ADMISSIBLE HYPOTHESES AND ENHANCED LEARNING

Mr. G. M. K. Hunt

Department of Philosophy, University of Warwick, Coventry CV4 7AL, England

ABSTRACT

This paper discusses strategies for moving through sequences of hypotheses, each one of which is produced in response to an experimental test of the previous member. Previous discussions of this issue have all agreed that hypotheses deductively incompatible with the evidence at stage ri cannot appear in the sequence beyond n This paper contends that this conclusion is untenable. The use of oversimplified models has led investigators into overlooking epistemological properties of more complex hypotheses which allow more sophisticated methodologies in testing and hypothesis generation. In particular, it is shown that testing already falsified hypotheses may give more experimental information than other, more traditional strategies. This is shown by considering a popular board game, but a realistic example is introduced to demonstrate the general importance and usefulness of the strategy.

In this short paper I discuss strategies in the testing of a sequence of hypotheses. Each of the hypotheses, except perhaps the first, is proposed after an experimental test of the preceding hypothesis. The aim of the strategy is to maximise the rate of convergence of members of the sequence of hypotheses to the correct hypothesis.

This problem has been discussed in the literature quite extensively under the general heading of Methodology of Science and to a lesser extent in the field of heuristics. Two fundamentally different views have been taken about the manner of proceeding in such a sequence. The first, associated with Carnap (Carnap, 1952), Reichenbach (Reichenbach, 1961) and many others following them is to put an evaluation on each possible n^{tn} hypothesis and choose that hypothesis which has a maximum value for this evaluation. Most often, but by no means always, the evaluations were simply the probabilities or confirmation of the alternative hypothesis, although this is by no means universal (e.g. Reichenbach, op cit). Such proposals are often termed inductivist.

A quite different view is taken by Popper (Popper, 1972) and followers who eschew a probabilistic or confirmatory evaluation function. Instead they propose that the function rises with the risk of falsification of a hypothesis. By adopting as the n^{tn} member of a sequence the hypothesis with the minumum risk of refutation, usually identified by maximising content of the hypothesis, we maximise the chance of refutation. This, in turn, maximises the expected rate of progress along the sequence of hypotheses and their convergence to the truth. The general stance of this latter group is that since there exist no objective measures of the confirmation of hypothesis by evidence, the appropriate evaluation is given by a methodologically derived function which reflects the a priori likelihood of refutation and hence progress along the sequence of hypotheses. Such a thesis concerning the appropriateness of the evaluation is called falsificationist.

Both approaches to the evaluation function share one important common feature. The evaluation takes a minimum value for any hypothesis which is deductively incompatible with the evidence to date. This assignment is justified in both approaches by the view that each proposed hypothesis must be a <u>possible</u> candidate for "the true hypothesis". The inductivist would restate this with term "probable" rather than possible but since "probable" entails "possible" the two views coincide.

The discussion of these topics has been rendered less helpful to investigators by the perhaps oversimple characterisation of hypothesis. They are, for much of the discussion, simply structured sets of sentences whose only relevant characteristic here is that they entail simple observation sentences which either do or do not accord with actual observations. The assessment of any hypothesis given a single piece of evidence is thus a two valued function - either inconsistent or consistent. While such a model of experimental testing has the virtue of simplicity it is, I contend, so unrealistic as to obscure the real and interesting problem of strategies for experimental testing even in simple contexts.

I will argue that when the two major schools of thought agree - that refuted hypotheses should be discarded - they are both wrong. In doing so I shall use hypotheses whose observational consequences are minimally more structured. These hypotheses will assign occupation states to a finite ordered set of cells. Accordingly the possibility arises of the hypotheses fitting the experimental data to a lesser or greater degree rather than a simple yes/no.

I have chosen such a form for the hypotheses because firstly they appear to represent the next simplest structure beyond unstructured hypothesis and secondly it is possible to find realistic and interesting examples of the use in scientific practice.

The most perspicuous way of proceeding is to produce a model with the characteristics of which I wish to discuss, then go on to outline a realistic example.

The example I will use is the game Mastermind. This game is for two players, a code maker and a code breaker. The code maker chooses four coloured counters from a. supply of six different colours, colour repetitions are allowed. The code of four colours is thought of as being ordered.

The code breaker conjectures a hypothesis as to the code and displays it on the board. The code maker then scores the hypothesis by displaying a black marker for each counter the code breaker has of the correct colour in the correct position, and a white marker for each counter, not yet scored, which is of a correct colour but in an incorrect position, by comparison with those code counters from which a black marker has not been awarded. The score is some measure of the nearness of the hypothesis to the truth.

For example, if the hidden code is Red, Red, Green, Green, a hypothesis Red, Blue, Green, Red would attract a black marker for the first Red counter, no marker for the Blue, a black marker for the Green and a white marker for the last Red. The scoring markers are not ordered, that is, one cannot deduce which of the counters earns which marks.

In the light of that score the code breaker conjectures a new code which is then scored, the process continuing until the code breaker scores four black markers - he has the correct hypothesis.

If we are to show that the code breaker may do better to hypothesise falsely it is necessary to show (1) after the first hypothesis some hypotheses can be known (deductively) to be false and (2) that a measure of better and worse guesses is available. In what follows, I shall assume, for simplicity, that the code maker is equally likely to choose any coloured marker in any position. The general situation is not changed by such a restriction, unless the code breaker has knowledge about the probability distribution.

To show (1) we need first to observe that there are 6^4 - 1296 possible codes, generated by an independent choice of one of six colours in each of the four positions. After choosing the first hypothesis the code breaker either gets no scoring markers or a combination of black and white markers. Unless the code breaker obtains four black markers, he knows that, at least, that hypothesis is false. If the code breaker does, improbably, get four black markers then he knows that all other possible codes are false (and the game is ended). Either way he knows that at least one hypothesis is false and thus (1) is satisfied. In fact an average of 1180 hypotheses are eliminated as the first hypothesis is scored.

We now turn to the second lemma: to show that some guesses as to the code are better than others and hence to show that some falsified hypotheses are better eliminators than some unfalsified hypotheses.

We first observe that in guessing a code and having it scored, the code breaker eliminates not only that guess (unless four black markers are scored) but, in general, many others as well. For example, if four red counters was guessed and no black markers obtained then it is certain that the hidden code does not contain any red counters in any position. A similar conclusion would follow if the guess contained only one red counter and no black or white counters were obtained. There are 671 codes containing at least one red counter, so if this is the code breaker's first guess all of these will have been eliminated. If this guess was not the first then some of the 671 will have already been eliminated. As most games last only five or so moves, in the course of which 1295 codes are eliminated there will be very few, if any, guesses which eliminate no other code but themselves.

I wish to suggest that the number of possibilities eliminated by a guess is a good measure of how valuable that guess has been in forwarding the aim of the code breaker. Since in this paper I will be concerned with the eliminative power of second guesses - there are no impossible codes for a first guess - I will use as a measure of eliminative power the proportion of remaining possible codes eliminated by that second guess. Because comparisons will only be made between second guesses which share a common first guess no metrical assumptions are made which vitiate the conclusions. The exception to this is where comparisons are made between the average eliminative powers of different ranges of guesses. These latter results are less important and will be discussed separately.

When playing a full game consisting of a sequence of guesses, it might well be that two successive guesses of lower eliminative power succeed jointly in eliminating more possible codes than an alternative pair of guesses each of higher eliminative power. This possibility cannot be ruled out a priori. This paper will restrict itself to the stepwise maximisation of eliminative power.

It now remains to demonstrate that a best guess, as measured by eliminative power, may be - and often is - a known false guess. Of course this can only happen on a second or subsequent guess as nothing is known as possible or impossible until the first guess is scored.

The method is to nominate a first guess and a score, e.g. Red, Red, Blue, Blue, score one Black marker and one White. It is now possible to deduce which codes the code maker may have chosen. They are simply these codes which would give that score against that guess. In this example there are 208 possibilities. Normally the next guess the code breaker will try will be one of these 208 codes.

For each of these 208 possible guesses and for each of the 208 possible codes (43264 pairs) it is possible to calculate the number of codes eliminated by the second guess. If it is assumed that all 208 codes are equally likely to have been chosen by the code maker, then it is possible to calculate the average eliminative power of each of the code breaker's possible second guesses. A code breaker might use such a measure of rejection power in choosing a second guess so as to maximise the expected rejection of wrong codes. It is possible to further calculate the average eliminative power of the 208 possible guesses the code breaker might have chosen. See Table 1 for results of Mastermind using only black markers. These results are more concisely displayed than the full game, but exhibit results typical of the full game.

The fact that the scoring procedure is not a simple refutation is essential for the arguments in this paper. For the black markers show that some counters are of the correct colour and in the correct position, while the white markers show that some remaining counters are of the correct colour but in incorrect positions. In a roundabout way the markers give a measure of distance from the truth. Indeed, two black and two white markers show that a simple interchange of two counters will give the correct code. Alternative, non-Mastermind, systems of scoring (e.g. use black marker only) give different measures and will be discussed later. They do, however, generally exhibit the interesting property which is the subject of this paper. This is because these scoring systems begin to take account in the experimental situation, of the structure of the hypothesis. My feeling is that after the simple yes/no given by an experiment to an unstructured deduction from a hypothesis, the next more complex observation statment would be a simple ordering of elements like the Mastermind code.

In Table 1 is listed the mean and maximum eliminative powers of both possible and impossible (i.e. already refuted) second guesses. Further work has shown that the striking results here appear if the scoring system of Mastermind is varied by using only black scoring markers or only white scoring markers, in the manner outlined in the game description, above.

In particular, modelling of an experiment designed to elucidate particular DNA sequences by the absorption, in solution, of matching test and subject DNA strings yields this same property: that use of already refuted hypotheses for further test results in more rapid convergence of hypotheses, by increasing the informational transfer at each experiment.

Of course ultimately, both in Mastermind and in molecular biology, one must eventually refer to unfalsified hypotheses for final confirmation.

REFERENCES:

1. Carnap, R, Continuum of Inductive Methods. Chicago Univ. Press, for example. (1952)

2. Popper, K. R., Objective Knowledge. Oxford: Clarendon Press, for example. (1972)

3. Reichenbach, H., Experience and Prediction. Chicago Univ. Press, for example.)1961)

TABLE 1 : Simplified Mastermind with black markers only

		Number of black markers obtained for first try			
		0	1	2	3
Mean eliminative power of second tries, any first try.	Possible Tries	0.640	0.617	0.627	0.395
	Imposs. Tries	0.553	0.590	0.553	0.442
Eliminative power of best try, any first try.	Possible Tries	0.640	0.617	0.627	0.395
	Imposs. Tries	0.581	0.637	0.637	0.555

For explanation see text.