George Boole and the Development of Probability Theory

Writing in the preface to the first edition of *The Logic of Chance* (1866), John Venn observed that little attention had been paid by mathematicians to the fundamental principles of probability theory. He wrote:

> *With regard to the remarks of the last few paragraphs, prominent exceptions must be made in the case of two recent works at least. The first of these is Professor De Morgan's Formal Logic. ... The other work to which I refer is the profound Laws of Thought of the late Professor Boole, to which somewhat similar remarks may in part be applied. Owing however to his peculiar treatment of the subject, I have scarcely anywhere come into contact with any of his expressed opinions.*

Given Boole's construction of an algebra of logic to investigate logical processes, it is perhaps not surprising that Boole should have been drawn to the study of probability theory. Boole's algebra of logic, in which algebraic techniques are applied to symbols representing classes, was eventually abstracted and systematized by others into the concept of a Boolean algebra. In particular, the algebra of subsets of a set, with its operations of intersection, union and complementation, is such a Boolean algebra. The modern theory of probability, at least as applied to discrete sample spaces, is expressed using the subsets of the sample space, with the operations above playing an important role.

Boole's greatest claim to fame is surely based on his book *An Investigation of the Laws of Thought on which are founded the Mathematical Theories of Logic and Probabilities*, or, more briefly *The Laws of Thought*, published in 1854. This book is 424 pages long, and six chapters of it, comprising 155 pages, are devoted to probability theory. Boole published a number of papers on probability theory between 1851 and 1862 (two years before his death) but nothing further on logic after *The Laws of Thought*. However, as we saw above, Venn spoke of Boole's *peculiar treatment of the subject* and it must be said that there are certain aspects of Boole's work on probability theory which seem extraordinary to the modern reader and have been the subject of much controversy since their first appeaance in the 1850's. In particular, one problem and its proposed solution by Boole have been investigated by several researchers over many years and it is this problem that we intend to discuss.

In *The Cambridge and Dublin Mathematical Journal*, vol. vi, November 1851, Boole published a short paper entitled *Proposed question in the theory of probabilities*. We quote:

> *Of those rigorous consequences of the first principles of the theory of probabilities the general utility of which has caused them to be ranked by Laplace among the great secondary principles of the science, none is more important than the following:-If an event E can*

only happen as the result of one of certain conflicting causes A_1, A_2, ..., A_n, then if c_i represent the probability of A_i, and p_i the probability that if A_i happen E will happen, the total probability of the event E will be represented by the sum $\sum c_i p_i$.

I am desirous of calling the attention of mathematicians to a question closely analogous to that of which the answer is conveyed in the above theorem; like it also, admitting of rigorous solution and susceptible of wide application. The question is the following:-If an event E can only happen as a consequence of some one or more of certain causes A_1, A_2, ..., A_n and if generally c_i represent the probability of the cause A_i, and p_i the probability that if the cause A_i exist the event E will exist, then the series of values c_1, c_2, ..., c_n, p_1, p_2, ..., p_n being given, required the probability of the event E.

This problem was restated as Problem VI on p.336 of *The Laws of Thought*.

In more modern terminology, we take $p(A_i) = c_i$ and $p(E|A_i) = p_i$. Then $p(E \cap A_i) = c_i p_i$. We assume that the statement in the first paragraph of Boole's paper implies that the events A_1, ..., A_n are pairwise mutually exclusive and that

$$E = E \cap A_1 \cup \ldots \cup E \cap A_n.$$

Then

$$p(E) = p(E \cap A_1) + \cdots + p(E \cap A_n) = c_1 p_1 + \cdots + c_n p_n.$$

Presumably, in the problem of the second paragraph, there is no implied assumption of being mutually exclusive.

Boole continued:

The motives which have led me, after much consideration, to adopt with reference to this question a course unusual in the present day, and not upon slight grounds to be revived, are the following. First, I propose the question as a test of the sufficiency of received methods. Secondly, I anticipate that its discussion will in some measure add to our knowledge of an important branch of pure analysis. However, it is upon the former of these grounds alone that I desire to rest my apology.

In *The Philosophical Magazine*, vol vi, October 1853, Cayley published a short paper entitled *On a question in the theory of probabilities.* He began:

The following question was suggested, either by some of Prof. Boole's memoirs on the subject of probabilities, or in conversation with him; I forget which; it seems to me a good instance of the class of questions to which it belongs.

Given the probability α that a cause A will act, and the probability p that A acting, the effect will happen; also the probability β that a cause B will act, and the probability q that B acting the effect will happen; required the probability of the effect.

The solution as given by Cayley is difficult to understand and little explanation is given. He proceeds:

Let λ be the probability that the cause A will act efficaciously; μ the probability that B acting will act efficaciously. Then

$$\begin{aligned} p &= \lambda + (1-\lambda)\mu\beta \\ q &= \mu + (1-\mu)\alpha\lambda \end{aligned}.$$

which determine λ and μ. The total probability ρ of the effect is given by

$$\rho = \lambda\alpha + \mu\beta - \lambda\mu\alpha\beta.$$

He then demonstrates that his method gives the correct answer when $\alpha = 1$. It is not at all clear what λ and μ represent, although some explanation is given in terms of an example involving a stormy, rainy and windy day.

Boole contributed a reply to Cayley's solution in *The Philosophical Magazine*, vol vii, January 1854, with his own paper entitled *Solution of a question in the theory of probabilities.* He objects to its solution, although admitting that it is true in certain special cases. He writes:

I think it to be one of the peculiar difficulties of the theory of probabilities, that its difficulties sometimes are not seen. The solution of a problem may appear to be conducted according to the principles of the theory as usually stated; it may lead to a result susceptible of verification in particular instances; and yet it may be an erroneous solution. The problem which Mr. Cayley has considered seems to me to afford a good illustration of this remark. Several attempts at its solution have been forwarded to me, all of them by mathematicians of great eminence, all of them admitting of particular verification, yet differing from each other and from the truth. Mr. Cayley's solution is the only published one I have seen, and I feel I must extend to it the same observations. But in doing this, I willingly add that I have two or three times attempted to solve the problem by the same kind of reasoning, and have not approached so near the truth as Mr. Cayley has done. To illustrate these remarks, I will first complete Mr. Cayley's solution, and give one or two apparent verifications, then exhibit the true solution; and lastly, make a few observations upon the general subject.

Denoting the required probability by u, rather than Cayley's ρ, Boole shows that, according to Cayley, u is a solution of the quadratic equation

$$\frac{(1-\alpha(1-p)-u)(1-\beta(1-q)-u)}{1-u} = (1-\alpha)(1-\beta).$$

This already seems curious, as the general assumption in probability theory is that all probabilities are rational numbers, whereas the solution above may involve a quadratic irrationality. (W. A. Whitworth, writing in his text *Choice and Chance*, 5th edition, 1901, tried to show in a particular case how one might argue if a particular probability were not rational (p. 179).)

Boole gives what he states is the true solution of the problem. It is a root u of the quadratic equation

$$\frac{(1-\alpha(1-p)-u)(1-\beta(1-q)-u)}{1-u} = \frac{(u-\alpha p)(u-\beta q)}{\alpha p+\beta q-u}.$$

Moreover, this root satisfies

$$u \geq \alpha p, \quad u \geq \beta q, \quad u \leq 1-\alpha(1-p), \quad u \leq 1-\beta(1-q), \quad u \leq \alpha p+\beta q$$

and is the unique root satisfying these inequalities.

He finishes the paper as follows:

> *. . . and of this I am fully assured, that no* general *method for the solution of questions in the theory of probabilities can be established which does not explicitly recognise, not only the special numerical bases of the science, but also those universal laws of thought which are the basis of all reasoning, and which, whatever they may be as to their essence, are at least mathematical as to their form. Such a method I have exhibited in a treatise now on the eve of publication, and to which I must refer for the investigation of the problem, the solution of which has been exemplified in this paper.*

In *The Laws of Thought*, Boole presented his method, alluded to above, for calculating probabilities, which is based on his logical calculus developed in the earlier part of the book. The method is most idiosyncratic and not at all easy to understand. In his book *A Treatise on Probability Theory*, (1921), J. M. Keynes, the famous economist, wrote:

> *In the following paragraph solutions are given of some problems posed by Boole in chapter XX of his* Laws of Thought*. Boole's own method of solving them is constantly erroneous, and the difficulty of his method is so great that I do not know of anyone but himself who has ever attempted to use it.*

Keynes states that the solutions of problems I-VI in chapter XX of *The Laws of Thought* are all erroneous. He also gives a good description, with valuable references, of this whole argument concerning the solutions of Boole's problems and their validity. Boole had corresponded with De Morgan on his problem and there seems to have been much discussion on what the precise hypotheses of the problem actually were. Perhaps De Morgan was uneasy about Boole's solutions and method, as Boole wrote in a letter to De Morgan of 23 February 1854:

...But at any rate satisfy your self on this point–whether the solutions my principle gives are ever false. If you find one instance in which they are I give it up. Are you satisfied with this declaration? I am sure that if there is any quality that I think you have in preeminence it is integrity in pursuit of the truth–but that is a quality in which I should be sorry to think myself your inferior. I don't think any man's mind ever was imbued with a more earnest desire to find out the truth and say it and nothing else, than mine while writing that book. And the very consciousness of this would make it not *painful to me to give up half my book if it were proved to be unfounded. However what I now ask of you both as a friend of truth & of me is to examine the questions fully– to settle it in your mind to make out whether I am right or wrong.*

In view of later well-directed criticism of Boole's solutions to probability theory problems in his book, he may in fact have been much less humble about his achievements than the sentiments expressed in this letter suggest.

Problem I of p.321 of *The Laws of Thought* is the one that caused most controversy and is related to Cayley's problem, although stated rather differently. It is:

The probabilities of two causes A_1 and A_2 are c_1 and c_2 respectively. The probability that if the cause A_1 present itself, an event E will accompany it (whether as a consequence of the cause A_1 or not) is p_1, and the probability that if the cause A_2 present itself, that event E will accompany it, whether as a consequence of it or not, is p_2. Moreover, the event E cannot appear in the absence of both the causes A_1 and A_2. Required the probability of the event E.

Denoting complementary events by primes, our data is

$$p(A_1) = c_1, \quad p(A_2) = c_2, \quad p(E|A_1) = p_1, \quad p(E|A_2) = p_2$$

and furthermore $E \cap A_1' \cap A_2' = \emptyset$. We must then find $p(E)$. Clearly, $p(E \cap A_1) = c_1p_1$ and $p(E \cap A_2) = c_2p_2$.

The solution given by Boole is that contained in his 1854 paper, involving the roots of a quadratic equation. As we remarked before, this seems suspicious as the probability may require a quadratic irrationality when it should be rational. Furthermore, what is even more surprising, if we assume that $E \cap A_1$ and $E \cap A_2$ are mutually exclusive, which is certainly compatible with the assumption that $E \cap A_1' \cap A_2' = \emptyset$, the solution is $p(E) = c_1p_1 + c_2p_2$, a result that cannot be obtained from Boole's formula, as it involves a denominator that is 0, whereas none of the other three terms need be 0. Thus, Boole's formula cannot possibly be correct. It is not entirely clear if Boole's problem is exactly the same as Cayley's, but he claims that his solution applies to both. The hypothesis that $E \cap A_1' \cap A_2' = \emptyset$ seems to be missing from the statement of Cayley's problem.

We would proceed to give the solution as follows. Let S denote the sample space. Then

$$S = (A_1 \cup A_2) \cup (A_1 \cup A_2)' = (A_1 \cup A_2) \cup (A_1' \cap A_2')$$

and thus we obtain

$$E = E \cap A_1 \cup E \cap A_2$$

from the data. Now we have

$$\begin{aligned} p(E) &= p(E \cap A_1 \cup E \cap A_2) = p(E \cap A_1) + p(E \cap A_2) - p(E \cap A_1 \cap A_2) \\ &= c_1 p_1 + c_2 p_2 - p(E \cap A_1 \cap A_2). \end{aligned}$$

Since we have no information about the event $E \cap A_1 \cap A_2$, the problem is indeterminate.

The publication of Boole's book in February, 1854, with its description of Boole's 'logical method' for calculating probabilities occasioned a most penetrating criticism by Henry Wilbraham, published in *The Philosophical Magazine*, supplement to vol. vii, June 1854. Wilbraham's paper has the title *On the theory of chances developed in Professor Boole's "Laws of Thought"*. The Royal Society catalogue of scientific papers mentions six more papers by Wilbraham, published between 1848 and 1857, none of which deals with probability. The last two papers appeared in the *Assurance Magazine* of 1857, one having the title *On the possible methods of dividing the net profits of a Mutual Life Assurance Company amongst the members.* Wilbraham (1825–1883) was educated at Harrow school and graduated at Cambridge in 1846 as seventh wrangler. He was elected a fellow of Trinity College, Cambridge in 1848. He was called to the Bar in 1851 and served as Registrar of the Court of Chancery for Lancashire and the Manchester district.

Wilbraham wrote:

> ... *The object of this paper is to show that Professor Boole does in a great number of questions relating to chances solvable by his method (or at least in those which are most difficult to treat by other methods), tacitly assume certain conditions expressed by the data of the problem, and to show how these assumed conditions may be algebraically expressed.*

He points out algebraically how Boole often makes assumptions in chapter XVII of his book that events are independent, without these assumptions being clearly stated in the data of the problem. He also shows how Boole makes certain tacit assumptions that enable him to solve problems that are otherwise indeterminate. His most damning condemnation of Boole's methods concerns Problem I, discussed above. In an analysis that seems much modern in spirit than anything displayed by his contemporary workers in probability theory, Wilbraham shows that the solution to the problem is precisely what has been given above, and therefore the problem cannot be solved. Curiously, Keynes states:

...Wilbraham gave as the solution $u = c_1 p_1 + c_2 p_2 - z$, where z is necessarily less than either $c_1 p_1$ or $c_2 p_2$ This solution is correct so far as it goes, but is not complete.

Here, u is just $p(E)$. In fact, Wilbraham says precisely what z is and his solution is complete. Keynes was well abreast of virtually all the literature on probability theory up to the end of the 19th century but he seems to have neglected to read Wilbraham's article in detail.

Wilbraham goes on to determine what he thinks are the assumptions tacitly made by Boole to render an indeterminate problem determinate. These are the two equations

$$\frac{p(A_1 \cap A_2 \cap E)}{p(A_1' \cap A_2 \cap E)} = \frac{p(A_1 \cap A_2' \cap E)}{p(A_1' \cap A_2' \cap E')}$$

and

$$\frac{p(A_1 \cap A_2 \cap E')}{p(A_1' \cap A_2 \cap E')} = \frac{p(A_1 \cap A_2' \cap E')}{p(A_1' \cap A_2' \cap E')}.$$

It is very tedious to check, but we found that the numerical information contained in these two equations, taken in conjunction with the probabilities deducible from the original data of the problem, does indeed lead to precisely the quadratic equation claimed by Boole to provide the solution.

Wilbraham remarks that the second equation, though perfectly arbitrary, is not unreasonable. But he states:

But the first of these equations appears to me not only arbitrary but eminently anomalous. In the form in which it stands as a relation among the chances of A_1, A_2 and E, no one, I should think, can contend that it is either deduced from the data of the problem or that the mind by the operation of any law of thought recognizes it as a necessary or most reasonable assumption.

He also subjects Cayley's solution of his probability problem to a similar analysis and finds that two similar equations involving quotients of probabilities have been tacitly introduced in order to obtain a solution. He finishes his paper with a withering, but unanswerable, analysis of Boole's logical method.

What, now, is the practical value of Professor Boole's logical method as applied to the theory of chances? In cases determinable by ordinary algebraical processes, his book gives a systematic and uniform method of solving the questions, though commonly a longer one than we should otherwise use; at least it appears to me that the really determinate problems solved in the book, as 2 and 3 of Chapter XVIII, might be more shortly solved without the logical equations. In these cases the originally assumed independence of simple events is unnecessary, none of the equations thereby consisting wholly of terms comprised in V. The disadvantage of Professor Boole's method in such cases is, that it does not show us whether *the problem is really determinate or requires further assumptions,- whether, in fact, the assumptions made are necessary or not. On the other hand, in cases not determinable by ordinary algebra, his system is this; he takes a general indeterminate problem, applies to it particular assumptions not definitely stated in his book, but which may be shown, as I have done, to be implied in his method, and with these assumptions solves it; that is to say, he solves a particular determinate case of an indeterminate problem, while his book may mislead the reader by making him suppose that it is the general problem which is being treated of. The question arises, Is the particular case thus solved a peculiarly valuable one, or one more worthy than any other of being solved? It is clearly not an assumption that must in all cases be true; nor is it one which, without knowing the connexion among the simple events, we can suppose more likely than any other to represent that connexion;...*

It seems to us that Wilbraham's analysis is most perspicacious and accurate, if perhaps a little sharp in its delivery. Its appearance spurred Boole to make a vigorous reply in the pages of *The Philosophical Magazine*. His first response to Wilbraham appeared in vol. viii, August 1854. He begins with the self-effacing remark:

Controversy is in every way so disagreeable to me, that it is with the most unfeigned reluctance I feel myself called upon to reply to the observations of Mr. Wilbraham inserted in the last Number of your Journal.

Quoting parts of Wilbraham's paper, he continues:

...I fear that the impression produced upon the mind of any person not acquainted with my work by such statemnts as the above would be, that I have introduced in a covert manner assumptions of the existence of which I was ignorant, or of the recognition of which I was afraid. It may be therefore right for me to state that I have, in the chapter containing the the demonstration of the general method for the solution of questions in probabilities ..., explicitly stated the principles upon which that demonstration proceeds, and with equal explicitness deduced from them the algebraical equations upon which the solution depends.... To prove that particular assumptions not definitely stated in my book are employed, it ought, I conceive, to have been shown that the principles which I have expressly stated are insufficient for the conclusions drawn from them. ...I...desire to consider simply whether Mr. Wilbraham's strictures affect in any way the validity of the method which I have published.

Concerning Wilbraham's detailed criticism of his solution of Problem I, he writes:

...Now I cannot but think that a cautious inquirer after truth, seeing that two hypotheses (still adopting Mr. Wilbraham's language), one of which appears to him "eminently anomalous", conduct to a solution which cannot by any known test be proved erroneous,

while two other hypotheses [those in Cayley's solution], which appear to him "perhaps not unreasonable" ... conduct to a solution which will not bear the test of examination... On the other hand, I affirm without hesitation that there is no case in which the equations deduced by Mr Wilbraham from my method of solution can be proved to be erroneous. They do not, indeed represent "hypotheses", but they are legitimate deductions from the general principles upon which the method is founded, and it is to those principles directly that attention ought to be directed.

I would request your readers to observe that I do not offer the above remarks as affording any proof that the principles upon which my method is established are true, but only as conclusive that Mr Wilbraham's objections against them, drawn from what to him appears to be the anomalous character of an equation to which they lead, are of no value whatever.

He then attempts to give the source of Wilbraham's erroneous judgments, speaking in very general, unquantitative terms. It seems clear that Boole did not really accept Wilbraham's proof that the problem is indeterminate but thinks it is legitimate to use essentially unstated hypotheses to arrive at a solution. It is not clear that he was really aware of the precise hypotheses he had used, in terms of probabilities. This is all the more reprehensible as he clearly states that his is *the* solution to the problem, when, as Wilbraham remarked, it is just one of many possible solutions. He also appears to claim that, as he knows of no case in which the answer he obtains is incorrect, his method is correct, which is surely curious logic for a professional mathematician.

In fact, somewhat later Hugh McColl wrote a paper in the *Journal of the London Mathematical Society*, vol XI, 1880, in which he investigated Boole's problem. He obtained Wilbraham's solution, although he appeared to be unaware of Wilbraham's paper. He argues that one can find cases in which Boole's value for the probability is wrong. He assumes that A_1 and A_2 are independent, and E to be more probable when both A_1 and A_2 exist, than when only one of them exists. He takes

$$c_1 = 0.1, \quad c_2 = 0.2, \quad p_1 = 0.6, \quad p_2 = 0.7.$$

Simple inequalities, whose use Boole had already advocated, lead to the estimate

$$0.18 \leq p(E) \leq 0.186,$$

whereas Boole's solution is 0.190697319 (in fact, Boole's solution involves a quadratic irrationality).

Boole made a second, shorter response to Wilbraham's criticisms in *The Philosophical Magazine*, vol. viii, September 1854. Wilbraham had suggested in his paper an alternative way of making assumptions in order to solve indeterminate problems such as that posed by Boole. (It cannot be said that this proposal seems very plausible.) Concerning Wilbraham's suggestion, Boole wrote:

> *If Mr. Wilbraham's method is both* correct *and* sufficient, *while mine is false, there must surely be some case in which the two would lead to different results, and in which, from the comparison of those results, my own may be proved to be erroneous. I would therefore request Mr. Wilbraham to endeavour to furnish an instance of this kind. . . . Should any method, even of limited application, be discovered which should lead to solutions satisfying the conditions to which I have referred, and yet different from those furnished by my own method, which is* not *of limited application, and which* always *causes those conditions to be satisfied, I should regard it as a very interesting and remarkable circumstance. But at present I am, as I have said, wholly ignorant of the existence of any such method.*

The arguments on the probability question faded from public view until Cayley published a further communication (*On a question in the theory of probabilities*) on the subject in *The Philosophical Magazine*, vol. XXIII, 1862.

Cayley began by resuming the discussion of his 1853 paper, whose solution had been objected to by Boole. He argued that there was a difference in the interpretation of the question, his solution referring to the 'causation' statement, with an assumed independence of events, whereas Boole's referred to the 'concomitance' statement, and stated that he thought Boole would agree with this. He drew attention to a paper of Dedekind of 1855, *Bemerkungen zu einer Aufgabe der Wahrscheinlichkeitsrechnung*, in *Crelle's Journal*, vol. L, 1855. In this paper, Dedekind had defended Cayley's paper and shown how, by making use of inequality relations and choosing the correct sign for the radical, a solution for the problem could be obtained.

Cayley then reproduced Boole's method of solution

> *. . . without attempting to explain (indeed I do not understand to my own satisfaction) the logical principles upon which it is based.*

He submitted a preliminary version of this paper to Boole, who replied a few days later with a series of eight observations or clarifications relating to Cayley's solution and his own method of solution, which Cayley included in the final version of the paper, together with his answers to Boole's points. In particular, Cayley expressed himself unable to understand what Boole meant by the probabilities of 'ideal events' in an 'ideal problem' that he considered.

Boole replied that it was not possible to explain what was the nature of the auxiliary quantities (the probabilities of ideal events) used in his solution. Furthermore, he wrote:

...I do not see any difficulty whatever in the conception of the ideal problem.

Cayley countered:

We thus join issue as follows: Prof. Boole says that there is no difficulty in understanding, I say that I do not understand, the rationale *of his solution.*

In a postscript to this paper, Cayley wrote:

I unaccountably did not recall to myself Mr. H. Wilbraham's paper ..., which contains a most valuable discussion of the question.

He then presented the equations that Wilbraham had shown to underly Boole's solution (and the different equations underlying his own solution) and says of one of these equations

...But it is not easy to explain the first of the equations (b); indeed Mr. Wilbraham remarked that it appeared not only arbitrary but eminently anomalous. The peculiarity in its form is, that it does not, like the others, when ABE, &c are considered as products, reduce itself to an identity; it seems to be a conclusion which, in support of his theory, Prof. Boole is bound to justify à posteriori.

The statement about ABE being regarded as products appears to refer to what happens if all events occurring in the equations are treated as independent.