

A procedure and a program for the trial-by-trial identification of hypotheses in concept learning

ALAN A. HARTLEY

Florida International University, Tamiami Trail, Miami, Florida 33174

An algorithm is described which identifies the subject's hypothesis on each trial of a concept learning task. The algorithm operates on classifications of the entire stimulus population. A computer program has been prepared to execute the algorithm.

The task of formulating an adequate model of concept learning would be much easier if the theorist had more direct access to the solution process than that afforded by manipulating variables and observing their effects on gross measures of performance such as trials or total errors to solution. Substantial advances in the understanding of attribute identification—the stage of concept learning in which the subject is theoretically assumed to be searching for the relevant dimensions—have been accompanied by the development of techniques which provide more direct avenues to the processing of information (cf. Falmagne, 1970; Levine, 1969; Nahinsky, 1970; Williams, 1971). Unfortunately, many of the techniques have been devised specifically for use with an attribute identification paradigm. In most cases, there is no readily available generalization to complete learning. A complete learning task is one in which the subject is told neither the dimensions that are relevant to the classification of stimuli nor the logical rule which maps combinations of values on the relevant dimensions onto the response categories.

TECHNIQUES FOR TAPPING INFORMATION PROCESSING

One common procedure is to require a verbal guess at the correct solution to the problem on each trial. This procedure is not particularly well suited to complete concept learning. It has been used primarily in attribute identification and similar tasks in which the *form* of the solution is carefully explained to the subject. So the subject knows that his tentative hypotheses should be constructed to agree with that form. The instructions in complete learning cannot standardize the subject's statements so easily. Moreover, it is unlikely that the unsophisticated subject can verbalize the complex logical connectives with sufficient precision to guarantee accurate communication between the subject and the experimenter.

I am grateful to William Batchelder for his advice and suggestions. Support during initial development of the program was provided by a NIMH predoctoral fellowship (MH54713-01). Please direct all correspondence to Department of Psychology, Florida International University, Miami, Florida 33174.

A method that has been successfully employed in the attribute identification of presence and conjunctive concepts is the insertion of blank trials into the stimulus sequence (e.g., Frankel, Levine, & Karpf, 1970; Nahinsky, 1970; Nahinsky & McGlynn, 1968; Nahinsky, Penrod, & Slaymaker, 1970; Nahinsky & Slaymaker, 1969). On a blank trial, a stimulus is presented and the subject offers a classification but no feedback is given. Blocks of blank trials are constructed so that the pattern of responses uniquely identifies the hypothesis or hypotheses on which the subject must be basing his responses. If the number of possible hypotheses is large or if the subject is not told the rule, the blank trials method is not feasible. The number of stimuli required to unambiguously identify an hypothesis would be quite large.

A very different method has been employed by Falmagne (1969, 1970, 1972). On each trial in a concept identification task with a presence solution, the subject was required to rate the likelihood that each possible solution was, in fact, correct. While this procedure allowed Falmagne to draw several powerful and remarkable inferences about the solution of presence problems, it is, in general, unworkable in the complete learning paradigm. Even if the subject were instructed that only two dimensions would be relevant, there would still be 270 possible solutions to a problem using stimuli with three dimensions each having three values.¹ This includes only rules constructed by mapping truth table categories onto the response classes. (See Bourne, 1970, for a discussion of the relation between conceptual rules and the logical truth table.) If this restriction were relaxed and all mappings of pairs of values on the two relevant dimensions onto the responses were considered, the number of possible solutions would be enormous.

One method that is applicable to complete learning problems has been used by Williams (1971). The task is attribute identification with a conjunctive rule. Williams allowed the subjects to select two dimensions of each successive stimulus. The values of the chosen dimensions (but no others) were displayed, and the subject was required to categorize the stimulus. The assumption was

that, in order to respond, the subjects would need to see the values the stimulus assumed on those dimensions which the subject currently thought to be relevant. This assumption generalizes without modification to complete learning. The signal inadequacy is that the method identifies the relevant dimensions but cannot reveal the subject's hypothesis about the mapping of the values of those dimensions onto the response classes.

The final method was developed for use in a variety of concept learning tasks. It not only provides an elegant resolution to the problem of accessing the process of solution but it also leads directly to an appealing criterion for the solution of the problem. In order to ascertain whether subjects changed their hypotheses on trials when they had predicted correctly, Suppes and Schlag-Rey (1965) required the subjects to categorize every stimulus, without feedback, on every trial. Any change in categorization from one trial to the next was assumed to indicate a change in the hypothesis. On each trial, the subject was given the correct classification for one stimulus.

THE FULL SORT METHOD

Suppes and Schlag-Rey simply compared the categorizations from adjacent trials. Although they do not discuss other uses of the procedure, which will be termed the *full sort* method, it is clear that a number of valuable inferences about the subject's processing of the information available are possible: (a) the classification rule or hypothesis which generated the pattern of responses can be determined; (b) from the hypotheses, the likelihood that newly available information about the solution is recognized or that prior information is retained can be measured, and passive states in which the subject is guessing or not actively processing can be identified (cf. Falmagne, 1970; Walters, Schmidt, Bornstein, & White, 1971); and (c) memory for prior feedback can be assessed if memory aids are not available. In addition, a criterion is available which should reliably identify the point at which the subject achieves solution. In the sections that follow, the criterion for solution and the extraction of hypotheses will be elaborated.

Criterion for Solution

The problems of requiring a verbalization of the correct hypothesis have been discussed. Another common response procedure is to require that the subject predict the classification of each successive stimulus. The criterion for success is set at some number of correct predictions. In addition to the limited information which the predictions convey, another difficulty with this procedure can be seen in an example. Suppose the subjects were required to correctly predict 9 consecutive stimuli in order to demonstrate that they

had learned the partition of a three-dimension, three-value stimulus set by a conditional rule. In that case, 21 of the 27 stimuli would be exemplars of the concept. With many specific forms of the conditional rule, a subject who can select the stimuli he will learn about and who works systematically through the stimuli using the rule "Everything is an exemplar" could meet the criterion for solution quite easily.

The criterion for solution suggested by the full sort method—correct classification of the entire set of stimuli—avoids the difficulties of the prediction and verbal hypothesis criteria.² The probability of a "chance" solution is greatly reduced and communication difficulties are eliminated.

Hypothesis Extraction

The primary advantage of the full sort method is that it allows the investigator to reconstruct the hypothesis which was used to generate the classification on each trial. The algorithm for identifying the hypothesis involves two searches through the responses to the stimuli. In the first, the stimuli are searched for all sets of stimuli that differ on the value of one and only one dimension and which were assigned to the same response class. The *conjunction* of the two unchanged values defines a partial hypothesis. For example, consider a stimulus set that varied on the dimensions of letter (M, B, or S), symbol (\$, /, or #), and figure (Circle, Triangle, or Square). If the stimuli "B \$ Circle," "B \$ Triangle," and "B \$ Square" (which differ on the figure dimension but not the letter or symbol dimensions) were all placed in the category "SEV," then "B and \$ is SEV" would have been identified as a partial hypothesis. When all such sets of stimuli have been examined, a search of conjunctive partial hypotheses that have been identified is begun. If all of the partial hypotheses that share a value on one dimension and differ on the other were assigned to the same class, then the *presence* of the unchanged value defines a new partial hypothesis which is a superset of the conjunctive components. For example, if "B and \$," "B and /," and "B and #" all indicated that a stimulus was to be assigned to the "SEV" category, then those partial hypotheses could be subsumed under the new presence partial hypothesis, "B is SEV."

Any logical rule (or hypothesis) can be described by a disjunctive combination of presence and conjunctive subrules (or partial hypotheses). For example, the biconditional rule, "B↔\$ is SEV," can be written as follows:

$$\left. \begin{array}{l} B \text{ and } \$ \\ S \text{ and } / \\ S \text{ and } \# \\ M \text{ and } / \\ M \text{ and } \# \end{array} \right\} \rightarrow \text{SEV.}$$

For completeness, the components of the "not SEV" category would be

$$\left. \begin{array}{l} B \text{ and } / \\ B \text{ and } \# \\ S \text{ and } \$ \\ M \text{ and } \$ \end{array} \right\} \rightarrow \text{not SEV.}$$

Similarly, the conditional rule, "B→\$ is SEV," can be written

$$\left. \begin{array}{l} B \text{ and } \$ \\ M \\ S \end{array} \right\} \rightarrow \text{SEV.}$$

Consequently, the basis for a subject's classification on a particular trial is completely defined by specifying the partial hypotheses that can be extracted together with any stimuli which do not "collapse" into partial hypotheses.

HEAD:

A PROGRAM FOR HYPOTHESIS EXTRACTION

A computer program (HEAD) has been prepared which executes the algorithm for identifying hypotheses. Two versions of the program are available. One is written in Digital Equipment Corporation BASIC-PLUS and the other in UNIVAC FORTRAN V.

Input/Output

HEAD accepts as input the subject's classification of the entire set of stimuli on each trial.³ From these responses, it computes the response rule that was used. The program also accepts and incorporates into the output identifying information (such as the subject's identification number and the condition to which he is assigned) and aspects of his performance (such as time to solve, number of erroneous classifications, and the stimulus seen on a trial).

Requirements

The BASIC-PLUS version runs in 8K of user-accessible core. It requires a BASIC compiler that allows file input/output and a configuration with virtual storage capability. Because of the size limitations, it has poor internal documentation. The FORTRAN version, though programmed in UNIVAC FORTRAN V, has been written to be compatible with any compiler which accepts ANSI FORTRAN. The single exception is that Hollerith words of the form 'nnnn' are used in DATA statements and arithmetic expressions.

Limitations

The particular dimensions and values of the stimuli output by the program are those used by the author. Modifications of the stimuli can be accomplished by replacing a small number of DATA statements. The most salient limitation of HEAD is that it is written for

stimuli which vary on three dimensions, each having three values. Although it would not be difficult to extend the program for more values on each dimension, changing the number of dimensions would require substantial revision.

Availability

Listings of both versions of the program as well as descriptive materials are available without charge from the author. Advice about possible modifications is included. For the BASIC-PLUS version, external documentation is also available.

REFERENCES

Bourne, L. E., Jr. Knowing and using concepts. *Psychological Review*, 1970, 77, 546-556.
 Falmagne, R. A direct investigation of hypothesis making behavior in concept identification. *Psychonomic Science*, 1969, 81, 436-440.
 Falmagne, R. Construction of a hypothesis model for concept identification. *Journal of Mathematical Psychology*, 1970, 7, 60-96.
 Falmagne, R. Memory process in concept identification. *Journal of Experimental Psychology*, 1972, 92, 33-42.
 Frankel, F., Levine, M., & Karpt, D. Human discrimination learning: A test of the blank-trials assumption. *Journal of Experimental Psychology*, 1970, 85, 342-348.
 Levine, M. Neo-noncontinuity theory. In G. Bower and J. Spence (Eds.) *The psychology of learning and motivation* (v.3). New York: Academic Press, 1969.
 Nahinsky, I. D. A hypothesis sampling model for conjunctive concept identification. *Journal of Mathematical Psychology*, 1970, 7, 293-316.
 Nahinsky, I. D., & McGlynn, F. D. Hypothesis sampling in conjunctive concept identification. *Psychonomic Science*, 1968, 11, 77-78.
 Nahinsky, I. D., Penrod, W. C., & Slaymaker, F. L. Relationship of component cues to hypotheses in conjunctive concept learning. *Journal of Experimental Psychology*, 1970, 83, 351-352.
 Nahinsky, I. D., & Slaymaker, F. L. Sampling without replacement and information processing following correct responses in concept identification. *Journal of Experimental Psychology*, 1969, 80, 475-482.
 Suppes, P., & Schlag-Rey, M. Observable changes of hypotheses under positive reinforcement. *Science*, 1965, 148, 661-662.
 Walters, S. B., Schmidt, S. W., Bornstein, R., & White, R. M. Jr. Effects of pretask frequency and conjunctive rule information on concept identification. *Journal of Experimental Psychology*, 1971, 89, 351-356.
 Williams, G. F. A model of memory in concept learning. *Cognitive Psychology*, 1971, 2, 158-184.

NOTES

1. There are $\binom{3}{2}$ or 3 ways to select the relevant dimensions, $\binom{3}{1} \binom{3}{1}$ or 9 ways to select one value on each relevant dimension, eight unique, nontrivial ways to assign the presence and/or absence of the two relevant values to the response classes, and two ways to order the two relevant values in determining the truth table assignment, yielding 432 solutions. However, the order of values affects only two rules (conditional and exclusion) which reduces the number of unique rules from 8 x 2 or 16 to 10 and the number of solutions from 432 to 270.
2. The author has found a single correct classification to be preferable to two or more. If the subject is not told the criterion, he may assume that, since the task is continuing, his classification was incorrect. If the subject is told the criterion, he may persist in using an incorrect classification assuming that he is correct but has not yet met the criterion.
3. The author has developed a package of programs which formulate and run concept learning experiments by means of a PDP 11/20 computer, with Tektronix 4010 peripheral terminals. HEAD has been written to accept the output of these programs for analysis. Listings of the programs are available from the author without charge.