In memory of Maria Panteki

How Boole broke through the top syntactic level

Wilfrid Hodges Herons Brook, Sticklepath, Okehampton January 2010 wilfrid.hodges@btinternet.com

1 Maria Panteki as I remember her

Maria Panteki came to Bedford College, University of London in around 1980 to take an MSc in Mathematics. I was on the Mathematics staff at Bedford from 1968 to 1984. In that year the college was closed down, and the assets and records of its Mathematics department were scattered around London University. Last year I was involved in an unsuccessful attempt to track some of them down. So I think it would be hopeless to try to dig out the official records on Maria, and I have to rely on memory.

She was a lively member of my Universal Algebra class. I'm told she had attended my Logic class before that, but it was a large class and I confess I don't remember. She was a close friend of my PhD student Cornelia Kalfa, and the two later became colleagues on the staff of the Aristotle University of Thessaloniki. To me as a logician it has been a particular point of pride that two logicians at the Aristotle University were students of mine.

She moved on from Bedford College to work with Ivor Grattan-Guinness on a group of mid nineteenth-century British mathematicians, some but not all of whom were also logicians. Her work in this field has become well known and justly praised for its scholarship and its penetration. She was an eager correspondent, and over the years she kept in touch with me at the logical end of this work. The flow of information was almost entirely from her to me. She explained to me the environment in which William Hamilton (of Edinburgh), Augustus De Morgan and George Boole worked. I learned about George Peacock, Duncan Gregory and Thomas Solly from her. Occasionally I could fill a small gap in her knowledge — I recall that she sent me an encyclopedia entry by one HWBJ whom she couldn't identify, and I gave her H. W. B. Joseph 1867-1943 (remembering that when I was a boy I was introduced to an old lady who I was told was Joseph's younger sister).

Her untimely death is a personal loss to many of us, and a real sadness for the history of mathematics. She would certainly have been delighted at the thoughtful conference on 'History of Modern Algebra: 19th century and later' which her colleagues at Thessaloniki dedicated to her memory. I add my own thanks to them for their warm hospitality.

2 History and mathematics

Maria and I didn't always agree in our assessments. In one of our discussions in 1999, she seemed puzzled by the link I drew between Peacock and Boole — more on this later. She wrote to me:

(1)

Since you mention Boole, I found not a single reference of his to Peacock, and I was greatly surprised. There was definitely a line of influence from P's symbolic algebra to B's algebraic logic, but as noted in my paper this line concerned mainly the elaboration of P's ideas by D. F. Gregory. ... Of course you have a specific prism to see their writings, that of model theory, a modern approach, whereas my own tends to be deeply historical, checking rather the background of these notions than their fruit.

(In passing I note the graceful syntax of Maria's last sentence above, which is more 19th than 20th century English. Clearly she absorbed more than mathematics from the sources that she studied!)

Maria is absolutely right to point to a difference between her approach as a historian of mathematics and mine as a mathematician interested in history. But I would phrase it a little differently. The difference between 'background' and 'fruit' — to use her words — seems to me the difference between tracing influences backwards in time and tracing them forwards. Both are difficult tasks that mathematicians like me should leave to the historians; my expertise in model theory gives me no specialist tools for either of these tasks.

But for me there is an important third task. The nineteenth century documents have to be measured not only against their context in history, but also against their context in the mathematical facts. The only reservation I would put on this is that if we read mathematical documents of an earlier age in the light of our own mathematics, we can easily misidentify the mathematical facts that the earlier documents are discussing. And here is the task: to identify what piece of mathematics a particular historical mathematician is discussing. For this you have to be a mathematician — otherwise you can hardly do more than describe the words and symbols, and this is not at all the same as locating the mathematical content. And of course you have to be a historian too — otherwise you can only describe how far the historical figure succeeded in grasping the mathematics that you know yourself.

In a paper for a recent conference on understanding traditional Indian logic, I illustrated this point by reconstructing some unpublished work of Lindenbaum and Tarski from the 1920s, [6]. Below I will document it with another example, this time taken from George Boole.

3 Boole's rule

Huge changes came over logic during the period 1830-1930 (taking rough dates). A question that has often worried me is to describe correctly the main differences between the earlier logic and the later. Popular accounts of the difference are often still based on the propaganda of the winning side in the battle between the old and the new, and this is never a good basis for reaching the truth.

George Boole introduced a certain rule in his *Mathematical Analysis of Logic* of 1847, [1]. The rule is strikingly different from the normal rules of traditional logic, but in modern logic it would hardly raise an eyebrow. So it serves as one criterion of the difference between the old logic and the new. For the remainder of this paper I will try to identify just what the rule was. This involves stating both its mathematical content and the justification that Boole thought he had for using it.

I ignore completely the question of its 'fruit' — I don't know the evidence that anybody else ever read this part of Boole's work, and I confess I haven't pursued the question. But we will need to look at the 'background', because it forms part of the evidence for Boole's intentions. Putting oneself into the mind of someone from a different culture is always hard, and everything I say is provisional. If I had a quarter of Maria's knowledge of the period, I'm sure I would have said some things differently.

Boole doesn't state the rule explicitly, but he calls attention to a particular case of it on page 67 of [1]:

(2) Let us represent the equation of the given Proposition under its most general form,

$$a_1t_1 + a_2t_2 \dots + a_rt_r = 0$$

... Now the most general transformation of this equation is

$$\psi(a_1t_1 + a_2t_2 \dots + a_rt_r) = \psi(0),$$

provided that we attribute to ψ a perfectly arbitrary character, allowing it even to involve new elective symbols, having *any proposed relation* to the original ones. (Boole's italics.)

The 'transformation' that Boole is invoking here is as follows:

(3)

Let $\psi(x)$ be a boolean function of one variable, and let *s* and *t* be boolean terms. Then from s = t we can derive $\psi(s) = \psi(t)$.

I will call this Boole's rule. Where we say 'boolean function' he speaks of 'elective symbols'; this is an important difference but I think it is irrelevant to our discussion. Also 'derive' just means we perform the transformation; without further investigation we can't assume that Boole intends the rule as a rule of derivation in the sense of modern logic, though he clearly intends something along those lines.

I divide my comments on the rule into two parts. The first part is about the rule itself with no particular reference to Boole. The second is about how Boole himself intended it.

4 Boole's rule in itself

There are three things to be said here. (1) The rule is syntactically 'deep'. (2) Nothing like it appears in traditional logic before Boole. (3) All modern systems of predicate logic use either it or some related deep rule.

4.1 The rule is deep

Consider the case where Boole's term $\psi(x)$ has the form fghjk(x), and where f, g etc. are elective symbols (or more generally 1-ary function symbols). Boole thinks of the term as built up by applying f to ghjk(x), which in turn is got by applying g to hjk(x) and so on. (His symbols 'operate upon' what follows them; [1] p. 15ff.) So parsing $\psi(x)$ gives a tree:



Then $\psi(s)$ and $\psi(t)$ have the same parsing, except that at the bottom they have respectively *s* and *t* in place of *x*. (The terms *s* and *t* might themselves be complex, so that the parsing of $\psi(s)$ and $\psi(t)$ could be continued downwards.) So the application of Boole's rule in this case involves making changes at the sixth level from the top. For every natural number *n* we can construct an example where the application of Boole's rule involves unpacking an expression down to *n* levels. This is what is meant by saying that Boole's rule is 'deep'.

Boole himself says in (2) above that ψ in his rule has a 'perfectly arbitrary character' and may involve new elective symbols. But at his date no logicians distinguished systematically between written expressions and what they stand for, so that the notion of parsing had no real purchase. This situation changed only in the 1920s, thanks to work of Post, Leśniewski, Tarski and others.

4.2 Traditional logic has no deep rules

The inference and transformation rules found in traditional aristotelian logic are never deep. Usually they assume that a sentence has one of the four forms

Every A is a B. No A is a B. Some A is a B.

Some *A* is not a *B*.

In any reasonable way of parsing these sentences, A and B will be near the top of the analysis. Some traditional logicians emphasise the fact that the rules of logic don't reach down inside the expressions put for A and B. One does meet some more complicated sentence forms, for example

If p then q.

Necessarily every *A* is a *B*. Every *A*, insofar as it is an *A*, is a *B*.

But none of these require a rule that reaches down to an arbitrarily deep level inside expressions.

There is really only one qualification that we need to make to this broad claim. Namely, traditional logicians accepted that in order to apply the rules of logic to a sentence, we often have to paraphrase the sentence first. So a sentence could in theory mean the same as 'Every A is a B', but have the A buried several levels down inside some contorted phrasing. But in practice we don't meet arbitrarily complex examples. Also — and this is an important point — traditional logicians rarely give us rules for paraphrasing. In the few cases where they do, the rules don't go deep into the syntax.

There are examples and references for all this in my paper [5]. In that paper I use 'top-level processing' as a name for the traditional belief — sometimes explicit and often implicit — that rules of logic have to apply to the top syntactic level of the expressions involved. Let me take up one of the examples described in that paper; it shows one of the most powerful and clear-headed attempts by a traditional logician to get around the restrictions imposed by top-level processing.

The example comes from Leibniz in the late 17th century. He wanted to justify the inference

(4) Painting is an art. Therefore a person who studies painting studies an art.

The problem is that in the second sentence, 'painting' has dropped to the position of object in a subordinate clause. Leibniz thought that the core issue was that in object position 'painting' is in an oblique case, i.e. not in the nominative case, either in Latin or in German. (This point is invisible in English.)

Leibniz understood that by quantifier rules (which happen not to be deep in our sense), it suffices to show:

(5)

All painting is an art. Titius studies some painting. Therefore Titius studies some art.

This brings 'painting' up into the main clause, but it is still not in the nominative case. Here is the paraphrase that Leibniz uses to solve the problem: (6)

All painting is an art. Some painting is a thing that Titius studies. Therefore some art is a thing that Titius studies.

The step of paraphrasing rests on what Leibniz sometimes calls 'linguistic analysis', and not on syllogistic logic ([7] p. 479f):

(7)

It should also be realized that there are *valid non-syllogistic inferences* which cannot be rigorously demonstrated in any syllogism unless the terms are changed a little, and this altering of the terms is the non-syllogistic inference. There are several of these, including arguments from the nominative to the oblique ... (Lebinz's emphasis)

Leibniz never offers rules for carrying out this kind of paraphrase. If he had done, I very much doubt they would have been deep.

Here is one reason why they would probably not have been deep. Leibniz is hoping to use paraphrase so as to extend the scope of a particular syllogistic rule,

Given $\alpha \rightarrow \beta$ and $\varphi(\alpha)$, if α is positive in $\varphi(\alpha)$, then infer $\varphi(\beta)$.

('Positive' appears as 'affirmative' in Leibniz's discussion.) If he had a deep paraphrasing rule to generalise the example above, he would have needed a method for recognising when an expression arbitrarily deep in the structure of a sentence is occurring positively. Maybe new discoveries will refute me, but I don't believe any general method for this was even considered before the twentieth century. (Special cases are mentioned by John of Salisbury in the twelfth century and Frege in the nineteenth.)

We will see below that Boole himself didn't regard his rule as beloning to traditional logic.

4.3 Boole's rule in modern calculi

Frege in his *Begriffsschrift* of 1879 ([4] §20, p. 50) gets the effect of Boole's rule by using modus ponens together with the axiom schema

(8)
$$c = d \rightarrow (\varphi(c) \rightarrow \varphi(d))$$

(our notation) where φ is a formula of any complexity. We can derive Boole's rule from (8) by considering the case

$$s = t \rightarrow ((\psi(s) = \psi(s)) \rightarrow (\psi(s) = \psi(t)))$$

and applying the axiom $\psi(s) = \psi(s)$ ([4] §21, p. 50) and propositional rules.

Not all modern calculi consider equality as a logical notion. For example Prawitz's *Natural Deduction* has no rules or logical axioms for equality. But Prawitz still has deep rules for the quantifiers, for example his rule $\forall I$) ([12] p. 20):

$$\frac{A}{\forall x A^{a}}$$

where the variable *x* can be buried arbitrarily far down in the formula *A*.

As far as I know, every general purpose calculus proposed for first-order predicate calculus has used deep rules for quantifiers. This is true even for the resolution calculus, where all the sentences have the form $\forall x_1 \dots \forall x_n \theta$ with θ quantifier-free. The reason is that in order to bring arbitrary sentences to this form we need to introduce Skolem functions, and so the variables may occur inside arbitrarily complex Skolem terms.

There are logical calculi that have Boole's rule only in a shallow form, and use the quantifier rules to take up the slack. One example is the logical calculus in Shoenfield's *Mathematical Logic* [13] p. 21.

Could there be a sound and complete proof calculus for predicate logic which has no deep rules? Curiously the answer is yes, but only in a roundabout way and by introducing extra symbols. For example to handle the application of Boole's rule to the term fghjk(x), we could introduce new function symbols *m*, *n*, *o*, *p* and axioms

 $m(x) = j(k(x)), \quad n(x) = h(m(x)), \quad o(x) = g(n(x)), \quad p(x) = f(o(x)).$

Then the application of Boole's rule is equivalent to deducing p(s) = p(t) from s = t, and this uses only top-level substitutions. Skolem showed that we can break down arbitrarily complex formulas in a similar way by adding new relation symbols. With his added symbols only shallow quantifier rules are needed.

In a way this is cheating. We eliminate deep rules, but only at the cost of changing the language. But the point is interesting, because this introduction of new symbols corresponds to part of what traditional logic handled by paraphrasing. (But only part of it. In [5] I gave several examples of traditional paraphrases that alter the domain.)

There is no evidence that Boole himself had any conception of a kind of logic that needs deep rules. He says that the 'purport' of his discussion of his rule 'will be more apparent to the mathematician than to the logician' ([1] p. 69). This is a good moment for us to go back to Boole and ask what he thought he was doing with his rule.

5 Boole's own understanding of his rule

We start with two negative points. Boole didn't regard his rule as justified either by 'common reason' or by the definitions of the expressions involved.

5.1 Not a rule of 'common reason'

At the end of his Preface Boole says ([1] p. 2):

(9)

In one respect, the science of Logic differs from all others; the perfection of its method is chiefly valuable as an evidence of the speculative truth of its principles. To supersede the employment of common reason, or to subject it to the rigour of technical forms, would be the last desire of one who knows the value of ... intellectual toil ...

With very few exceptions, traditional aristotelian logic had no metatheorems. Logicians deduced results by chains of reasoning where every step was obvious to 'common reason'. The very few metatheorems that one does find in traditional logic (like the *peiorem* rule of Theophrastus or the Laws of Distribution) are essentially summaries of families of facts that we can check directly. Aristotelian logicians saw themselves as codifying our inbuilt rules of reasoning, not finding new ways of reasoning to the same conclusions.

Boole's remark about the purport of his rule being more apparent to mathematicians than to logicians should be read in this context. Apparently he thought of his rule as a mathematical 'technical form', not an instance of common reason. We can see this from the fact that he felt a need to justify its use mathematically. He says ([1] p. 69):

(10)

The purport of the last investigation will be more apparent to the mathematician than to the logician. As from any mathematical equation an infinite number of others may be deduced, it seemed to be necessary to shew that when the original equation expresses a logical Proposition, every member of the derived series, even when obtained by expansion under a functional sign, admits of exact and consistent interpretation.

There is more to unpick here than I can handle in this short essay. But if we look at the context, it is clear that he is saying that his mathematical discussion on p. 68 shows that certain consequences of Boole's rule 'admit of exact and consistent interpretation'. So in some sense he is justifying the rule. The discussion on p. 68 uses a metatheorem that he derived on p. 60f by means of a highly speculative application of Maclaurin's theorem. The general form of his argument on p. 68 is: Boole's rule applied to equations E gives equations F. If we paraphrase the equations F by means of Maclaurin's theorem, we can see that what they say is a special case of what the equations E said. So all's well with the world.

If this is how he proposes to justify his rule, then he clearly doesn't regard the rule as belonging to 'common reason'. This is interesting because of the contrast with Frege's view in the *Begriffsschrift*. Frege certainly accepted the traditional aristotelian view that we mentioned after quotation (9) above. In his view, logic starts with self-evident truths and deduces from them other truths by means of deduction rules that are self-evidently correct (where 'correct' means that they never lead from truth to falsehood). And we saw earlier that Boole's rule is virtually one of Frege's logical axiom schemas. So Frege would surely have regarded Boole's rule as self-evidently true.

My personal sympathies are entirely with Frege on this one. I have been trying — without any success so far — to interest some cognitive scientists in the question, since I regard self-evidence as a cognitive notion. (Not everyone does.)

5.2 Not based on definition

Today we might well justify Boole's rule by stating the necessary and sufficient conditions for an equation to be true, and then showing (probably by induction on the complexity of ψ) that the rule applied to a true equation always yields a true equation. I failed to find in Boole any hint of a justification along these lines.

A look at Boole's historical context may throw some light on this. The next subsection will give some of the evidence for Boole's debt to George Peacock on questions of foundations. So it was interesting to see how unclear Peacock is about equations. On p. 8f of the 1830 edition of his *Treatise on Algebra* [8] he says:

(11)

The sign =, placed between two quantities or expressions, indicates that they are equal or equivalent to each other: it may indicate the identity or absolute equality of the quantities between which it is placed: or it may shew that one quantity is equivalent to the other, that is, if they are both of them employed in the same algebraic operation, they will produce the same result: or it may simply mean, as is not uncommonly the case, that one quantity is the result of an operation, which in the other is indicated and not performed.

Here he distinguishes three notions: (1) 'a = b' means that the quantity a is 'identical' with the quantity b, (2) 'a = b' means that if F is any algebraic operation then F(a) is 'the same' as F(b), (3) 'a = b' means that b is the result of performing the operation indicated by a. This is chaotic. For example, what is the difference between 'equal', 'identical' and 'the same'? How are we to tell whether '2 + 2 = 4' means that 2 + 2 is identical with 4, or that the result of adding 2 to 2 is 4? The chaos continues into Peacock's second edition twelve years later ([10] p. 4):

(12)
= [denotes] equality, or the result of any operation or operations.
... The sum of 271, 164, and 1023, or the result of the addition of these numbers to each other, is equal to 1458.

Here 'the result of the addition' and 'is equal to' appear in the same clause, conflating two of his previous notions. (I warmly thank Marie-José Durand-Richard for helping me with these references, though she may not agree with the conclusion I draw from them.)

Note also Peacock [10] p. 198: Given

$$a_1A_1 = \alpha_1A_2, \quad a_2A_2 = \alpha_2A_3, \quad \dots \quad a_{n-1}A_{n-1} = \alpha_nA_n$$

he finds the value *x* of $a_n A_1 / A_n$ as

$$x = (\alpha_1 \alpha_2 \dots \alpha_n / a_1 a_2 \dots a_{n-1})$$

Remarkably, his proof removes the = altogether and uses the theory of proportions.

In short it seems that Peacock had several notions of what an equation is, none of them very precise, and he saw no need to clarify the relations between these notions. My guess is that he could get away with this because he thought of the mathematical content as living in the terms and their interpretation; the equation sign, where it wasn't just part of an algorithm, was a device that was useful for commenting on the mathematics, but wasn't strictly part of the mathematics. But here I am speculating. The non-speculative point is that Peacock is evidence for a mathematical environment in which it would have seemed quite unnatural to justify Boole's rule by reference to the definition of 'equation'.

5.3 Based on absence of contradiction

Boole's remark (10) about 'consistent interpretation' was not meant lightly. Already on page 4 of [1] he had said

(13)

We might justly assign it as the definitive character of a true Calculus, that it is a method resting upon the employment of Symbols, whose laws of combination are known and general, and whose results admit of a consistent interpretation.

So there is good reason to hope that his notion of 'consistent interpretation' will throw light on his view of Boole's rule.

The notion of 'consistent interpretation' comes from Peacock. In Peacock it means something fairly precise: a 'consistent interpretation' of + and - is one that (i) applies to a class *C* of quantities that contains the natural numbers, and agrees with the interpretation of these symbols on the natural numbers, and (ii) makes true in *C* the basic identities of algebra that were true in the natural numbers. For example when + and - are given their usual interpretations on the integers, the distributive law and the identity x - x = 0 (both of which were true on the natural numbers) remain true even when the variables are interpreted as standing for integers, possibly negative. This seems to be the meaning of 'consistent interpretation' at [8] p. xxvii, [9] p. 226, [10] p. vii and [11] footnote p. 10. De Morgan picked up the phrase; at [3] p. 208 he says 'I believe that symbolic algebra will never cease to dictate results which must be capable of consistent interpretation'. Andrew Bell used it in his *Elements of Algebra* from 1839.

Boole was certainly happy to ally himself with Peacock's symbolical algebra. The opening words of his [1] (p. 3) are:

(14)

They who are acquainted with the present state of the theory of Symbolical Algebra, are aware, that the validity of the processes of analysis does not depend upon the interpretation of the symbols which are employed, but solely upon the laws of their combination. Every system of interpretation which does not affect the truth of the relations supposed, is equally admissible ...

We saw in (1) that Maria would prefer to say that Boole's debt was to Duncan Gregory rather than to Peacock. Boole studied with Gregory, and very likely he learned Peacock's work through Gregory. But I know of nothing in Boole's use of symbolical algebra that he could have got from Gregory better than from Peacock.

To return to Boole's use of 'consistent interpretation': Boole can't mean exactly the same by it as Peacock did. In both the passages (10) and (13) he is talking about the 'consistent interpretation' of derived results in a calculus, and this would make no sense in Peacock's usage (at least as I read Peacock). But the context allows us to read Boole as meaning something similar to Peacock but a little looser. Boole means not just that the usual identities come out true ((ii) above), but also that when standard mathematical transformations are applied, the results never contradict each other. This is what he was showing on his page 68. There he showed that some results of applying Boole's rule and some results of applying Maclaurin's theorem are consistent with each other, when they are read in terms of Boole's logical interpretation of the elective symbols.

I believe Boole's view is as follows. We reason in certain ways. These ways can lead us to contradict ourselves. But ([2] p. 160):

(15)

we are nevertheless so formed that we can, by due care and attention, perceive when [logical consistency] is violated, and when it is regarded.

Thus we have it in our power to avoid contradictions; and this is our best guarantee of the 'truth' of a calculus. I think this is exactly what Boole is saying at the quotation (13).

Thus: neither Boole's rule nor Maclaurin's theorem is an example of 'common reason'. But both of them come naturally to any trained mathematician, because they are used all over the place in analysis. To justify their application to logic, the best test is that taken together, they never yield contradictory results. Of course Boole's calculation on p. 68 doesn't prove that no contradictions arise. But as with set theory today, the more we apply 'due care and attention' without finding any contradictions, the less likely it is that there are any.

In short, Boole adopted Boole's rule because it was used in analysis and it didn't give any trouble when it was transferred to logic. The modern notion of a rule of deduction, which demonstrably never leads from truths to falsehoods, is nowhere to be seen.

References

[1] George Boole, *The Mathematical Analysis of Logic: being an Essay towards a Calculus of Deductive Reasoning*, Macmillan, Barclay, & Macmillan, Cambridge 1847.

[2] George Boole, 'On belief in its relation to the understanding', in *George Boole: Selected Manuscripts on Logic and its Philosophy*, ed. I. Grattan-Guinness and G. Bonnet, Birkhäuser, Basel 1997, pp. 157-161.

[3] Augustus De Morgan, 'On fractions of vanishing or infinite terms', *The Quarterly Journal of Pure and Applied Mathematics* 1 (1857) 204-210.

[4] Gottlob Frege, Begriffsschrift, Nebert, Halle 1879.

[5] Wilfrid Hodges, 'Traditional logic, modern logic and natural language', *Journal of Philosophical Logic* 38 (2009) 589-606.

[6] Wilfrid Hodges, 'Tarski on Padoa's method: A test case for understanding logicians of other traditions', in *Logic, Navya-Nyāya and Applications: Homage to Bimal Krishna Matilal*, ed. Mihir K. Chakraborty et al., College Publications, London 2008, pp. 155-169.

[7] G. W. Leibniz, *New Essays on Human Understanding*, translated and edited by Peter Remnant and Jonathan Bennett, Cambridge University Press, Cambridge 1996.

[8] George Peacock, A Treatise on Algebra, Deighton, Cambridge 1830.

[9] George Peacock, 'Report on the recent progress and present state of certain branches of analysis', in *Report of the Third Meeting of the British Association for the Advancement of Science*, John Murray, London 1834, pp. 185-352.

[10] George Peacock, A Treatise on Algebra Volume I: Arithmetical Algebra, Deighton, Cambridge 1842.

[11] George Peacock, A Treatise on Algebra Volume II: On Symbolical Algebra, Deighton, Cambridge 1845.

[12] Dag Prawitz, Natural Deduction, Almqvist & Wiksell, Stockholm 1965.

[13] Joseph Shoenfield, Mathematical Logic, Addison-Wesley, Reading Mass. 1967.