Teaching Robots How to Discover What Humans Want

Louise F. Gunderson and Donald E. Brown

Department of Systems and Information Engineering,
University of Virginia, Charlottesville, VA
lfg4a@virginia.edu, brown@virginia.edu

ABSTRACT

One of the problems in the introduction of robots into human societies, is teaching the robots about human preferences. In this paper, a data mining methodology, developed for the analysis of criminal preferences, is extended to allow robots to learn about human preferences.

1 INTRODUCTION

One of the problems in the introduction of robots into human societies, is the "house training" of the robot. Some of the "house training" rules are obvious:

- Robots should not run into furniture;
- Robots should not break things unnecessarily
- The lawn mowing robot should not run over the cat (at least not with the blade on).

However, some of the rules dealing with people are more subtle and difficult to implement. For example, some rules might be:

- Don't startle the people
- Don't annoy the people.

In fact, these rules may be more than difficult, they may be impossible, since what is annoying for one human may not be annoying for another. The robot needs

- A classification system for the idiosyncrasies of people and
- A method of placing people in these classification.

In short, the robots need a way to interact with people that is similar to the interaction between humans and dogs. Previous work has explored the use of clickers, of the type used to train dogs, in the training of robots (Kaplan, *et al*, 2001). In this paper, a methodology for data mining individual human preferences using a clicker system is discussed. This data mining methodology was originally designed for use in modeling human criminal preferences in computer targets (Brown and Gunderson, 2002).

2 APPROACHES TO FINDING HUMAN PREFERENCES

In the last few years there has been research in to the derivation of human preferences to use in multi-agent models. These models attempt to predict behavior in a specific geographic or social environment Recent work on the simulation of recreational behavior, used the judgments of experts to construct the models (Itami *et al*, 2000). In cases where there is a survey of a representative population, then statistical methods can be used to discover the preferences of the agents (Bhat, 2000; Train,1998) (Gimblett, et al, 1998). However, the surveying of human households is very difficult and puts an undue burden on the occupants of that household. It would be much better if the robots could actually derive the human preferences for themselves.

The data-mining method described below was designed to find the target preferences of criminals (Brown and Gunderson, 2001), (Brown and Gunderson, 2000). This is a situation in which the derivation of these preferences can only be accomplished by examination of the data caused by the behavior of the criminals. It is unlikely that criminals will accurately reveal their preferences to an investigator. Even if they would, the population of criminals is likely to come from prison populations, and therefore to be biased, in that only the criminals who got caught are represented. In order to deal with this type of data, the cluster specific salience discovery method (CSSD) was designed to discover the preferences of the criminals from data about their targets.

3 MODELING HUMAN PREFERENCE

Let us consider a vacuum cleaning robot living in a house with two people. We will assume that the robot can tell them apart and that the people have some way of communicating "good robot" or "bad robot, no biscuit" to the robot. This could be as easy as a personalized key chain remote. Let us further assume that the robot has a noise making capacity and some speed control.

One of the humans (human A) is nervous and does not like to be startled. They prefer that the robot make a

noise and they prefer that it move slowly. The other human (human B) does not want to know that the robot exists. They prefer that the robot be silent, but they do not care about the speed with which it moves. Neither of the humans wants the robot to vacuum in their presence. The problem facing the robot is that of how to discover these preferences without being told explicitly.

Judgment analysis is an *a posteriori* method of assessing how a decision maker formed a judgment (Cooksey, R.W.,1996 Pg. 42). This theory is based on the work of Egon Brunswik, who viewed the decision-maker as being embedded in an ecology from which he received cues as to the true state of things (Brunswik, E., 1956, Pg. 5). These cues are probabilistically related to the actual state of events. Brunswik's original theory has been extended into many judgment domains including meteorological forecasting (Stewart, T.R. *et al*, 1989), social welfare judgments (Dagleish, 1988), the understanding of risk judgments (Bushell and Dalgleish, 1993), and medical decision making (Wigton, 1988). More information about Brunswik's work on human cognition can be found in Hammond and Stewart's review (Hammond and Stewart, 2001)

In judgment analysis, the judgment process is represented by the lens model (Cooksey, 1996, Pg. 12). To discuss this model, let us consider the simple example of estimating the distance to a child's building block, lying on a table. In this model, the actual distance to the block is an environmental (distal) variable (y_e). The observer has a series of observable (proximal) cues (c_i) relating to this distal variable, such as the size of the retinal representation of the block, the differences in the image in the right and left eyes, and the blurring of the image. These cues have a correlation to the actual state (Ecological Validity). The subject weights the cues and uses a function of these weighted cues to make a judgment as to the true state (y_s). This cue weighting has a correlation to the relationship of the cues to the actual state (Cue Utilization Validity). The actual achievement (performance) in the judgment task can be used to update the weights placed on the cues in future judgment tasks. This model is shown graphically in Figure 1.

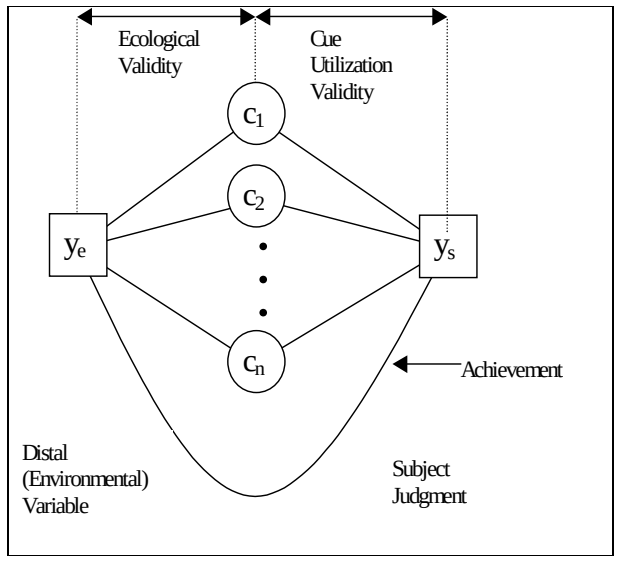

Figure 1 - Lens Model

In this case the robots is estimating the human's judgment process. It uses the features of their behavior to determine the preferences of the humans in their environment.. Some example features would be:

- Level of noise
- Speed of movement

Let us consider the example above. Each human makes a judgment as to the "goodness" of the robot's behavior, which they communicate to the robot. This judgment can be represented by a weighted sum:

$$y_s = \sum_{i=1}^{n} w_i c_i$$

Where y_s = the judgment of the target s
w_i = the weighting of cue i
c_i = the ith cue
n = the total number of cues

Looking only at the cases where the robot was rewarded, i.e. a judgment of "good robot", we would see for human A, that the rewards would be centered on some noise, low speed, vacuum off. For human B, the rewards would be centered on low noise, any speed, vacuum off. This is shown graphically below

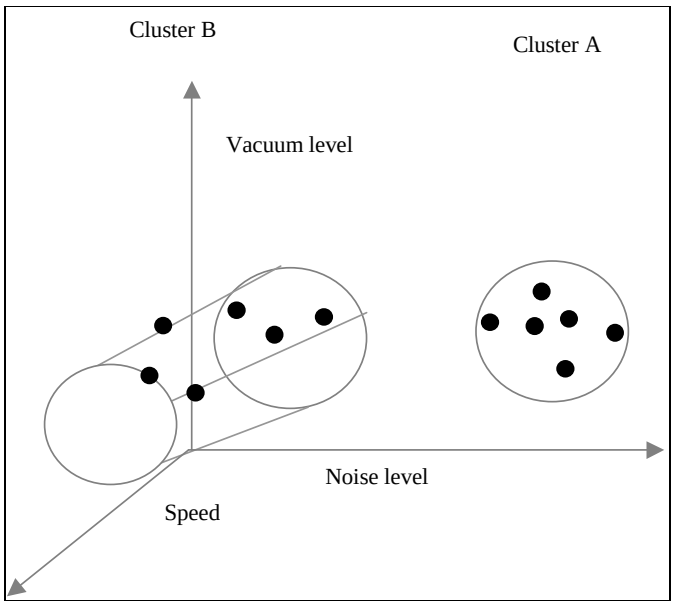

Figure 2 - Analysis of Two Behavior Patterns

The cue weighting term can range from 0, for cues that the human considers irrelevant, to very large, for cues that the human considers as salient to the choice of a behavior. This term then becomes representative of the salience of the cue to the human, and is termed the salience weighting of the feature.

For an interval feature, if the salience weighting approaches zero, the distribution of the events in the feature space will be uniform. If the feature is categorical, then the events will be uniformly distributed among the categories. For a non-zero salience weighting, the events will be grouped around the maximum preference, with the tightness of the grouping being proportional to the strength of the salience.

So, this model suggests that the events caused by a specific human should have the following characteristics:

1. A relatively small variance along any feature for which the human has a relatively large salience weighting.
2. A relatively large variance along any feature for which the human has a relatively small salience weighting.

This allows us to use the grouping of the rewards to determine the preferences of the humans. One method of grouping objects according to perceived similarities is clustering (Everitt, 1993; Hartigan, 1975; Jain and Dubes, 1988). Clustering algorithms are used in a wide array of classification problems, including the analysis of social networks and spatial distribution of organisms (Gordon, 1999).

However, this leads to a problem. To explain this let refer back to Figure 2, with the two humans, A, who cares about 3 features, and B, who cares about only 2 of these features. If the cluster formed by human A is considered, then 3 features are required to discover the cluster. If the cluster formed by human B is considered, then only 2 of the features are required to discover the cluster and the addition of the third feature can obscure the cluster structure.

This results in a series of clusters with different cluster specific salience weighting (Mirkin, 1999). The next section discusses a method to determine the appropriate cluster specific salience weighting for each of the clusters formed by the actions of criminals with specific preferences. This methodology can be used with any clustering algorithm

3.1 Selecting a Clustering Algorithm and Stopping Rule

A clustering algorithm discovers groups with some type of similarity from the data. Different clustering algorithms have different properties and problems. For this analysis, it is important that the algorithm should not be biased towards forming spherical clusters, because the elongation of the clusters yields valuable information about the cluster specific salience weight. However, some agglomerative hierarchical methods, for example centroid clustering and Ward's method, tend to impose this spherical shape (Everitt, 1993, Pg. 68-69) so to avoid this tendency to spherical shapes, *single clustering* (also called *nearest neighbor clustering*) was chosen.

While *single clustering* does meet all of the criteria discussed above, it poses two particular problems. First, if there are too many random points in the data set, it will tend to fuse them into a single large cluster. The only indication that this has happened is the large variances of this large cluster. Second, this method does not help in the case of overlapping clusters.. Another type of clustering, that is less prone to these problems is mixture model clustering.

In mixture models clustering methods, a probability distribution is used to determine the clusters, rather than a distance metric. Hartigan proposed a method for finding these clusters based on the means algorithm (Hartigan, 1975, pgs 113- 125). MacLachlan and Basford proposed an expectation maximization method to solve for the mixing proportions and parameters (McLachan and Basford, 1988).

Even after the selection of the appropriate clustering algorithm, there remains the problem of the selection of an appropriate number of clusters. For agglomerative

hierarchical clustering, Milligan and Cooper (1985) tested 30 stopping rules on non-overlapping clusters. They found that the best stopping criteria was the Calinski and Harabasz index (Calinski and Harabasz, 1974). This stopping rule uses the variance ratio criteria (VRC), which is the ratio of the between group sum of squares (BGSS) and the within group sum of squares (WGSS).

$$VRC = \frac{\frac{BGSS}{k-1}}{\frac{WGSS}{n-k}}$$

Where k = number of clusters
n = number of observations

The VRC is calculated for increasing numbers of clusters. The first number of clusters for which the VRC shows a local maximum (or at least a rapid rate of increase) is chosen as the appropriate number of clusters. For mixture model clustering, the Bayesian Information Criteria (BIC) is used to select the number of clusters (Fraley and Raftery, 2002)

3.2 Description of CSSW Method

After it has been determined that all of the variables are salient to at least one of the clusters, the following method is used to determine the appropriate cluster-specific salience weighting.

1. A "cutoff" variance (v) is chosen for all dimensions, where all dimensions = n.
2. A "cutoff" number (s) is chosen for the smallest number of point in a cluster.
3. The observations are clustered in all dimensions and the VRC or the BIC is calculated for all possible numbers of clusters.
4. The appropriate number of clusters is chosen.
5. The within cluster variance is calculated for each cluster with more than s points for all of the dimensions.
6. If a cluster is identified, for which the variance is less than v for all n variables, this cluster is identified and removed from the data set.
7. The remaining data is clustered in all possible subsets of n-1 variables.
8. The process is repeated until the dimensionality of all the clusters has been identified.

This method is shown graphically below.

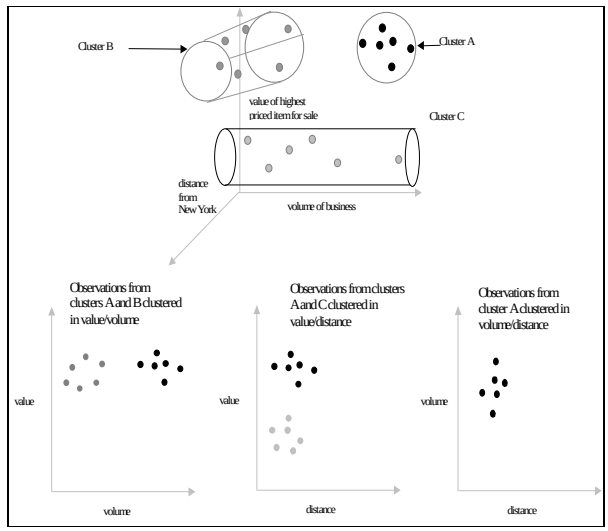

Figure 3 - Cluster Specific Salience Weighting

4 CLUSTERING RESULTS WITH SYNTHETIC DATA

It has been recommended that all clustering algorithms be tested with synthetic data with known characteristics (Milligan, 1996). A series of test cases were constructed to test the CSSD method. These were designed to represent a robot working in an area with 4 humans (agents), each with different preferences. In each of the cases 4 agents created 250 observations each. Each of the agents had a preference for each of the 4 features. This preference could be:

- N(0.25,0.03),
- N(0.5,0.03),
- N(0.75,0.03), or
- U(0,1).

For the first 10 models all the agents had gaussian preferences for all of the features. For models 10 – 19 one out of the possible 16 preferences was uniform, for models 20 – 29 two, for models 30 – 29 three, and for 30 – 40 four. All of the preferences were chosen out of a discrete distribution.

The clusters in each of the models were identified four times. The first time was with the Agglomerative Hierarchical clustering (Single Link) alone (as implemented in SAS). The second time was with the CSSD and Agglomerative Hierarchical Clustering (Single Link). Then the clusters in the models were identified with a mixture model alone (Mclust as implemented in R) Finally they were identified with CSSD and mclust. The metric for performance was the Jaccard index

$$J = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{is} + b_{ij} + c_{ij}}$$

Where a_{ij}= 1 if x_i is in the same cluster as x_j in both the prediction and observation set, 0 otherwise

b_{ij}=1 if x_i is in a different cluster than x_j in the prediction, same cluster in the observation, 0 otherwise

c_{ij}=1 if x_i is in the same cluster as xj in the prediction, different cluster in the observation, 0 otherwise.

The results are shown in the table below .

% Error	Jaccard-AHC (Single)	Jaccard-CSSD-AHC (Single)	Jaccard mclust	Jaccard CSSD mclust
0.00	1	1	1	1
6.25	0.790	0.847	0.818	0.970
12.50	0.749	0.932	0.702	0.936
18.75	0.542	0.848	0.723	0.939
25.00	0.515	0.780	0.605	0.963
Total	0.719	0.881	0.770	0.962

These results show that this algorithm can be used to separate preferences that have the same characteristics as human preferences.

5 CONCLUSIONS

In this paper, a method has been shown that will allow robots to use data mining to discover the preferences of the humans in their environment. The use of data mining allows the robot to learn these preferences in a natural way. Rather than typing in commands, or dealing with an interface, this would be like training a dog. In fact, many dog trainer recommend the use of clickers in training. The method has applicability across a wide range of robotic tasks and allows for the human to make the robot response to personal needs and preferences. More work needs to be done in the refinement of the data mining and the implementation of the method, but this is a beginning of an easier way for robots to learn what humans want.

REFERENCES

Bhat, C. 2000. "Incorporating Observed and Unobserved Heterogeneity in Urban Work Travel Mode Choices Modeling." *Transportation Science.* 34(2):228-238.

Brown, D., Gunderson, L. 2001. "Interactive Analysis of Computer Crimes." *Transaction of the IEEE, Part A.* **31**(4), Pgs. 311 – 318.

Brown, D. and Gunderson, L. 2002. "Using Data Mining to Discover the Preferences of Computer Criminal." in *Advances in Computers, Vo. 56*. San Diego, CA: Academic Press. Pgs. 344 – 374.

Brown, D., Gunderson, L. Evans, M. 2000. "Interactive Analysis of Computer Crimes." *Computer*. 33(8):69 – 77.

Brunswik, E. 1956. *Perception and the Representative Design of Psychological Experiments*. Berkeley, CA:University of California Press.

Bushell, H, and Dalgleish, L. I. 1993. "Assessment of Risk by Employees in Hazardous Workplaces." *Safety Science Monitor*. 3.

Calinski, T. and Harabasz, J. 1974. "A Dendrite Method for Cluster Analysis." *Communications in Statistics.* 3(1):1-27.

Cooksey, R.W. 1996. *Judgment Analysis*. Berkeley, CA:Academic Press.

Dagleish, L. I. 1988. "Decision making in Child Abuse Cases: Application of Social Judgment Theory and Signal Detection Theory." In *Human Judgment: The SJT View, Advances In Psychology, Vol. 54*. ed B. Brehmer, C.R.B. Joyce. Amsterdam, The Netherlands:Elsevier Science Publishers B.V.

Everitt, B.S. 1993. *Cluster Analysis*. London,England:Edward Arnold.

Fraley, C. and Raftery, A. E. (2002) *MCLUST: Software for Model-Based Clustering, Density Estimation and Discriminant Analysis*, Technical Report No. 415, Department of Statistics, University of Washington, Seattle, WA.

Gimblett H. R., Merton, R. T. and Itami, R. M. 1998. "A Complex Systems Approach to Simulating Human Behavior Using Synthetic Landscapes." C*omplexity International* 6, Available online at <http://life.csu.edu.au/complex/ci/vol6/gimblett/gimblett.html>

Gordon, A. D. 1999. *Classification: 2nd Edition*. New York, NY:Chapman & Hall, CRC..

Hammond, K. R. and Stewart, T. R. 2001. *The Essential Brunswik*. New York, NY: Oxford University Press.

Hartigan, J. A. 1975. *Clustering Algorithms*. New York, NY:Wiley.

Itami, R. M., Gimblett, H. R. "Intelligent Recreation Agents in a Virtual GIS World." In Proceedings of Complex Systems 2000 Conference. November 19-22, 2000. University of Otago, Dunedin, New Zealand. Available online at http://srnr.arizona.edu/~gimblett/NZ_RBSIM_Final 2 .html>

Jain, A. and Dubes R. 1988. *Algorithms for Clustering Data*. Englewood Cliffs, NY :Prentice Hall.

Kaplan, F., Oudyer, P-Y., Kubinyi, E. and Miklosi, A. 2001. "Taming robots with clicker training : a solu-

tion for teaching complex behaviors." In *Proceedings of the 9th European workshop on learning robots*, LNAI, Springer, M. Quoy, P. Gaussier, and J. L. Wyatt, Eds.

McLachlan, G. J. and Basford, K. E. 1988. *Mixture Models: Inference and Applications to Clustering*, Marcel Dekker, Inc: New York and Basel.

Mirkin, B. 1999. "Concept Learning and Feature Selection Based on Square-Error Clustering." *Machine Learning* 35:25-39.

Stewart, T. R., Moninger, W. R., Grassia, J., Brady, R. H., and Merrem, F. H. 1989. "Analysis of Expert Judgment in a Hail Forecasting Experiment." *Weather and Forecasting* 4:24-34.

Train K. 1998. "Recreation Demand Models with Taste Differences." *Land Economics* 74: 230-239.

Wigton, R. S. 1988. "Applications of Judgment Analysis and Cognitive Feedback to Medicine." In *Human Judgment: The SJT View, Advances In Psychology, Vol. 54*. ed B. Brehmer, C.R.B. Joyce. Amsterdam, The Netherlands:Elsevier Science Publishers B.V.