CHAPTER XVIII

INTRODUCTION

Nothing so like as eggs; yet no one, on account of this apparent similarity, expects the same taste and relish in all of them. 'Tis only after a long course of uniform experiments in any kind, that we attain a firm reliance and security with regard to a particular event. Now where is that process of reasoning, which from one instance draws a conclusion, so different from that which it infers from a hundred instances, that are no way different from that single instance? This question I propose as much for the sake of information, as with any intention of raising difficulties. I cannot find, I cannot imagine any such reasoning. But I keep my mind still open to instruction, if any one will vouchsafe to bestow it on me.—HUME.[1]

1. I HAVE described Probability as comprising that part of logic which deals with arguments which are rational but not conclusive. By far the most important types of such arguments are those which are based on the methods of Induction and Analogy. Almost all empirical science rests on these. And the decisions dictated by experience in the ordinary conduct of life generally depend on them. To the analysis and logical justification of these methods the following chapters are directed.

Inductive processes have formed, of course, at all times a vital, habitual part of the mind's machinery. Whenever we learn by experience, we are using them. But in the logic of the schools they have taken their proper place slowly. No clear or satisfactory account of them is to be found anywhere. Within and yet beyond the scope of formal logic, on the line, apparently, between mental and natural philosophy, Induction has been admitted into the organon of scientific proof, without much help from the logicians, no one quite knows when.

2. What are its distinguishing characteristics? What are the qualities which in ordinary discourse seem to afford strength to an inductive argument?

[1] *Philosophical Essays concerning Human Understanding.*

I shall try to answer these questions before I proceed to the more fundamental problem—What ground have we for regarding such arguments as rational?

Let the reader remember, therefore, that in the first of the succeeding chapters my main purpose is no more than to state in precise language what elements are commonly regarded as adding weight to an empirical or inductive argument. This requires some patience and a good deal of definition and special terminology. But I do not think that the work is controversial. At any rate, I am satisfied myself that the analysis of Chapter XIX. is fairly adequate.

In the next section, Chapters XX. and XXI., I continue in part the same task, but also try to elucidate what sort of assumptions, *if* we could adopt them, lie behind and are required by the methods just analysed. In Chapter XXII. the nature of these assumptions is discussed further, and their possible justification is debated.

3. The passage quoted from Hume at the head of this chapter is a good introduction to our subject. Nothing so *like* as eggs, and after a *long* course of uniform experiments we can expect with a firm reliance and security the same taste and relish in all of them. The eggs must be like eggs, and we must have tasted many of them. This argument is based partly upon *Analogy* and partly upon what may be termed *Pure Induction*. We argue from Analogy in so far as we depend upon the *likeness* of the eggs, and from Pure Induction when we trust the *number* of the experiments.

It will be useful to call arguments *inductive* which depend in any way on the methods of Analogy and Pure Induction. But I do not mean to suggest by the use of the term *inductive* that these methods are necessarily confined to the objects of phenomenal experience and to what are sometimes called empirical questions; or to preclude from the outset the possibility of their use in abstract and metaphysical inquiries. While the term *inductive* will be employed in this general sense, the expression *Pure Induction* must be kept for that part of the argument which arises out of the repetition of instances.

4. Hume's account, however, is incomplete. His argument could have been improved. His experiments should not have been too uniform, and ought to have differed from one another

as much as possible in all respects save that of the likeness of the eggs. He should have tried eggs in the town and in the country, in January and in June. He might then have discovered that eggs could be good or bad, however like they looked.

This principle of varying those of the characteristics of the instances, which we regard in the conditions of our generalisation as non-essential, may be termed *Negative Analogy*.

It will be argued later on that an increase in the *number* of experiments is *only* valuable in so far as, by increasing, or possibly increasing, the variety found amongst the non-essential characteristics of the instances, it strengthens the Negative Analogy. If Hume's experiments had been *absolutely* uniform, he would have been right to raise doubts about the conclusion. There is no process of reasoning, which from one instance draws a conclusion different from that which it infers from a hundred instances, if the latter are known to be in *no* way different from the former. Hume has unconsciously misrepresented the typical inductive argument.

When our control of the experiments is fairly complete, and the conditions in which they take place are well known, there is not much room for assistance from Pure Induction. If the Negative Analogies are known, there is no need to count the instances. But where our control is incomplete, and we do not know accurately in what ways the instances differ from one another, then an increase in the mere number of the instances helps the argument. For unless we know for certain that the instances are perfectly uniform, each new instance *may* possibly add to the Negative Analogy.

Hume might also have weakened his argument. He expects no more than the same taste and relish from his eggs. He attempts no conclusion as to whether his stomach will always draw from them the same nourishment. He has conserved the force of his generalisation by keeping it narrow.

5. In an inductive argument, therefore, we start with a number of instances similar in some respects AB, dissimilar in others C. We pick out one or more respects A in which the instances are similar, and argue that some of the other respects B in which they are also similar are likely to be associated with the characteristics A in other unexamined cases. The more comprehensive the essential characteristics A, the greater the variety amongst the non-essential characteristics C, and the less comprehensive the characteristics B which we seek to associate with A, the stronger is the likelihood or probability of the generalisation we seek to establish.

These are the three ultimate logical elements on which the probability of an empirical argument depends,—the Positive and the Negative Analogies and the scope of the generalisation.

6. Amongst the generalisations arising out of empirical argument we can distinguish two separate types. The first of these may be termed *universal induction.* Although such inductions are themselves susceptible of any degree of probability, they affirm *invariable* relations. The generalisations which they assert, that is to say, claim universality, and are upset if a single exception to them can be discovered. Only in the more exact sciences, however, do we aim at establishing universal inductions. In the majority of cases we are content with that other kind of induction which leads up to laws upon which we can generally depend, but which does not claim, however adequately established, to assert a law of more than probable connection.[1] This second type may be termed *Inductive Correlation.* If, for instance, we base upon the data, that this and that and those swans are white, the conclusion that *all* swans are white, we are endeavouring to establish a universal induction. But if we base upon the data that this and those swans are white and that swan is black, the conclusion that *most* swans are white, or that the probability of a swan's being white is such and such, then we are establishing an inductive correlation.

Of these two types, the former—universal induction—presents both the simpler and the more fundamental problem. In this part of my treatise I shall confine myself to it almost entirely. In Part V., on the Foundations of Statistical Inference, I shall discuss, so far as I can, the logical basis of inductive correlation.

7. The fundamental connection between Inductive Method and Probability deserves all the emphasis I can give it. Many writers, it is true, have recognised that the conclusions which we reach by inductive argument are probable and inconclusive. Jevons, for instance, endeavoured to justify inductive processes by means of the principles of inverse probability. And it is true also that much of the work of Laplace and his followers was

[1] What Mill calls 'approximate generalisations.'

directed to the solution of essentially inductive problems. But it has been seldom apprehended clearly, either by these writers or by others, that the validity of every induction, strictly interpreted, depends, not on a matter of fact, but on the existence of a relation of probability. An inductive argument affirms, not that a certain matter of fact *is* so, but that *relative to certain evidence* there is a probability in its favour. The validity of the induction, relative to the original evidence, is not upset, therefore, if, as a fact, the truth turns out to be otherwise.

The clear apprehension of this truth profoundly modifies our attitude towards the solution of the inductive problem. The validity of the inductive method does *not* depend on the success of its predictions. Its repeated failure in the past may, of course, supply us with new evidence, the inclusion of which will modify the force of subsequent inductions. But the force of the old induction *relative to the old evidence* is untouched. The evidence with which our experience has supplied us in the past may have proved misleading, but this is entirely irrelevant to the question of what conclusion we ought reasonably to have drawn from the evidence then before us. The validity and reasonable nature of inductive generalisation is, therefore, a question of logic and not of experience, of formal and not of material laws. The actual constitution of the phenomenal universe determines the character of our evidence ; but it cannot determine what conclusions *given* evidence *rationally* supports.

CHAPTER XIX

THE NATURE OF ARGUMENT BY ANALOGY

All kinds of reasoning from causes or effects are founded on two particulars, viz. the constant conjunction of any two objects in all past experience, and the resemblance of a present object to any of them. Without some degree of resemblance, as well as union, 'tis impossible there can be any reasoning.—HUME.[1]

1. HUME rightly maintains that some degree of resemblance must always exist between the various instances upon which a generalisation is based. For they must have this, at least, in common, that they are instances of the proposition which generalises them. Some element of analogy must, therefore, lie at the base of every inductive argument. In this chapter I shall try to explain with precision the meaning of Analogy, and to analyse the reasons, for which, rightly or wrongly, we usually regard analogies as strong or weak, without considering at present whether it is possible to find a *good* reason for our instinctive principle that likeness breeds the expectation of likeness.

2. There are a few technical terms to be defined. We mean by a *generalisation* a statement that all of a certain definable class of propositions are true. It is convenient to specify this class in the following way. If $f(x)$ is true for all those values of x for which $\phi(x)$ is true, then we have a generalisation about ϕ and f which we may write $g(\phi, f)$. If, for example, we are dealing with the generalisation, "All swans are white," this is equivalent to the statement, "'x is white' is true for all those values of x for which 'x is a swan' is true." The proposition $\phi(a).f(a)$ is an *instance* of the generalisation $g(\phi, f)$.

By thus defining a generalisation in terms of propositional functions, it becomes possible to deal with all kinds of generalisa-

[1] *A Treatise of Human Nature.*

tions in a uniform way; and also to bring generalisation into convenient connection with our definition of Analogy.

If some one thing is true about both of two objects, if, that is to say, they both satisfy the same propositional function, then to this extent there is an *analogy* between them. Every generalisation $g(\phi, f)$, therefore, asserts that one analogy is always accompanied by another, namely, that between all objects having the analogy ϕ there is also the analogy f. The set of propositional functions, which are satisfied by both of the two objects, constitute the *positive analogy*. The analogies, which would be disclosed by complete knowledge, may be termed the *total positive analogy*; those which are relative to partial knowledge, the *known positive analogy*.

As the positive analogy measures the resemblances, so the negative analogy measures the differences between the two objects. The set of functions, such that each is satisfied by one and not by the other of the objects, constitutes the *negative analogy*. We have, as before, the distinction between the *total negative analogy* and the *known negative analogy*.

This set of definitions is soon extended to the cases in which the number of instances exceeds two. The functions which are true of *all* of the instances constitute the positive analogy of the set of instances, and those which are true of *some only*, and are false of others, constitute the negative analogy. It is clear that a function, which represents positive analogy for a group of instances taken out of the set, may be a negative analogy for the set as a whole. Analogies of this kind, which are positive for a sub-class of the instances, but negative for the whole class, we may term *sub-analogies*. By this it is meant that there are resemblances which are common to some of the instances, but not to all.

A simple notation, in accordance with these definitions, will be useful. If there is a positive analogy ϕ between a set of instances $a_1 \ldots a_n$, whether or not this is the total analogy between them, let us write this—

$$\underset{a_1 \ldots a_n}{\mathrm{A}} (\phi).[1]$$

[1] Hence $\underset{a_1 \ldots a_n}{\mathrm{A}} (\phi) \equiv \phi(a_1) . \phi(a_2) \ldots \phi(a_n) \equiv \prod_{x=a_1}^{x=a_n} \phi(x)$.

And if there is a negative analogy ϕ', let us write this—

$$\underset{a_1 \ldots a_n}{\bar{A}} (\phi').$$[1]

Thus $\underset{a_1 \ldots a_n}{A} (\phi)$ expresses the fact that there is a set of characteristics ϕ which are common to all the instances, and $\underset{a_1 \ldots a_n}{\bar{A}} (\phi')$ that there is a set of characteristics ϕ' which is true of at least one of the instances and false of at least one.

3. In the typical argument from analogy we wish to generalise from one part to another of the total analogy which experience has shown to exist between certain selected instances. In all the cases where one characteristic ϕ has been found to exist, another characteristic f has been found to be associated with it. We argue from this that any instance, which is known to share the first analogy ϕ, is likely to share also the second analogy f. We have found in certain cases, that is to say, that both ϕ and f are true of them; and we wish to assert f as true of other cases in which we have only observed ϕ. We seek to establish the generalisation $g(\phi, f)$, on the ground that ϕ and f constitute between them an observed positive analogy in a given set of experiences.

But while the argument is of this character, the grounds, upon which we attribute more or less weight to it, are often rather complex; and we must discuss them, therefore, in a systematic manner.

4. According to the view suggested in the last chapter, the value of such an argument depends partly upon the nature of the conclusion which we seek to draw, partly upon the evidence which supports it. If Hume had expected the same degree of nourishment as well as the same taste and relish from all of the eggs, he would have drawn a conclusion of weaker probability. Let us consider, then, this dependence of the probability upon the *scope* of the generalisation $g(\phi, f)$,—upon the comprehensiveness, that is to say, of the condition ϕ and the conclusion f respectively.

The more comprehensive the condition ϕ and the less comprehensive the conclusion f, the greater *à priori* probability do we attribute to the generalisation g. With every increase in ϕ this probability increases, and with every increase in f it will diminish.

[1] Hence $\underset{a_1 \ldots a_n}{\bar{A}} (\phi') \equiv \sum_{x=a_r}^{x=a_s} \phi'(x) \,.\, \sum_{x=a_r'}^{x=a_s'} \overline{\phi'(x)}.$

The condition $\phi(\equiv \phi_1\phi_2)$ is more comprehensive than the condition ϕ_1, relative to the general evidence h, if ϕ_2 is a condition independent of ϕ_1 relative to h, ϕ_2 being independent of ϕ_1, if $g(\phi_1, \phi_2)/h \neq 1$, *i.e.* if, relative to h, the satisfaction of ϕ_2 is not inferrible from that of ϕ_1.

Similarly the conclusion $f(\equiv f_1f_2)$ is more comprehensive than the conclusion f_1, relative to the general evidence h, if f_2 is a conclusion independent of f_1, relative to h, *i.e.* if $g(f_1, f_2)/h \neq 1$.

If $\phi \equiv \phi_1\phi_2$ and $f \equiv f_1f_2$, where ϕ_1 and ϕ_2 are independent and f_1 and f_2 are independent relative to h, we have—

$$g(\phi_1, f)/h = g(\phi_1\phi_2, f) \,.\, g(\phi_1\bar{\phi}_2, f)/h$$
$$\ngtr g(\phi, f)/h,$$

and

$$g(\phi, f)/h = g(\phi, f_1f_2)/h$$
$$= g(\phi f_1, f_2)/h \,.\, g(\phi, f_1)/h$$
$$\leq g(\phi, f_1)/h,$$

so that

$$g(\phi, f_1)/h \geq g(\phi, f)/h \geq g(\phi_1, f)/h.$$

This proves the statement made above. It will be noticed that we cannot necessarily compare the *à priori* probabilities of two generalisations in respect of more and less, unless the condition of the first is included in the condition of the second, and the conclusion of the second is included in that of the first.

We see, therefore, that some generalisations stand *initially* in a stronger position than others. In order to attain a given degree of probability, generalisations require, according to their scope, different amounts of favourable evidence to support them.

5. Let us now pass from the character of the generalisation *à priori* to the evidence by which we support it. Since, whenever the conclusion f is complex, *i.e.* resolvable into the form f_1f_2 where $g(f_1, f_2)/h \neq 1$, we can express the probability of the generalisation $g(\phi, f)$ as the product of the probabilities of the two generalisations $g(\phi f_1, f_2)$ and $g(\phi, f_1)$, we may assume in what follows, that the conclusion f is simple and not capable of further analysis, without diminishing the generality of our argument.

We will begin with the simplest case, namely, that which arises in the following conditions. First, let us assume that our knowledge of the examined instances is complete, so that we know of every statement, which is about the examined instances, whether it is true or false of each.[1] Second, let us assume that

[1] If $\psi(a)$ is a proposition and $\psi(a) = h \,.\, \theta(a)$, where h is a proposition not involving a, then we must regard $\theta(a)$, not $\psi(a)$ as the statement *about a*.

all the instances which are known to satisfy the condition ϕ, are also known to satisfy the conclusion f of the generalisation. And third let us assume that there is nothing which is true of *all* the examined instances and yet not included either in ϕ or in f, *i.e.* that the positive analogy between the instances is exactly co-extensive with the analogy ϕf which is covered by the generalisation.

Such evidence as this constitutes what we may term a perfect analogy. The argument in favour of the generalisation cannot be further improved by a knowledge of additional instances. Since the positive analogy between the instances is exactly coextensive with the analogy covered by the generalisation, and since our knowledge of the examined instances is complete, there is no need to take account of the negative analogy.

An analogy of this kind, however, is not likely to have much practical utility; for if the analogy covered by the generalisation, covers the *whole* of the positive analogy between the instances it is difficult to see to what *other* instances the generalisation can be applicable. Any instance, about which everything is true which is true of all of a set of instances, must be identical with one of them. Indeed, an argument from perfect analogy can only have practical utility, if, as will be argued later on, there are some distinctions between instances which are *irrelevant* for the purposes of analogy, and if, in a perfect analogy, the positive analogy, of which we must take account, need cover only those distinctions which are relevant. In this case a generalisation based on perfect analogy might cover instances numerically distinct from those of the original set.

The law of the Uniformity of Nature appears to me to amount to an assertion that an analogy which is perfect, except that mere differences of position in time and space are treated as irrelevant, is a valid basis for a generalisation, two total causes being regarded as the *same* if they only differ in their positions in time or space. This, I think, is the whole of the importance which this law has for the theory of inductive argument. It involves the assertion of a generalised judgment of irrelevance, namely, of the irrelevance of mere position in time and space to generalisations which have no reference to particular positions in time and space. It is in respect of such position in time or space that 'nature' is supposed 'uniform.' The significance of the law and the nature of its justification, if any, are further discussed in Chapter XXII.

6. Let us now pass to the type which is next in order of simplicity. We will relax the first condition and no longer assume that the *whole* of the positive analogy between the instances is covered by the generalisation, though retaining the assumption that our knowledge of the examined instances is complete. We know, that is to say, that there are some respects in which the examined instances are all alike, and yet which are not covered by the generalisation. If ϕ_1 is the part of the positive analogy between the instances which is *not* covered by the generalisation, then the probability of this type of argument from analogy can be written—

$$g(\phi, f) \Big/ \underset{a_1 \ldots a_n}{\mathrm{A}} (\phi\phi_1 f).$$

The value of this probability turns on the comprehensiveness of ϕ_1. There are some characteristics ϕ_1 common to all the instances, which the generalisation treats as unessential, but the less comprehensive these are the better. ϕ_1 stands for the characteristics in which all the instances resemble one another outside those covered by the generalisation. To reduce these resemblances between the instances is the same thing as to increase the differences between them. And hence any increase in the Negative Analogy involves a reduction in the comprehensiveness of ϕ_1. When, however, our knowledge of the instances is complete, it is not necessary to make separate mention of the negative analogy $\underset{a_1 \ldots a_n}{\bar{\mathrm{A}}} (\phi')$ in the above formula. For ϕ' simply includes all those functions about the instances, which are not included in $\phi\phi_1 f$, and of which the contradictories are not included in them; so that in stating $\underset{a_1 \ldots a_n}{\mathrm{A}} (\phi\phi_1 f)$, we state by implication $\underset{a_1 \ldots a_n}{\bar{\mathrm{A}}} (\phi')$ also.

The whole process of strengthening the argument in favour of the generalisation $g(\phi, f)$ by the accumulation of further experience appears to me to consist in making the argument approximate as nearly as possible to the conditions of a perfect analogy, by steadily reducing the comprehensiveness of those resemblances ϕ_1 between the instances which our generalisation disregards. Thus the advantage of additional instances, derived from experience, arises not out of their number as such, but out of their tendency to limit and reduce the comprehensiveness of ϕ_1, or, in other words, out of their tendency to increase the negative analogy ϕ', since $\phi_1\phi'$ comprise between them whatever is not covered by ϕf. The more numerous the instances, the less comprehensive are their superfluous resemblances likely to be. But a single additional instance which greatly reduced ϕ_1 would increase the probability of the argument more than a large number of instances which affected ϕ_1 less.

7. The nature of the argument examined so far is, then, that the instances all have some characteristics in common which we have ignored in framing our generalisation; but it is still assumed that our knowledge about the examined instances is complete. We will next dispense with this latter assumption, and deal with the case in which our knowledge of the characteristics of the examined instances themselves is or may be incomplete.

It is now necessary to take explicit account of the known negative analogy. For when the known positive analogy falls short of the total positive analogy, it is not possible to infer the negative analogy from it. Differences may be known between the instances which cannot be inferred from the known positive analogy. The probability of the argument must, therefore, be written—

$$g(\phi, f) \Big/ \underset{a_1 \ldots a_n}{\mathrm{A}}(\phi\phi_1 f)\ \underset{a_1 \ldots a_n}{\bar{\mathrm{A}}}(\phi'),$$

where $\phi\phi_1 f$ stands for the characteristics in which all n instances $a_1 \ldots a_n$ are *known* to be alike, and ϕ' stands for the characteristics in which they are *known* to differ.

This argument is strengthened by any additional instance or by any additional knowledge about the former instances which diminishes the known superfluous resemblances ϕ_1 or increases the negative analogy ϕ'. The object of the accumulation of further experience is still the same as before, namely, to make the form of the argument approximate more and more closely to that of perfect analogy. Now, however, that our knowledge of the instances is no longer assumed to be complete, we must take account of the mere *number* n of the instances, as well as of our specific knowledge in regard to them; for the more numerous the instances are, the greater the opportunity for the *total* negative analogy to exceed the *known* negative analogy. But

the more complete our knowledge of the instances, the less attention need we pay to their mere number, and the more imperfect our knowledge the greater the stress which must be laid upon the argument from number. This part of the argument will be discussed in detail in the following chapter on Pure Induction.

8. When our knowledge of the instances is incomplete, there may exist analogies which are known to be true of some of the instances and are not known to be false of any. These sub-analogies (see § 2) are not so dangerous as the positive analogies ϕ_1, which are known to be true of *all* the instances, but their existence is, evidently, an element of weakness, which we must endeavour to eliminate by the growth of knowledge and the multiplication of instances. A sub-analogy of this kind between the instances $a_r \ldots a_s$ may be written $\underset{a_r \ldots a_s}{\mathrm{A}}(\psi_k)$; and the formula, if it is to take account of all the relevant information, ought, therefore, to be written—

$$g(\phi, f) \Big/ \underset{a_1 \ldots a_n}{\mathrm{A}}(\phi\phi_1 f)\ \underset{a_1 \ldots a_n}{\bar{\mathrm{A}}}(\phi')\Pi\left\{\underset{a_r \ldots a_s}{\mathrm{A}}(\psi_k)\right\},$$

where the terms of $\Pi\left\{\underset{a_r \ldots a_s}{\mathrm{A}}(\psi_k)\right\}$ stand for the various sub-analogies between sub-classes of the instances, which are not included in $\phi\phi_1 f$ or in ϕ'.

9. There is now another complexity to be introduced. We must dispense with the assumption that the whole of the analogy covered by the generalisation is known to exist in all the instances. For there may be some instances within our experience, about which our knowledge is incomplete, but which show *part* of the analogy required by the generalisation and nothing which contradicts it; and such instances afford some support to the generalisation. Suppose that ${}_b\phi$ and ${}_bf$ are *part* of ϕ and f respectively, then we may have a set of instances $b_1 \ldots b_m$ which show the following analogies:

$$\underset{b_1 \ldots b_m}{\mathrm{A}}({}_b\phi\,{}_b\phi_1\,{}_bf)\ \underset{b_1 \ldots b_m}{\bar{\mathrm{A}}}({}_b\phi')\Pi\left\{\underset{b_r \ldots b_s}{\mathrm{A}}({}_b\psi_k)\right\},$$

where ${}_b\phi_1$ is the analogy not covered by the generalisation, and so on, as before.

The formula, therefore, is now as follows:

$$g(\phi, f) \Big/ \prod_{a, b \ldots} \left\{ \underset{a_1 \ldots a_n}{\mathrm{A}} ({}_a\phi_a\phi_{1a}f) \underset{a_1 \ldots a_n}{\bar{\mathrm{A}}} ({}_a\phi') \right\} \Pi \left\{ \underset{a_r b_s \ldots}{\mathrm{A}} (\psi_k) \right\}.$$

In this expression ${}_a\phi$, ${}_af$ are the whole *or* part of ϕ, f; the product $\prod\limits_{a, b \ldots}$ is composed of the positive and negative analogies for each of the sets of instances $a_1 \ldots a_n$, $b_1 \ldots b_m$, etc.; and the product Π contains the various sub-analogies of different sub-classes of all the instances $a_1 \ldots a_n$, $b_1 \ldots b_m$, etc., regarded as *one* set.[1]

10. This completes our classification of the *positive* evidence which supports a generalisation; but the probability may also be affected by a consideration of the negative evidence. We have taken account so far of that part of the evidence only which shows the whole or part of the analogy we require, and we have neglected those instances of which ϕ, the condition of the generalisation, or f, its conclusion, or part of ϕ or of f is *known to be false*. Suppose that there are instances of which ϕ is true and f false, it is clear that the generalisation is ruined. But cases in which we know *part* of ϕ to be true and f to be false, and are ignorant as to the truth or falsity of the rest of ϕ, weaken it to some extent. We must take account, therefore, of analogies

$$\underset{a_1' \ldots a_n'}{\mathrm{A}} ({}_{a'}\phi_{a'}\bar{f}),$$

where ${}_{a'}\phi$, part of ϕ, is true of all the set, and ${}_{a'}f$, part of f, is false of all the set, while the truth or falsity of some part of ϕ and f is unknown. The negative evidence, however, can strengthen as well as weaken the evidence. We deem instances favourably relevant in which ϕ and f are both false together.[2]

Our final formula, therefore, must include terms, similar to those in the formula which concludes § 9, not only for sets of instances which show analogies ${}_a\phi_af$, where ${}_a\phi$ and ${}_af$ are parts of ϕ and f, but also for sets which show analogies ${}_a\bar{\phi}_af$, or analogies ${}_a\bar{\phi}_a\bar{f}$, where ${}_a\phi$ and ${}_af$ are the whole or part of ϕ and f, and $\bar{\phi}\,\bar{f}$ are the contradictories of ϕ and f.[1]

[1] Even if we want to distinguish between the sub-analogies of the a set and the sub-analogies of the b set, this information can be gathered from the product Π.

[2] I am disposed to think that we need not pay attention to instances for which part of ϕ is known to be false, and part of f to be true. But the question is a little perplexing.

It should be added, perhaps, that the theoretical classification of most empirical arguments in daily use is complicated by the account which we reasonably take of generalisations previously established. We often take account indirectly, therefore, of evidence which supports in some degree other generalisations than that which we are concerned to establish or refute at the moment, but the probability of which is relevant to the problem under investigation.

11. The argument will be rendered unnecessarily complex, without much benefit to its theoretical interest, if we deal with the most general case of all. What follows, therefore, will deal with the formula of the third degree of generality, namely—

$$g(\phi, f) \Big/ \underset{a_1 \ldots a_n}{\mathrm{A}} (\phi\phi_1 f) \underset{a_1 \ldots a_n}{\bar{\mathrm{A}}} (\phi') \Pi \left\{ \underset{a_r \ldots a_s}{\mathrm{A}} (\psi_k) \right\},$$

in which no *partial* instances occur, *i.e.* no instances in which part only of the analogy, required by the generalisation, is known to exist. In this third degree of generality, it will be remembered, our knowledge of the characteristics of the instances is incomplete, there is more analogy between the instances than is covered by the generalisation, and there are some sub-analogies to be reckoned with. In the above formula the incompleteness of our knowledge is implicitly recognised in that $\phi\phi_1 f\phi'$ are not between them entirely comprehensive. It is also supposed that all the evidence we have is positive, no knowledge is assumed, that is to say, of instances characterised by the conjunctions ${}_a\bar{\phi}_af$, ${}_a\phi_a\bar{f}$, or ${}_a\bar{\phi}_a\bar{f}$, where ${}_a\phi$ and ${}_af$ are part of ϕ and f.

An argument, therefore, from experience, in which, on the basis of examined instances, we establish a generalisation applicable beyond these instances, can be strengthened, if we restrict our attention to the simpler type of case, by the following means:

(1) By reducing the resemblances ϕ_1 known to be common to all the instances, but ignored as unessential by the generalisation.

(2) By increasing the differences ϕ' known to exist between the instances.

[1] Where the conclusion f is simple and not complex (see § 5), some of these complications cannot, of course, arise.

(3) By diminishing the sub-analogies or unessential resemblances ψ_k known to be common to some of the instances and not known to be false of any.

These results can generally be obtained in two ways, either by increasing the number of our instances or by increasing our knowledge of those we have.

The reasons why these methods seem to common sense to strengthen the argument are fairly obvious. The object of (1) is to avoid the possibility that ϕ_1 as well as ϕ is a necessary condition of f. The object of (2) is to avoid the possibility that there may be some resemblances additional to ϕ, common to all the instances, which have escaped our notice. The object of (3) is to get rid of indications that the total value of ϕ_1 may be greater than the known value. When $\phi\phi_1 f$ is the *total* positive analogy between the instances, so that the known value of ϕ_1 is its total value, it is (1) which is fundamental; and we need take account of (2) and (3) only when our knowledge of the instances is incomplete. But when our knowledge of the instances is incomplete, so that ϕ_1 falls short of its total value and we cannot infer ϕ' from it, it is better to regard (2) as fundamental; in any case every reduction of ϕ_1 must increase ϕ'.

12. I have now attempted to analyse the various ways in which common practice seems to assume that considerations of Analogy can yield us presumptive evidence in favour of a generalisation.

It has been my object, in making a classification of empirical arguments, not so much to put my results in forms closely similar to those in which problems of generalisation commonly present themselves to scientific investigators, as to inquire whether ultimate uniformities of method can be found beneath the innumerable modes, superficially differing from another, in which we do in fact argue.

I have not yet attempted to justify this way of arguing. After turning aside to discuss in more detail the method of Pure Induction, I shall make this attempt; or rather I shall try to see *what sort* of assumptions are capable of justifying empirical reasoning of this kind.

CHAPTER XX

THE VALUE OF MULTIPLICATION OF INSTANCES, OR PURE INDUCTION

1. It has often been thought that the essence of inductive argument lies in the multiplication of instances. "Where is that process of reasoning," Hume inquired, "which from one instance draws a conclusion, so different from that which it infers from a hundred instances, that are no way different from that single instance?" I repeat that by emphasising the number of the instances Hume obscured the real object of the method. If it were strictly true that the hundred instances are *no* way different from the single instance, Hume would be right to wonder in what manner they can strengthen the argument. The object of increasing the number of instances arises out of the fact that we are nearly always aware of *some* difference between the instances, and that even where the known difference is insignificant we may suspect, especially when our knowledge of the instances is very incomplete, that there may be more. Every new instance *may* diminish the unessential resemblances between the instances and by introducing a new difference increase the Negative Analogy. For this reason, and for this reason only, new instances are valuable.

If our premisses comprise the body of memory and tradition which has been originally derived from direct experience, and the conclusion which we seek to establish is the Newtonian theory of the Solar System, our argument is one of Pure Induction, in so far as we support the Newtonian theory by pointing to the great number of consequences which it has in common with the facts of experience. The predictions of the Nautical Almanack are a consequence of the Newtonian theory, and these predictions are verified many thousand times a day. But even here the force of the argument largely depends, not on the mere number of these predictions, but on the knowledge that the circumstances in which they are fulfilled differ widely from one another in a vast number of important respects. The *variety* of the circumstances, in which the Newtonian generalisation is fulfilled, rather than the number of them, is what seems to impress our reasonable faculties.

2. I hold, then, that our object is always to increase the Negative Analogy, or, which is the same thing, to diminish the characteristics common to all the examined instances and yet not taken account of by our generalisation. Our method, however, may be one which certainly achieves this object, or it may be one which possibly achieves it. The former of these, which is obviously the more satisfactory, may consist either in increasing our definite knowledge respecting instances examined already, or in finding additional instances respecting which definite knowledge is obtainable. The second of them consists in finding additional instances of the generalisation, about which, however, our definite knowledge may be meagre; such further instances, if our knowledge about them were more complete, would either increase or leave unchanged the Negative Analogy; in the former case they would strengthen the argument and in the latter case they would not weaken it; and they must, therefore, be allowed some weight. The two methods are not entirely distinct, because new instances, about which we have some knowledge but not much, may be known to increase the Negative Analogy a little by the first method, and suspected of increasing it further by the second.

It is characteristic of advanced scientific method to depend on the former, and of the crude unregulated induction of ordinary experience to depend on the latter. It is when our definite knowledge about the instances is limited, that we must pay attention to their number rather than to the specific differences between them, and must fall back on what I term Pure Induction.

In this chapter I investigate the conditions and the manner in which the mere repetition of instances can add to the force of the argument. The chief value of the chapter, in my judgment, is negative, and consists in showing that a line of advance, which might have seemed promising, turns out to be a blind alley, and that we are thrown back on known Analogy. Pure

Induction will not give us any very substantial assistance in getting to the bottom of the general inductive problem.

3. The problem of generalisation[1] by Pure Induction can be stated in the following symbolic form:

Let h represent the general *à priori data* of the investigation; let g represent the generalisation which we seek to establish; let $x_1x_2\ldots x_n$ represent instances of g.

Then $x_1/gh=1$, $x_2/gh=1\ldots x_n/gh=1$; given g, that is to say, the truth of each of its instances follows. The problem is to determine the probability $g/hx_1x_2\ldots x_n$, *i.e.* the probability of the generalisation when n instances of it are given. Our analysis will be simplified, and nothing of fundamental importance will be lost, if we introduce the assumption that there is nothing in our *à priori data* which leads us to distinguish between the *à priori* likelihood of the different instances; we assume, that is to say, that there is no reason *à priori* for expecting the occurrence of any one instance with greater reliance than any other, *i.e.*

$$x_1/h=x_2/h=\ldots=x_n/h.$$

Write
$$g/hx_1x_2\ldots x_n=p_n$$
and
$$x_{n+1}/hx_1x_2\ldots x_n=y_{n+1};$$
then
$$\frac{p_n}{p_{n-1}}=\frac{g/hx_1\ldots x_n}{g/hx_1\ldots x_{n-1}}=\frac{gx_n/hx_1\ldots x_{n-1}}{g/hx_1\ldots x_{n-1}\cdot x_n/hx_1\ldots x_{n-1}}$$
$$=\frac{x_n/ghx_1\ldots x_{n-1}}{x_n/hx_1\ldots x_{n-1}}$$
$$=\frac{1}{y_n}.$$

$\therefore \dfrac{p_n}{p_{n-1}}=\dfrac{1}{y_n}$, and hence $p_n=\dfrac{1}{y_1y_2\ldots y_n}\cdot p_0$, where $p_0=g/h$, *i.e.* p_0 is the *à priori* probability of the generalisation.

[1] In the most general sense we can regard any proposition as the generalisation of all the propositions which follow from it. For if h is any proposition, and we put $\phi(x)\equiv$ 'x can be inferred from h' and $f(x)\equiv x$, then $g(\phi, f)\equiv h$. Since Pure Induction consists in finding as many instances of a generalisation as possible, it is, in the widest sense, the process of strengthening the probability of any proposition by adducing numerous instances of known truths which follow from it. The argument is one of Pure Induction, therefore, in so far as the probability of a conclusion is based upon the number of independent consequences which the conclusion and the premisses have in common.

It follows, therefore, that $p_n > p_{n-1}$ so long as $y_n > 1$. Further,

$$x_1x_2\ldots x_n/h = x_n/hx_1x_2\ldots x_{n-1}\,.\,x_1x_2\ldots x_{n-1}/h$$
$$= y_n\,.\,x_1x_2\ldots x_{n-1}/h$$
$$= y_ny_{n-1}\ldots y_1.$$

$$\therefore\ p_n = \frac{p_0}{y_1y_2\ldots y_n} = \frac{p_0}{x_1x_2\ldots x_n/h}$$
$$= \frac{p_0}{x_1x_2\ldots x_ng/h + x_1x_2\ldots x_n\bar{g}/h}$$
$$= \frac{p_0}{g/h + x_1x_2\ldots x_n/\bar{g}h\,.\,\bar{g}/h}$$
$$= \frac{p_0}{p_0 + x_1x_2\ldots x_n/\bar{g}h(1-p_0)}.$$

This approaches unity as a limit, if $x_1x_2\ldots x_n/\bar{g}h\,.\,\frac{1}{p_0}$ approaches zero as a limit, when n increases.

4. We may now stop to consider how much this argument has proved. We have shown that if each of the instances necessarily follows from the generalisation, then each additional instance increases the probability of the generalisation, so long as the new instance could not have been predicted with certainty from a knowledge of the former instances.[1] This condition is the same as that which came to light when we were discussing Analogy. If the new instance were identical with one of the former instances, a knowledge of the latter would enable us to predict it. If it differs or may differ in analogy, then the condition required above is satisfied.

The common notion, that each successive verification of a doubtful principle strengthens it, is formally proved, therefore, without any appeal to conceptions of law or of causality. *But we have not proved* that this probability approaches certainty as a limit, or even that our conclusion becomes more likely than not, as the number of verifications or instances is indefinitely increased.

5. What are the conditions which must be satisfied in order that the rate, at which the probability of the generalisation increases, may be such that it will approach certainty as a

[1] Since $p_n > p_{n-1}$ so long as $y_n \neq 1$.

limit when the number of independent instances of it are indefinitely increased? We have already shown, as a basis for this investigation, that p_n approaches the limit of certainty for a generalisation g, if, as n increases, $x_1x_2\ldots x_n/\bar{g}h$ becomes small compared with p_0, *i.e.* if the *à priori* probability of so many instances, assuming the falsehood of the generalisation, is small compared with the generalisation's *à priori* probability. It follows, therefore, that the probability of an induction tends towards certainty as a limit, when the number of instances is increased, provided that

$$x_r/x_1x_2\ldots x_{r-1}\bar{g}h < 1-\epsilon$$

for all values of r, and $p_0 > \eta$, where ϵ and η are finite probabilities, separated, that is to say, from impossibility by a value of some finite amount, however small. These conditions appear simple, but the meaning of a 'finite probability' requires a word of explanation.[1]

I argued in Chapter III. that not all probabilities have an exact numerical value, and that, in the case of some, one can say no more about their relation to certainty and impossibility than that they fall short of the former and exceed the latter. There is one class of probabilities, however, which I called the numerical class, the ratio of each of whose members to certainty can be expressed by some number less than unity; and we can sometimes compare a non-numerical probability in respect of more and less with one of these numerical probabilities. This enables us to give a definition of 'finite probability' which is capable of application to non-numerical as well as to numerical probabilities. I define a 'finite probability' as one which *exceeds* some numerical probability, the ratio of which to certainty can be expressed by a finite number.[2] The principal method, in which a probability can be proved finite by a process of argument, arises either when

[1] The proof of these conditions, which is obvious, is as follows:

$$x_1x_2\ldots x_n/\bar{g}h = x_n/x_1x_2\ldots x_{n-1}\bar{g}h\,.\,x_1x_2\ldots x_{n-1}/\bar{g}h < (1-\epsilon)^n,$$

where ϵ is finite and $p_0 > \eta$ where η is finite. There is always, under these conditions, some finite value of n such that both $(1-\epsilon)^n$ and $\frac{(1-\epsilon)^n}{\eta}$ are less than any given finite quantity, however small.

[2] Hence a series of probabilities $p_1p_2\ldots p_r$ approaches a limit L, if, given any positive finite number ϵ however small, a positive integer n can always be found such that for all values of r greater than n the difference between L and p_r is less than $\epsilon.\gamma$, where γ is the measure of certainty.

its conclusion can be shown to be one of a finite number of alternatives, which are between them exhaustive or, at any rate, have a finite probability, and to which the Principle of Indifference is applicable; or (more usually), when its conclusion is *more* probable than some hypothesis which satisfies this first condition.

6. The conditions, which we have now established in order that the probability of a pure induction may tend towards certainty as the number of instances is increased, are (1) that $x_r/x_1x_2 \ldots x_{r-1}\bar{g}h$ falls short of certainty by a finite amount for all values of r, and (2) that p_0, the *à priori* probability of our generalisation, exceeds impossibility by a finite amount. It is easy to see that we can show by an exactly similar argument that the following more general conditions are equally satisfactory:

(1) That $x_r/x_1x_2 \ldots x_{r-1}\bar{g}h$ falls short of certainty by a finite amount for all values of r beyond a specified value s.

(2) That p_s, the probability of the generalisation relative to a knowledge of these first s instances, exceeds impossibility by a finite amount.

In other words Pure Induction can be usefully employed to strengthen an argument if, after a certain number of instances have been examined, we have, from some other source, a finite probability in favour of the generalisation, and, assuming the generalisation is false, a finite uncertainty as to its conclusion being satisfied by the next hitherto unexamined instance which satisfies its premiss. To take an example, Pure Induction can be used to support the generalisation that the sun will rise every morning for the next million years, provided that with the experience we have actually had there are finite probabilities, however small, *derived from some other source*, first, in favour of the generalisation, and, second, in favour of the sun's *not* rising to-morrow assuming the generalisation to be false. Given these finite probabilities, obtained otherwise, however small, then the probability can be strengthened and can tend to increase towards certainty by the mere multiplication of instances provided that these instances are so far distinct that they are not inferrible one from another.

7. Those supposed proofs of the Inductive Principle, which are based openly or implicitly on an argument in inverse probability, are all vitiated by unjustifiable assumptions relating to the magnitude of the *à priori* probability p_0. Jevons, for

instance, avowedly assumes that we may, in the absence of special information, suppose any unexamined hypothesis to be as likely as not. It is difficult to see how such a belief, if even its most immediate implications had been properly apprehended, could have remained plausible to a mind of so sound a practical judgment as his. The arguments against it and the contradictions to which it leads have been dealt with in Chapter IV. The demonstration of Laplace, which depends upon the Rule of Succession, will be discussed in Chapter XXX.

8. The prior probability, which must always be found, before the method of pure induction can be usefully employed to support a substantial argument, is derived, I think, in most ordinary cases—with what justification it remains to discuss—from considerations of Analogy. But the conditions of valid induction as they have been enunciated above, are quite independent of analogy, and might be applicable to other types of argument. In certain cases we might feel justified in assuming *directly* that the necessary conditions are satisfied.

Our belief, for instance, in the validity of a logical scheme is based partly upon inductive grounds—on the *number* of conclusions, each seemingly true on its own account, which can be derived from the axioms—and partly on a degree of self-evidence in the axioms themselves sufficient to give them the initial probability upon which induction can build. We depend upon the initial presumption that, if a proposition appears to us to be true, this is by itself, in the absence of opposing evidence, *some reason* for its *being* as well as appearing true. We cannot deny that what appears true is sometimes false, but, unless we can assume some substantial relation of probability between the appearance and the reality of truth, the possibility of even probable knowledge is at an end.

The conception of our having *some* reason, though not a conclusive one, for certain beliefs, arising out of direct inspection, may prove important to the theory of epistemology. The old metaphysics has been greatly hindered by reason of its having always demanded demonstrative certainty. Much of the cogency of Hume's criticism arises out of the assumption of methods of certainty on the part of those systems against which it was directed. The earlier realists were hampered by their not perceiving that lesser claims in the beginning might yield them what they wanted in the end. And transcendental philosophy has partly arisen, I believe, through the belief that there is no knowledge on these matters short of certain knowledge, being combined with the belief that such certain knowledge of metaphysical questions is beyond the power of ordinary methods.

When we allow that probable knowledge is, nevertheless, real, a new method of argument can be introduced into metaphysical discussions. The demonstrative method can be laid on one side, and we may attempt to advance the argument by taking account of circumstances which seem to give *some* reason for preferring one alternative to another. Great progress may follow if the nature and reality of objects of perception,[1] for instance, can be usefully investigated by methods not altogether dissimilar from those employed in science and with the prospect of obtaining as high a degree of certainty as that which belongs to some scientific conclusions; and it may conceivably be shown that a belief in the conclusions of science, enunciated in any reasonable manner however restricted, involves a preference for some metaphysical conclusions over others.

9. Apart from analysis, careful reflection would hardly lead us to expect that a conclusion which is based on no other than grounds of pure induction, defined as I have defined them as consisting of repetition of instances merely, could attain in this way to a high degree of probability. To this extent we ought all of us to agree with Hume. We have found that the suggestions of common sense are supported by more precise methods. Moreover, we constantly distinguish between arguments, which we call inductive, upon other grounds than the number of instances upon which they are based; and under certain conditions we regard as crucial an insignificant number of experiments. The method of pure induction may be a useful means of strengthening a probability based on some other ground. In the case, however, of most scientific arguments, which would commonly be called inductive, the probability that we are right, when we make predictions on the basis of past experience, depends not so much on the number of past experiences upon which we rely, as on the degree in which the circumstances of these experiences

[1] A paper by Mr. G. E. Moore entitled, "The Nature and Reality of Objects of Perception," which was published in the *Proceedings of the Aristotelian Society for 1906*, seems to me to apply for the first time a method somewhat resembling that which is described above.

resemble the known circumstances in which the prediction is to take effect. Scientific method, indeed, is mainly devoted to discovering means of so heightening the known analogy that we may dispense as far as possible with the methods of pure induction.

When, therefore, our previous knowledge is considerable and the analogy is good, the purely inductive part of the argument may take a very subsidiary place. But when our knowledge of the instances is slight, we may have to depend upon pure induction a good deal. In an advanced science it is a last resort,—the least satisfactory of the methods. But sometimes it must be our first resort, the method upon which we must depend in the dawn of knowledge and in fundamental inquiries where we must presuppose nothing.