The Influence of Prior Theories on the Ease of Concept Acquisition

Michael J. Pazzani & David Schulenburg
Department of Information and Computer Science
University of California, Irvine

Abstract

The finding that conjunctive concepts are easier for human subjects to learn than disjunctive concepts is reported in most introductory books on cognitive psychology. In this paper, we report some conditions under which this finding may not be true. In particular, we demonstrate that the prior causal knowledge of subjects can influence the rate of concept learning. We report on an experiment that indicates that disjunctive concepts which are consistent with prior knowledge take fewer trials to learn than conjunctive concepts which are not consistent with prior knowledge. We present a computer model of this learning task.

Introduction

In concept identification tasks, it has been found that conjunctive concepts require fewer trials to learn than disjunctive concepts (Bruner, Goodnow, & Austin, 1956). More recently, it has been suggested (e.g., Murphy & Medin, 1985; Schank, Collins, & Hunter, 1986; Pazzani, in press) that a person's prior knowledge influences the speed or accuracy of learning. These claims are in part responsible for interest in explanation-based approaches to learning (DeJong & Mooney, 1986). More recently, a number of experiments have shown that with proper background knowledge people are capable of the single-instance generalization predicted by explanation-based learning (Ahn, Mooney, Brewer, & DeJong, 1987).

In this paper, we explore the interaction between the prior knowledge and the logical form of concepts. We first present an experiment in which these factors interact. Then, we present a computer model of this learning task.

Ease of Concept Acquisition: An Experiment

The purpose of this experiment was to investigate the interaction between prior knowledge and the acquisition of conjunctive and disjunctive concepts. Subjects were divided into two groups. The *Inflate* group had to perform a prediction task. This group observed photographs of a person and a balloon and had to learn to predict under which conditions a balloon could successfully be inflated. The second group, *Alpha,* used the same materials, but had a concept identification task that required learning which photographs belonged to a category called "alpha." These groups were then divided into conjunctive and disjunctive groups. The conjunction to be learned was that a balloon whose size is small <u>and</u> whose color is yellow was an alpha (or could be inflated). The disjunction to be learned was that a person whose age is adult <u>or</u> a person who is stretching a balloon is an alpha (or could inflate the balloon). Note that for the prediction task, the conjunctive concept is not consistent with prior knowledge while the disjunctive concept is. It is also important to stress that the prior background knowledge[1] (e.g., adults are stronger than children, stretching a balloon makes it easier to inflate) is not sufficient for subjects to deduce the

1. In a prior experiment (Pazzani, 1987), we asked subjects a series of True-False question about which balloons are easier to inflate. Almost all subjects indicated that adults could inflate balloons more easily than children and that a balloon that had been stretched was easier to inflate. Subjects also indicated that the color of a balloon, or dipping a balloon in water did not affect the ease of inflation. Some subjects also responded that long skinny balloons were harder to inflate than round balloons.

correct relationship in the absence of any data. There are a number of possible consistent relationships including a conjunctive one (adults can only inflate balloons that have been stretched).

The *Alpha* subjects serve as a control group to rule out the possibility that age and stretching are more salient than color and size. In a previous experiment (Pazzani, 1987), we have shown that for single attribute discriminations (e.g., action = stretching), prior background knowledge does not affect the concept identification task but does affect the prediction task. Subjects in this prior experiment took approximately the same number of trials to learn that a photograph of a person stretching a balloon was an alpha as to learn that a photograph of a person measuring a balloon with a ruler is an alpha. However, subjects required fewer trials to learn that a person could inflate a balloon that had been stretched than to learn than a person could inflate a balloon that had been measured.

We made the following predictions about the outcome of the experiments.

- Subjects in the *Alpha* conjunction category would take fewer trials than subjects in the *Alpha* disjunction category. (Conjunctions are easier to learn than disjunctions.)
- Subjects in the *Inflate* disjunction category would take fewer trials than those in the *Inflate* conjunction category. (Consistent concepts are easier to learn than inconsistent concepts.)
- Subjects in the *Inflate* disjunction category would take fewer trials than those in the *Alpha* disjunction category. (Prior knowledge facilitates learning.)

Subjects. The subjects were 88 male and female undergraduates attending the University of California, Irvine who participated in this experiment to receive extra credit in an introductory psychology course. Each subject was tested individually. Subjects were randomly assigned to one of the four conditions.

Stimuli. The stimuli consisted of pages from a photo album. Each page consisted of a close-up photograph of a balloon which varied in color (yellow or purple) and size (small or large) and a photograph of a person (either an adult or a 5 year-old child) doing something to the balloon (either dipping it in water or stretching it). For the *Inflate* subjects, the back of the page of the photo album had a picture of the person with a balloon that had been inflated or a balloon that had not been inflated. For the alpha subject, a card with the words "Alpha" or "Not Alpha" was on the reverse side of each page.

Procedures. Subjects were shown a page from the photo album and asked to make a prediction (or classification). Then the card was turned over and the subject saw the correct answer. Then the subject was presented with another card. This process was repeated until the subjects were able to predict or classify correctly on 6 consecutive trials. We recorded the number of the last trial on which the subject made an error. The pages were presented in a random order, subject to the constraint that the first page was always a positive example. If the subject exhausted all pages, the process was repeated until the correct answer was made or until 50 cards were presented. If the subject did not obtain the correct answer after 50 trials, we recorded this as the last error being made on trial 50.

Results. The results of this experiment (see Figure 1) confirmed our predictions. Figure 1 clearly illustrates that the task of learning a predictive relationship is influenced by prior theory. This effect is so strong, that it dominates the well-known finding that conjunctive concepts are easier to learn than disjunctive concepts. The interaction between the learning task and the concept to be acquired is significant at the 0.01 level $F(3,84) = 22.07$. However, the overall effect of either variable is not significant.

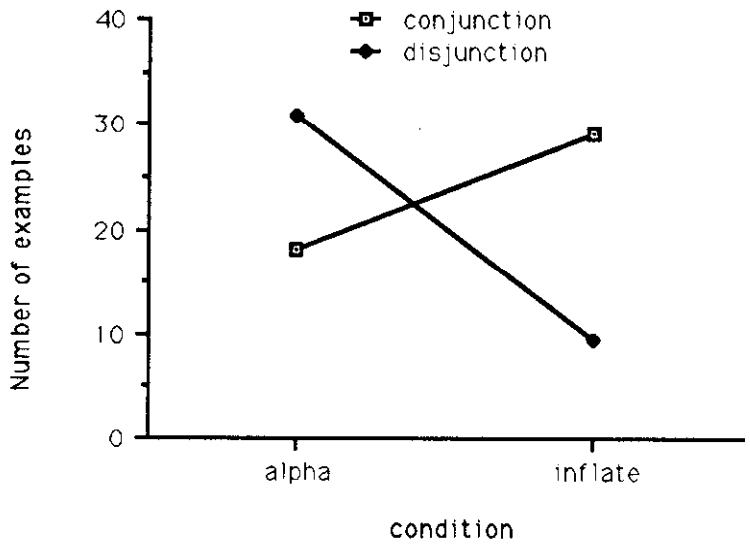

Figure 1. The ease of acquiring predictive (inflate) and descriptive (alpha) concepts. The disjunctive relationship is consistent with prior knowledge on the ease of inflating balloons, while the conjunctive relationship violates these beliefs.

Analysis of the data with a Scheffe´ test confirmed our three predictions (the results are significant at the 0.05 level):

- The *Alpha* conjunction category required significantly fewer trials than the *Alpha* disjunction category (18.0 vs. 30.8).
- The *Inflate* disjunction category required significantly fewer trials than the *Inflate* conjunction category (9.4 vs 29.1).
- The *Inflate* disjunction category required significantly fewer trials than the *Alpha* disjunction category (9.4 vs 30.8).

Discussion . The findings are consistent with our previous finding that concepts consistent with prior knowledge require fewer examples to learn accurately than concepts that are not consistent with prior knowledge. The result is especially important since it demonstrates that prior knowledge dominates the commonly accepted view that disjunctive concepts are more difficult to learn than conjunctive concepts. The result of the classification task with the same stimuli rules out an alternative explanation for these findings based on cue salience (Bower & Trabasso, 1968). This experiment raises important issues for empirical learning methods including neural network models (Rumelhart, Hinton, & Williams, 1986). The learning rules of purely empirical methods do not take the prior knowledge of the learner into account.

The experiment also points out inadequacies of current explanation-based methods (e.g., Mitchell, Kedar-Cabelli, & Keller, 1986) that assume that the background theory is sufficiently strong to prove why a particular outcome occurred. Purely explanation-based approaches to learning predict that subjects would be capable of learning from a single example. The background knowledge of our subjects seems to be able to identify what factors of the situation *might* influence the outcome of an attempt to inflate a balloon. However, they needed a number of examples to determine which of these factors were relevant and whether the factors were

necessary or sufficient. In the next section, we discuss a method of integrating empirical and explanation-based learning that makes use of this weaker sort of domain knowledge.

Explanation-based Learning with Weak Theories of Tendencies

Much knowledge about the world does not consist of universally true generalizations (Mackie, 1974; Goodman, 1983). Instead, it consists of less certain *ceteris paribus* generalizations of tendencies that occur in the absence of other factors. Research in cognitive psychology has demonstrated that prior background knowledge influences the rate or accuracy of learning. However, computational models of learning that make use of prior knowledge have for the most part assumed that this background knowledge consists of universally true generalizations. Here we relax this assumption by considering that background knowledge consists of tendencies or influences.[2] We have constructed a learning system called POSTHOC that uses this sort of background knowledge to propose hypotheses that are then tested against further data. This background knowledge is also used to revise hypotheses that fail to make accurate predictions. POSTHOC is also capable of performing classification tasks for which its background knowledge is irrelevant.

To model the previous experiment, the following two influences are used as background knowledge:

```
(easier (strong-actor)  (inflate balloon))
(easier (less-elastic)  (inflate balloon))
```

An example in POSTHOC consists of a set of attributes and an outcome. For example, an adult successfully inflating a small yellow balloon that had been stretched is represented as:[3]

```
((size . small) (color . yellow) (age . adult) (act . stretch)) =>
   (inflate balloon)
```

and a small purple balloon that had been dipped in water by a child that is classified as a *not alpha* is represented as:

```
((size . small) (color . purple) (age . child) (act . dip)) =>
   (not alpha)
```

In addition, POSTHOC has a set of inference rules that indicate which features used to describe an example are needed to identify when an influence is present in an example. The following inference rules are used to model the results of the previous experiment:

```
(influence (act . stretch)  (less-elastic))

(influence (old-actor) (strong-actor))
(influence (age . adult) (old-actor))
```

These rules state that stretching a balloon tends to make it less elastic; that older actors tend to be stronger actors; and that adults are old.

POSTHOC maintains a single hypothesis that consists of a disjunction of conjunctions. For example, the following represents the hypothesis that adults can inflate any balloon, or children

2. In Pazzani (in press), we consider how this sort of knowledge might be acquired.

3. This representation of the potentially salient features of an example is admittedly over simplistic. Here, we concentrate on the processes of learning and this simple uniform representation facilitates the implementation of the empirical learning component of POSTHOC at the expense of over-simplifying the representation of background knowledge.

can inflate a yellow balloon:

```
( ((age . adult))  ((age . child)(color . yellow)) ) =>
   (inflate balloon)
```

When the current hypothesis makes an error, a set of rules examine the hypothesis and the incorrectly classified example, and revises the hypothesis. Thus, POSTHOC is an incremental hill-climbing model of human learning of the type advocated in (Langley, Gennari, & Iba, 1987).

There are three sets of rules. One set deals with errors of commission in which a positive example is falsely classified as a negative example. This rule set makes the hypothesis more general. The second rule set deals with errors of omission in which a negative example is falsely classified as a positive example. This rule set makes the hypothesis more specific. The final rule set creates an initial hypothesis when the first positive example is encountered. Within each rule set, the rules are ordered by priority. The rule sets in POSTHOC are:

Initializing Hypothesis:

```
1.    IF there is an influence that is present in the example
      THEN initialize the hypothesis to a single conjunct
      representing the features of that influence.

2.    IF TRUE
      THEN initialize the hypothesis to a conjunction of all
      features of the initial example.
```

The first rule determines if there are features of the example that would influence the outcome of a positive example. This is accomplished by chaining backward from the rules that indicate that a certain outcome (e.g., inflating a balloon) is easier under certain conditions. The conditions are verified by chaining backward via influence rules to find features that are indicative of an influence. For example, if the initial positive example is an adult successfully inflating a small yellow balloon that had been stretched, POSTHOC would try to establish that the strength of the actor is an influential factor. This can be established by showing that the actor is strong. The fact that the actor is strong can be verified because the example indicates that the actor is adult. The initial hypothesis is that adults can inflate balloons. In this example, since there is more than one influence present, one is selected at random. This is true of all of the rules in all of the rule sets. The second rule in this rule set initializes the hypothesis to the first positive example. This occurs if there are no influences present that would account for outcome. This is true for the classification task because there are no rules that indicate factors that influence whether or not something is classified as an alpha. This can also occur for the prediction task if no influence predicted by prior knowledge is present in a positive example.

Errors of Omission

```
1.    IF the hypothesis was formed with background knowledge
      AND there are features that indicate an additional influence
      THEN create a new conjunct of the features indicative of the
      influence.

2.    IF the hypothesis is a single conjunct
      AND a feature of the conjunct is not in the example
      AND the conjunct consists of more than one feature
      THEN drop the feature from the conjunct
```

3. IF TRUE
THEN add a new conjunct consisting of a random feature from the example and simplify the hypothesis.

The first rule applies only if the current hypothesis is consistent with prior knowledge and the features of the example indicate the presence of an additional factor. This additional factor is assumed to be a multiple sufficient cause (cf. Kelley (1971)) and a new conjunct is added. This rule would add a second conjunct (act . stretch) to the hypothesis (age . adult) if an example of a child inflating a balloon that had been stretched is encountered. The new hypothesis indicates that adults can inflate a balloon or anyone can inflate a balloon that has been stretched.

The second rule is a variant of the wholist strategy in (Bruner, et al., 1956) that drops a single feature rather than all features that differ between the misclassified example and the hypothesis. In case of ties, one is selected at random. Subjects in the Alpha group of the experiment learned conjunctive concepts more slowly than the wholist strategy would predict.

The final rule forms an additional conjunct from a random feature of the example when hypotheses consistent with background knowledge and conjunctive hypotheses have been ruled out. The simplification of the hypothesis affects the form of the hypothesis to make it more concise and understandable but does not affect the rate or accuracy of the hypothesis. It consists of a number of simplification rules (e.g., X or XY <=> X).

Errors of Commission

1. IF the hypothesis was formed with background knowledge AND for each true conjunct there are features not present in the current example that would be necessary for an influence THEN modify the conjuncts by adding the additional features that are indicative of the influence.

2. IF TRUE
THEN specialize each true conjunct of the hypothesis by adding the inverse of a feature of the example that is not in the conjunct and simplify.

The first rule adds a multiple necessary cause to the hypothesis (Kelley, 1971). For example, if the hypothesis is that all adults can inflate balloons, an error will occur on an example of an adult not inflating a large yellow balloon that has been dipped in water. The hypothesis is modified by finding an additional factor which could affect the outcome that is not present in the example (stretching the balloon) and asserting that this is necessary to inflate the balloon. The new hypothesis consists of a single conjunct that represents the prediction that adults can only inflate balloon that have been stretched.

The second rule specializes a hypothesis by adding additional features to each true conjunct. For example, if the hypothesis were yellow balloons or purple balloons that had been dipped in water are alphas:

```
( ((color . yellow))  ((color . purple) (act . dip)) ) =>
   alpha
```

and the following example is encountered:

```
((size . small) (color . yellow) (age . child) (act . dip)) =>
   (not alpha)
```

then the example will be falsely classified as an alpha because (color . yellow) is true.

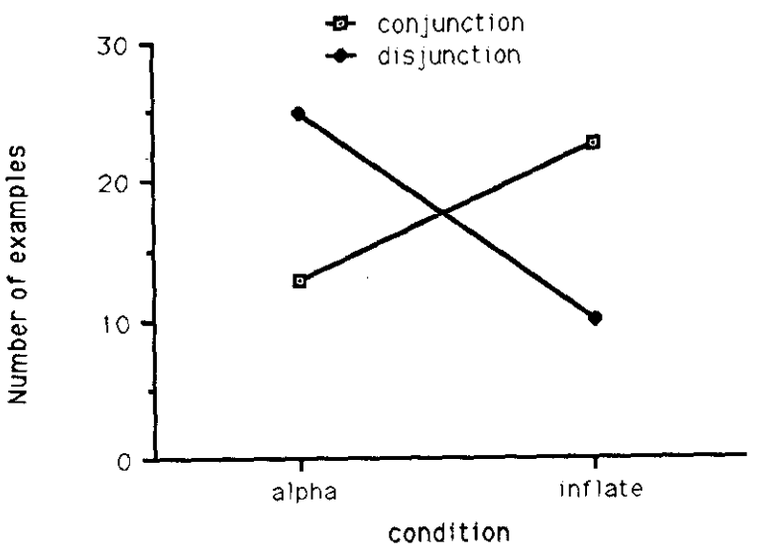

Figure 2. The result of a computer simulation of the learning experiment.

This hypothesis is modified by finding the inverse of a feature of the example (e.g., size) and asserting that this is necessary when the color is yellow. If this change turns out to be incorrect, later examples will force further revision of the hypothesis.

```
( ((color . yellow)(size . large)) ((color . purple)(act . dip)) )
   => alpha
```

Results

We ran 200 trials of the POSTHOC on each of the four conditions from the experiment. The results of this simulation are shown in Figure 2. Analysis of the data from this simulation confirms the same three predictions from the human learning experiment. In the absence of prior knowledge, conjunctions are easier to learn than disjunctions, concepts consistent with background knowledge are easier to learn than concepts that violate prior knowledge, and prior knowledge facilitates learning. Inconsistent conjunctive concepts (e.g., only small yellow balloons can be inflated) are more difficult for POSTHOC to acquire because the initial hypothesis typically includes a number of irrelevant attributes (e.g., `(age . adult)`) predicted to be relevant by the weak domain theory. These irrelevant attributes are incrementally dropped from the hypothesis when they cause errors.

Conclusions

We have presented experimental evidence that prior knowledge influences the ease of concept acquisition. This experiment suggests additional experiments which we are in the process of running. Our model predicts that consistent conjunctive concepts are easier to acquire than inconsistent conjunctive concepts or inconsistent disjunctive concepts. In addition, if there are redundant relevant cues (Bower & Trabasso, 1968), our model predicts that subjects will attend to features consistent with their prior knowledge. Initial results on these experiments appear to confirm these predictions. A computer model of this task has practical applications as well since it extends the capabilities of explanation-based learning systems to deal with weaker domain theories by including an empirical component to test and revise hypotheses.