Laboratory #1. Inference ('guess the process').

We use inference every day to make probabilistic statements from past experience. We might, for example, notice that relatively small snowflakes have begun to fall, and from this expect a big snowfall. Of course not all snow storms that start with small flakes end up dropping a large amount of snow, but the association is frequent enough to be captured in a rule of thumb ("Big snow, little snow; Little snow, big snow") and so serve as a rough guide to how much snow we might expect from a storm.

Inductive methods, where we proceed from the known to the unknown are used to generate knowledge about our world, including the world of living things. Any inductive method can lead us astray, however, because there is no necessary reason that a generalization to the unknown must be correct. We might, for example, generalize that all mammals bear live young as Europeans did platypus from Australia forced discarding the generalization. This process of inferring and discarding ("conjectures and refutations") is a very effective way of learning about the natural world, especially if we use some well established rules that make the whole process more efficient, and we are conscious of how easily we become attached to our own conjectures.

The purpose of this lab is to provide you with experience in casual and systematic application of inductive methods.

Specific goals of this lab are:

1. Compare experimental with observational methods of testing and developing generalizations.
2. Compare formal decision trees with more casually designed experiments to test and develop generalizations.

The use of formal decision trees in scientific research was advocated most strongly by Platt (1964), who argues that other methods are slow and ineffective. Here is the first part of Platt's article, which describes the explicit use of decision trees in making inferences.

Scientists these days tend to keep up a polite fiction that all science is equal. Except for the work of the misguided opponent whose arguments we happen to be refuting at the time, we speak as though every scientist's field and methods of study are as good as every other scientist's, and perhaps a little better. This keeps us all cordial when it comes to recommending each other for government grants.

But I think anyone who looks at the matter closely will agree that some fields of science are moving forward very much faster than others, perhaps by an order of magnitude, if numbers could be put on such estimates. The discoveries leap from the headlines--and they are real advances in complex and difficult subjects, like molecular biology and high-energy physics....

Why should there be such rapid advances in some fields and not in others? I think the usual explanations that we tend to think of--such as the tractability of the subject, or the quality or education of the men drawn into it, or the size of research contracts--are important but inadequate. I have begun to believe that the primary factor in scientific advance is an intellectual one. These rapidly moving fields are fields where a particular method of doing scientific research is systematically used and taught, an accumulative method of inductive inference that is so effective that I think it should be given the name of "strong inference." I believe it is important to examine this method, its use and history and rationale, and to see whether other groups and individuals might learn to adopt it profitably in their own scientific and intellectual work.

In its separate elements, strong inference is just the simple and old-fashioned method of inductive inference that goes back to Francis Bacon. The steps are familiar to every college student and are practised, off and on, by every scientist. The difference comes in their systematic application. Strong inference consists of applying the following steps to every problem in science, formally and explicitly and regularly:

1) Devising alternative hypotheses;
2) Devising a crucial experiment (or several of them), with alternative possible outcomes, each of which will, as nearly as possible, exclude one or more of these hypotheses;
3) Carrying out the experiment so as to get a clean result;
1') Recycling the procedure, making subhypotheses or sequential hypotheses to refine the possibilities that remain; and so on.

It is like climbing a tree. At the first fork, we choose--or in this case, "nature" or the experimental outcome chooses--to go to the right branch or the left; at the next fork, to go left or right; and so on. There are similar branch points in a "conditional computer program," where the next move depends on the result of the last calculations.

continued... (over)

...Continued from previous page

And there is a "conditional inductive tree" or "logical tree" of this kind written out in detail in many first-year chemistry books, in the table of steps for qualitative analysis of an unknown sample, where the student is led through a real problem of consecutive inference: Add reagent A; if you get a red precipitate it is subgroup alpha and filter and add reagent B; if not, you add the other reagent B'; and so on.

On any new problem, of course, inductive inference is not as simple and certain as deduction, because it involves reaching out into the unknown. Steps 1 and 2 require intellectual inventions, which must be cleverly chosen so that hypotheses, experiment, outcome, and exclusion will be related in a rigorous syllogism; and the question of how to generate such inventions is one which has been intensively discussed elsewhere (2,3). What the formal schema reminds us to do is to try to make the inventions, to take the next step, to proceed to the next fork, without dawdling or getting tied up in irrelevancies.

It is clear why this makes for rapid and powerful progress. For exploring the unknown, there is no faster method; this is the minimum sequence of steps.

Reprinted from:
J.R. Platt (1964). Strong Inference. *Science* **146**: 347-353. © AAAS, Washington, D.C.

Does Platt's method of strong inference work better than other, less formal methods? To find out, we will play 3 versions of a game called inferential cards. The third version uses Platt's method of strong inference. If Platt's method is better we'll be able to measure the difference by comparing the 3 versions.

The rules for inferential cards are simple.

--One person, ("Nature") develops a rule for distributing cards into two or more piles.
--This person then shuffles a deck of cards, takes cards one at a time off the top of the deck, and places them face up in piles according to the rule.
--The other players ("the scientists") observe how cards are placed.
--As soon as possible, the scientists guess the rule for placement.
--The person placing the cards replies "yes" or "no" to each guess or "conjecture."
--The game continues until the scientists correctly guess the rule, or until the deck is exhausted.

Laboratory #1. Inference ('guess the process')

To start, form up into groups of 3 (groups of 4 or 2 will also work, but 5 is too many!). The dealer ("Nature") thinks up a rule, then places cards according to the rule. The scientists work together to guess what the rule might be, then present that rule to "Nature." In true decision-theoretic fashion "Nature" can **only answer YES or NO** to each conjecture posed by the scientists. When the rule is guessed (or the deck exhausted) write the rule on the form that has been drawn up for this purpose (at the end of this lab). Fill out the names of the people who have seen the rule, and record the number of cards that were placed before the rule was correctly guessed. Record this number of cards placed under "random" cards. Fold the paper so that names are on the outside and rules on the inside, out of sight. Hand this in to the lab instructor--we will be using these again later.

After one game, switch roles so that the dealer (who made up the rule) is now a scientist. The new dealer makes up a new rule. When this has been guessed, write the rule and the score on a new piece of paper, folded with rule on the inside, names on the outside. Pass this also to the lab instructor.

The next step will be to compare this method, based on observations from cards in random order, to a more experimental method, based on selected cards. We will find out whether the scientists can guess the correct rule more efficiently by selecting cards for "Nature" to classify.

To make the comparison, we will use rules that have already been guessed by the observational method. One person ("Nature") should obtain one of these rules from the lab instructor, then read it silently. The scientists, who now hold the deck of cards, decide which cards to show to "Nature" for placement in piles. As soon as possible, the scientists make a conjecture as to the rule. As before, "Nature" can only reply "NO" or "YES." When the rule is guessed, add everyone's name to the list of people who have seen the rule. Record, as "selected cards," the number of cards to guess the rule. Did the scientists do better (fewer cards) than with the randomly presented cards?

Everyone should have a chance to use the experimental method, so appoint another person to be "Nature" and repeat this with another rule. "Nature" should go find the lab instructor in order to exchange the rule just guessed for a new rule that your group has not seen. The scientists again select cards for "Nature" to place according to the rule. Conjectures continue until the rule is guessed. The number of cards to guess the rule is recorded as "selected cards." Now that you have seen the rule, add your names to the sheet for this rule.

The last step will be to see if Platt's method reduces the number of cards to guess the rule. By now we have arrived at a fairly sophisticated style of "conjecture and refutation" where we deliberately choose a sequence of cards in order to infer the correct rule as efficiently as possible. We have remained casual about how we develop this sequence. We have been using an informal decision tree, rather than working from a formal decision tree. Platt (1964) states that a formal decision tree (strong inference) works more efficiently. We'll test this.

Delegate someone to be "Nature." While "Nature" is off exchanging an old rule for a new rule, the scientists re-read Platt's 3-step program, devise alternative hypotheses (step 1), and devise a crucial experiment to distinguish these hypotheses (step 2). Platt recommends "multiple working hypotheses" which you will find makes step 2) difficult. Because you have already had experience with rules for placing cards, you can use hypotheses based on your experience for the first pair of hypotheses. Here is an example of Platt's multiple working hypothesis cycle.

1) H1: Cards by suit (4 piles).
   H2: Cards by color (2 piles).
2) Test: Ace of spades and ace of clubs presented to "Nature"
3) Result: Cards in separate piles so H2 rejected, H1 retained.
4) H1: Cards by suit. (H1 retained).
   H3: Face cards in separate piles (new hypothesis added, based on previous result.
5) Test: Queen of diamonds and Queen of clubs presented to "Nature"

Repeat the cycle by devising pairs of hypotheses, presenting "crucial" cards to nature, and reporting the result. Keep a record of your sequence of hypotheses, treatments, and outcomes as you will need this later. A hypothesis must be retained in two cycles in order to be proposed to Nature as the rule. Otherwise, one could reject H1 on the first cycle without using a card.

When the rule is guessed, add the names in your group to the list for this rule. Then write down the number of Platt "crucial cards" required to guess the rule. Was there an improvement over the two previous methods used to infer the rule?

Table 1.1. Number of cards until rule is guessed correctly.

	random	selected	crucial	test
Rule 1				
Rule 2				
etc…				
Mean number				
Standard deviation				

At the end of the lab results will be tabulated and sent to you, along with the mean and standard deviation for each column Fill in the table lines 1 and 2 with two examples, then show the mean and standard deviation for each column.

Optional. Platt recommends multiple working hypotheses at each step, rather than sequence of mutually exclusive Yes/No decisions on a single hypothesis at each step. Platt does not mention mutually exclusive Yes/No hypotheses, nor does he give reasons for multiple working hypotheses. Studies have shown that people tend to use "confirmation bias" when inferring rules for patterns. That is, they stick with a conjecture and try to show that it is correct, rather than trying to discredit it. They become attached to their idea.

One way to correct this is to use Yes/No decisions to try to disconfirm a working hypothesis, so as not to become attached to one idea. If your group is interested, try using the method of Yes/No disconfirmation, rather than a sequence of mutually exclusive pairs as shown above. To do this, you will need to form Yes/No pairs on a single hypothesis at Step 1). At step 2) you will need to select one or more "crucial cards" to present for placement in piles by "Nature" to try to disconfirm your single hypothesis.

If you are interested in finding out whether multiple working hypotheses work better than Yes/No disconfirmatory tests on a single hypothesis then send "Nature" to obtain a rule from the lab instructor. Make sure the rule has already been tested by random cards, selected cards, and Platt crucial cards. Keep a record of your full sequence of hypothesis tests, treatments (cards presented), and outcomes in your lab notebook. Record the number of cards required on the rule sheet, as "test (either/or) cards" for comparison with the random, selected, and Platt crucial cards methods.

Write-up.

Hand in a copy of Table 1.1, first 3 columns. Add test (either/or) column if you did this.
The results of the lab will be tabulated at the end of the session, or sent later to everyone via e-mail.

Hand in one example of a Platt (crucial card) “logical tree.”

Provide a brief written comparison of the numerical results in Table 1.1.

Discuss the following questions.

1. Did a selected sequence of cards improve the efficiency of inference, as compared to a random sequence? If so why? If not, why not?
2. Did a Platt decision tree improve the efficiency of inference, as measured by number of cards, relative to an informal (selected card) tree? If so why? If not, why not?
3. You might also want to comment on Platt's (1964) use of "men" instead of "people" in the third paragraph of the article. (This part will not be marked).

Extra: Add a graph showing the results for all the rules used in your section.

Inferential Cards.

The rule should be kept out of sight until guessed. NO clues.

"Nature" is allowed to say "yes" or "no" in good decision theoretic form, as with statistical analysis.

Write out the rule here. When your group is done, use the reverse side to write names of people (either as a guesser or as Nature) who have seen this rule.

Before placing cards according to the rule, make sure no one has seen the rule (names on reverse).

RULE:

Number of random cards ____

Number of selected cards ____

Number of crucial cards ____ (multiple working hypotheses)

Extra
Number of test (either/or) cards _____   (Number of experiments ______)

Number of __________________ cards
(selected/test/crucial)

Inferential Cards

Names of people who have seen this rule.

Names ______________________________ ___________________________

Names ______________________________ ___________________________

Names ______________________________ ___________________________

Names ______________________________ ___________________________

Names ______________________________ ___________________________

Names ______________________________ ___________________________

Names ______________________________ ___________________________

Names ______________________________ ___________________________

Names ______________________________ ___________________________

Names ______________________________ ___________________________