Human Cognitive Emulation

by Lee Rumpf
2006-2007 Period 3 TJHSST Computer Systems Techlab

-Abstract-

This project attempts to accurately model human responses to stimuli. Using a survey format and decision tree learning, this research hopes to produce a unique response to a stimuli based on information gained about the user. Analyzed solely on its own, the ramification of this project can perhaps draw broad conclusions about groups of people and how they respond. When combined with other techniques of emulating human though patterns, computer programs can come closer to representing accurately human responses.

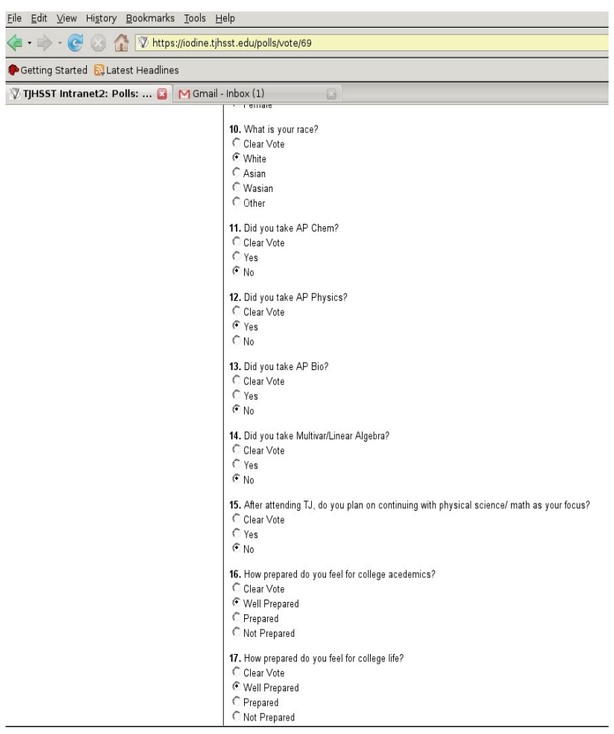
-Procedure and Results-

To get data, a survey was posted on the TJ Intranet and filled out by over 250 seniors. It asked 14 questions about various aspects of their academic and social life, followed by three questions of interest. The questions of interest were 1) After attending TJ, do you plan on continuing with physical science/ math as your focus? 2) How prepared do you feel for college academics? And 3) How prepared do you feel for college life? After the data was gathered it was converted to a csv file and used to make a tree. To test, the user simply follows the tree down from “trunk” to “branch”. If, when the user reaches the bottom of the tree, the answer the tree gives to the qualifier is equal to the users answer, the program was a success. When all was done, the program produced mostly accurate answers/predictions, but because of the detail of the tree the best results actually arise from the math within ID3. The level of the tree determines importance, and so Time in the Systems Lab, Whether or not a student takes higher math, and Time in Sports were the most important factors in question 1, 2, and 3 respectively.

-Methods and Procedures-

ID3 is the heart of Decision Tree Learning. It was developed in 1975 and I use it in LISP

```
(defun id3 (examples target.attribute attributes)
  (let (firstvalue a partitions)
   (setq firstvalue (get.value target.attribute (first
examples)))
   (cond
    ((every #'(lambda(e)(eq firstvalue (get.value
target.attribute e)))
            examples)
     firstvalue)
    ((null attributes)
     (most.common.value target.attribute examples))
    (t
     (setq partitions
           (loop for a in attributes collect (partition a
examples)))
     (setq a (choose.best.partition target.attribute
partitions))
     (cons (first a)
           (loop for branch in (cdr a) collect
                 (list (first branch)
                       (id3 (cdr branch)
                            target.attribute
                            (remove (first a) attributes)))))))))
```



An Intranet survey, left, was used to gather data from over half of the Senior class about their academic and social lives, relating to future plans and feelings about college.

Entropy is used to calculate the information gain of each classifier. To the right is a graph of entropy vs. p. Where p is the proportion of examples positive and negative