler theorems", 2008. Online presentation <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#toddler>.
- [37] Aaron Sloman. Evolution of minds and languages. What evolved first and develops first in children: Languages for communicating, or languages for thinking (Generalised Languages: GLs)?, 2008. <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0702>.
- [38] Aaron Sloman, 'Architectural and representational requirements for seeing processes and affordances', in *Computational Modelling in Behavioural Neuroscience: Closing the gap between neurophysiology and behaviour.*, eds., D. Heinke and E. Mavritsaki, 303–331, Psychology Press, London, (2009). <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0801>.
- [39] Aaron Sloman, 'Some Requirements for Human-like Robots: Why the recent over-emphasis on embodiment has held up progress', in *Creating Brain-like Intelligence*, eds., B. Sendhoff, E. Koerner, O. Sporns, H. Ritter, and K. Doya, 248–277, Springer-Verlag, Berlin, (2009). <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0804>.
- [40] Aaron Sloman and Jackie Chappell, 'Computational Cognitive Epigenetics (Commentary on [13])', *Behavioral and Brain Sciences*, **30**(4), 375–6, (2007). <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0703>.
- [41] Elizabeth Spelke, 'Core Knowledge', *American Psychologist*, 1233–1243, (November 2000).
- [42] G.J. Sussman, *A computational model of skill acquisition*, American Elsevier, San Francisco, CA, 1975. <http://dspace.mit.edu/handle/1721.1/6894>.
- [43] A. N. Whitehead and B. Russell, *Principia Mathematica Vols I – III*, Cambridge University Press, Cambridge, 1910–1913.
- [44] Heike Wiese, 'The Co-Evolution of Number Concepts and Counting Words', *Lingua*, **117**, 758–772, (2007).
- [45] D. Winterstein, *Using Diagrammatic Reasoning for Theorem Proving in a Continuous Domain*, Ph.D. dissertation, University of Edinburgh, School of Informatics., 2005. <http://www.era.lib.ed.ac.uk/handle/1842/642>.