REALTIME CLUSTERING OF UNLABELLED SENSORY DATA FOR TRAINEE STATE ASSESSMENT

Keith W Brawner[a] and Avelino Gonzalez[b]

[a]Army Research Laboratory – Human Research & Engineering Directorate, Orlando, FL 32826
[b]University of Central Florida, Orlando, FL 32826
[a]Keith.W.Brawner@us.army.mil
[b]Gonzalez@ucf.edu

ABSTRACT

The grand challenge of Intelligent Tutoring Systems (ITS) development is that of creating a computer tutor as good as a human tutor. This difficult task may be broken into several parts. The first task real instructors perform prior to making instructional decisions is assessing the state of the trainee. Thus, the first consideration in the construction of an ITS is obtaining meaningful data from sensors and interpreting them in order to assess trainee emotional state. The interpretation of sensor data is the significant problem in this area, with the problem of sensor data mostly having been reduced to sensor selection. The machine learning methods for interpreting unlabelled sensor data are significantly more sparse than the sensors available, and their selection is far from straightforward. In this paper, Growing Neural Gas (GNG) methods, two types of incremental clustering, and Adaptive Resonance Theory (ART) will be evaluated against each other on fabricated and realtime data streams of trainees' state in order to determine the best selection of methods to accomplish this task.

Keywords: Intelligent Tutoring, Realtime Data Streams, Trainee State Assessment, Machine Learning

1. INTRODUCTION

Decisions by instructors regarding the trainee can be divided into two basic questions: "How well is the trainee doing?", and "What should be done about it?". In this regard, a computer instructor is no different than a human one. As computer-based instruction gains popularity, there is high demand for algorithms and methods to autonomously make these types of decisions. If a computer instructor can make these types of assessments as optimally as a human instructor, computation algorithms can be constructed to optimize learning goals more effectively than humans. Each of these are stepping stones for projects in self-directed, computer-enabled, automated, learning.

Computers hold an advantage over human instructors as they can make decisions based on sensory data not available to the human. For instance, humans cannot see in the Infrared spectrum, measure heart rate, measure galvanic skin response, or read Electroencephalography (EEG) activity levels. While humans still hold the advantage when processing visual and sound data, designers of ITSs should consider the advantages they have in order to design more effective instructional machines. The sensors chosen for automated trainee assessment should perform their individual tasks according to their ability, taking advantage of the computer's strengths.

This data is not without challenges. Sensory input to a computer instructor is likely to be usable in only a very narrow time window. Sensory data varies significantly from individual to individual and from day to day. Additionally, the data which is processed by sensors comes with no interpreted meaning. Because of the high variation in the data, a constructed model is not likely to be useful long term, and models constructed with this data must be built quickly and not reused. The difficulty in the creation of a unified, emotional model for affective sensing is a large part of why emotional sensors have not been integrated into the ITS domain (Arroyo, Cooper, Burleson, Woolf, Muldner & Christopherson, 2009).

Two traditional methods for dealing with this problem are reprocessing by incremental rebaselining, in the case of 'microclusters' (Aggarwal & Yu, 2008), or averaging over a Hamming window (Papadimitriou, Brockwell & Faloutsos, 2004). Reprocessing based on a new, changed, data stream appears to be a simple solution, but only creates more problems. Issues of reprocessing include: when to reprocess, determination of when data shift is too extreme to be usable, and the reprocessing window being a time when decisions cannot be made on that type of data. These issues point to the idea that reprocessing will not solve the general problem. Using data in a window presents its own problems, but it is generally accepted that usable conclusions can be conferred via windowed data (Hore, Hal & Goldgof, 2007).

The solution to the generalized problem lies with machine learning classifiers that are able to adapt to changes in a data stream in realtime. These methods build a model of the data seen so far, use it to make decisions, and discard, log, or internalize it after a training session is complete. Using this type of solution provides an accurate assessment of trainee state, but creates the problem of dealing with a realtime data stream.

2. REALTIME DATA STREAM PROBLEMS

Processing realtime sensory data is becoming increasingly easier as computation gains speed and takes advantage of multiple cores. Even though computers are becoming faster, the primary trouble with data streams does not lie in the speed of computation. Algorithms dealing with realtime data streams must deal with the critical issues of potentially infinite length, concept detection, concept evolution, and concept drift (Beringer & Hüllermeier, 2006). Each of these presents their own problems for classification.

Sensory data is inherently unlabelled. Sensor measurements, while accurate, do not natively imply anything about the item which they are measuring. Just as a nuclear engineer can use a temperature sensor to measure the temperature of a reactor core, the sensor does not tell him the implications of that temperature. An individual, or algorithm in place of an individual, must give meaning to the collected data. In the same fashion, measurements of heart rate variability in a trainee are not interpreted for their implications on learning and training. However, asking the trainee their state at each heartbeat, second, or minute is not only impractical, but a task that distracts from learning. Labels for state data must be computer-generated, infrequently polled from the user, or a combination of both.

2.1. Infinite Length

The fundamental problem when dealing with realtime data streams is the potentially infinite length. Any machine learning algorithm that relies on viewing all data to make a decision at current time will fail to process a potentially infinite stream. Although the data stream is not infinite in the real sense, the practical implication of infinite length is the presence of a point of diminishing returns. For any algorithm that depends on the presence all data to make a decision there is a point where there is too much data to process and make a decision before more data is presented. The implication at this point is that the algorithm can no longer perform in realtime. While certain algorithms may be used up to a hundred, thousand, or million data points the presence of the limiting number is not recoverable. The infinite length problem immediately discounts the use of traditional systems of machine-learning such as Bayesian Networks. Training algorithms that iterate over all previous inputs, such as backpropagation with Artificial Neural Networks (ANNs), also must be discarded.

As an example, in the domain of trainee state assessment from sensor streams, the duration of a given scenario-based training event is unknown and the sampling rate can be over 500 Hz with two dimensions.

2.2. Concept Detection

The second problem in dealing with realtime data streams is the presentation of a new, previously unknown, class of data. Because there is no fundamental limit to the types of observations, there must be a method for detection of the novelty. This makes many traditional, unmodified, machine learning methods impractical, such as Support Vector Machines and Decision Trees, which are unable to dynamically respond to the development of new, previously unseen, classes of data.

In the domain of interest, the number of moods or their representation is not known prior to observation. Even if the number of states is known *a priori*, there is no guarantee that any of the states will be represented in a given period, that one state will correspond to one location within the observational space, or that one state will correspond to only one location of the sampling space.

2.3. Concept Drift

The third issue when dealing with realtime data streams is the problem of the underlying concepts changing over time. In the domain of interest, a particular mood of a trainee can represent itself differently over time. For instance, in an unpublished Galvanic Skin Response (GSR) classification experiment, spikes were observed in a GSR data stream of people exposed to fear-based imagery. However, after the spike had been resolved and the image had been rescinded, both the next rest state and next image presented demonstrated elevated GSR measures. This is not indicative of the subject evidencing strong emotion in the rest state or stronger emotion in response to fear-based imagery, but instead evidence of a change in baseline observation. This further rules out ideas such as Hamming window baselining as a possible solution. The experiment that dealt with this datastream evidenced poor performance using NeuroEvolution of Augmenting Topologies, which has elements of ANNs and Genetic Algorithms, to attempt classification. This indicates that such methods would not be successful, even with realtime modifications. A 2010 study, pictured below, found similar observance patterns among GSR data.

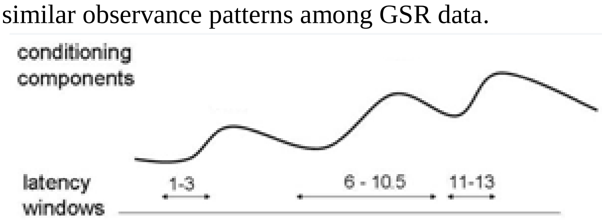

Figure 1: Example of concept drift changing baseline measurements in one-dimensional, time-aligned, GSR data (Gao, Raine, Venables, Dawson & Mednick, 2010)

2.4. Concept Evolution

The final problem present in realtime data streams is concept evolution. This is observed when an already evolved concept presents itself in a way different from before. In the domain of text mining, this is presented by using different words to express the same concept (Masud et. al. 2010). In the domain of trainee state modeling, this may represent itself as a changed baseline in one dimension and a different baseline in another; however, it is unknown what types of sensory measurements are likely to experience the evolution of one concept differently in the represented tested sampling space.

3. POTENTIAL SOLUTIONS

The nature of this problem lends itself to several branches of machine learning. Unsupervised methods can be used in order to cluster inherently unlabelled sensory information. Information can then be attached to the clusters in a semi-supervised manner to give meaning to the data. For instance, unsupervised methods can point out general data trends and use another algorithm, or infrequent survey data, to give these established categories various implications. Supervised or statistical Machine Learning can process clusters with attached meaning to recommend instructional decisions for learning goals or outcomes.

The components of this section will summarize algorithms that will be tested against the first portion of the adaptive tutoring problem; continuous, realtime, noisy, sensor data.

3.1. Incremental Clustering

Clustering is a type of unsupervised learning operating on the principle of determining the distance from an observed point in sample space to other points relative to it. If the observed point is within a threshold distance of similarity to other points, it is assumed to have a relationship to them and is clustered together. Examples of clustering include hierarchal clustering, graph clustering, and k-means (Jain, 2008).

Two types of incremental clustering algorithms are used for this paper as a baseline measurement of performance. Both of these are k-means variants adapted for incremental use on data streams of potentially infinite length. The performance of these algorithms is expected to flag on the problems of drift and evolution, but provides a reasonable baseline approach upon which to compare the other methods presented.

The first clustering algorithm used in this paper is adapted from http://gromgull.net, and closely follows the approach of online agglomerate clustering (Guedalia., London & Werman, 1998). This approach is summed up in their work as, for each point:

Move the closest centroid towards the datapoint
Merge the two closest centroids, if appropriate
    Creates one redundant centroid
Set redundant centroid equal to the datapoint

The second method is custom-developed, but follows the basic point recognition algorithm pseudo-code:

For each point
    Compare point to all known clusters
    If no cluster is within vigilance
        create new cluster here
    else
        move matched cluster up to <delta> in the direction of the recent point

3.2. Growing Neural Gas (GNG)

Neural Gas is a method of finding the optimal data representations based on feature vectors. It was inspired by the successes of ANNs and Self Organizing Maps, and has been used as an alternative to clustering. An incremental version has been created in the form of GNG (Holmstrom 2002), and is expected to outperform baseline clustering on the problems of interest. The general algorithm is below, for each new data point, although the specific implementation in that experiment may be slightly different, as the Modular toolkit for Data Processing (MDP) 2.6 implementation was used:

If appropriate (current point does not correspond to known information), create new reference arc, store error
Else increment age of all arcs in this area, move existing arcs towards new data, establish new ages for arcs
Remove Aged arcs
If any non-emanating arcs exist, remove them
If it is the time to add a new point (due to timing)
    Add a new reference point, halve the distances of the existing arcs to this point, scale the existing errors
Compute non-Hamiltonian path of all arcs (depth first)
For this point against each class:
    If there are few related nodes, compute the probability of the point belonging to the lowest error class
    Else determine the modified shape of the cluster it is most likely to belong to

3.3. Adaptive Resonance Theory (ART)

ART is a type of neural network architecture which classifies objects based on the activation of nodes in the structure. It was developed to classify data in a one-pass learning environment (Carpenter & Grossberg, 1995), and has performance roughly equivalent to neural networks with significantly reduced training time. In its most basic form ART draws n-dimensional hypercubes around similar input patterns, where n is the dimension of the input data. Matched data is the data that falls within the smallest hypercube or of the class of the closest available hypercube. Hypercubes are expanded to compensate for new data in accordance with parameter settings. Although sometimes viewed as a disadvantage, ART systems are capable of one-pass learning, and are consequentially sensitive to the input order of data. It is expected that ART systems will respond well to the problem of evolution, but poorly to the problem of drift.

3.4. Parameters used in this experiment

The parameters used in this experiment for online agglomerate clustering was a maximum cluster number of 15. The incremental k-means approach used a delta parameter, corresponding to a numeric value of how much to change in response to a new point, of 10%. It also used a vigilance parameter, corresponding to a maximum percentage error before creating a new cluster of 20%. These values were chosen based upon minor experimental sampling.

The ART algorithm used in these experiments was complement coded, with no category maximum, fast learning, a vigilance of .9, and a bias of .000001. More information about ART parameters can be found in the original paper (Carpenter & Grossberg, 1995) or about this specific implementation via the pyrobot users group.

The GNG algorithm used in this paper is implemented as part of the Open Source Modular Toolkit for Data Processing, with the following parameters: alpha = .1, lamda=20, eps_beta=.5, max_nodes=400, max_age=10, epilon_n=.06, d=0.995. These all all roughly recommended parameter settings. More information about these parameters can be found in the MDP 2.6 website: http://mdp-toolkit.sourceforge.net/, or in the published paper (Zito, Wilbert, Wiskott & Berkes, 2009).

4. PROBLEM SPACE

In order to test each algorithm, several datasets were obtained. The benchmark datasets include: a fabricated two dimensional set of predefined shapes, a fabricated two dimensional point-drawing of shapes representing movement from one class of data to the next, and a two dimensional real dataset of electrocardiology (ECG) and GSR data taken during an experiment, and a three dimensional real dataset of EEG-classified trainee state.

4.1. Ordered set of point-drawn shapes

A set of predefined shapes will be chosen to represent the problem space of generic, unsupervised, classification. The points are presented in order of shape appearance in order to establish how well each classification algorithm responds to the development of new data classes over time. Each shape is Gaussianly distributed among its boundaries with the exception of the outlined circle, which has a Gaussianly distributed ring. Points are drawn between shapes in order to simulate the gradual change of classes. This more accurately simulates the trainee state space by not providing drastically different shift. Although there is no correct answer to this dataset, the general classification of observed shapes should be a goal. While these are quickly apparent to human eyes, it is important to note that these patterns are presented to each algorithm with very little memory. This dataset tested 4250 unique points.

4.2. Randomized set of point-drawn shapes

This set uses the same set of datapoints as described in section 4.1, with the order of presentation randomized. This tests how well each algorithm remembers the states that it has previously seen. As in 4.1, this dataset tests 4250 unique points.

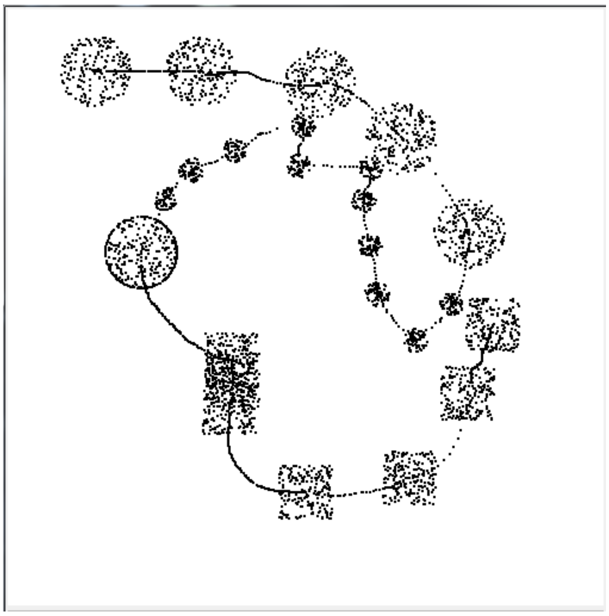

Figure 2: A set of predefined shapes, used in shape order for dataset 4.1, and used in randomized order for dataset 4.2

4.3. Feature-extracted EEG data

A set of ECR and GSR data captured in another, currently unpublished, experiment is available to the author. The experiment involved a user interacting with a digital character in well-, and ill-defined scenarios. One of the three scenarios was deliberately made frustrating for the user in order to evince particular emotional state. A random user was selected for the use of their data, in order to gauge an idea of performance on other, similar, datasets. Technologies of eMotive EEG, Galvanic Skin Response (GSR), and electrocardiology (ECG) data were captured in this experiment. It is expected that the user demonstrated various cognitive states during the various phases of the experiment. These were, in order, a rest period, a scenario which was chosen randomly among three possible, and a survey. The performance on this set of data can be used to gauge an idea of performance on other, similar, datasets. This dataset used 9,342 points, corresponding to 38 minutes of actual data collection time.

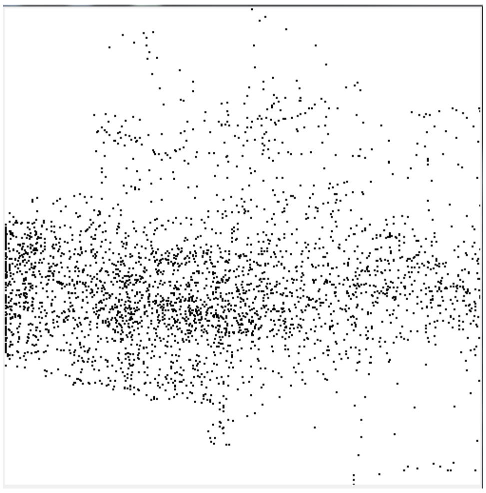

Figure 3: Classified EEG data. Short Term Excitement (STE) is along the x-axis, Boredom/Engagement is along the y-axis.

4.4. ECG and GSR data

In the experiment described above, the ECG and GSR data was collected as an additional measurement of trainee state. This dataset tested 400,000 datapoints, corresponding to 24 minutes of actual data collection time.

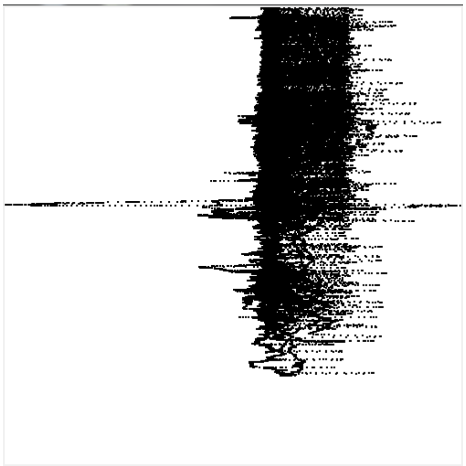

Figure 4: Unfiltered ECG and GSR data. ECG data is along the x-axis, GSR data is along the y-axis

5. RESULTS

5.1. Set of ordered, predefined shapes

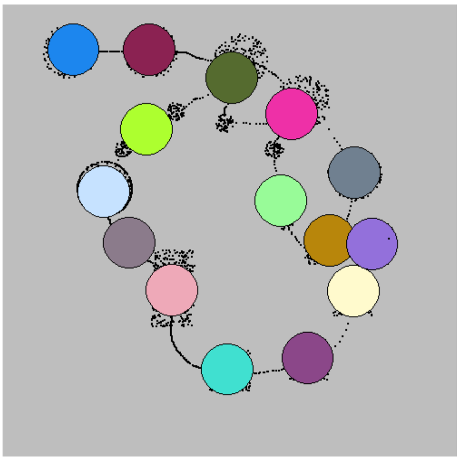

Figure 5: Online agglomerate shaped clustering

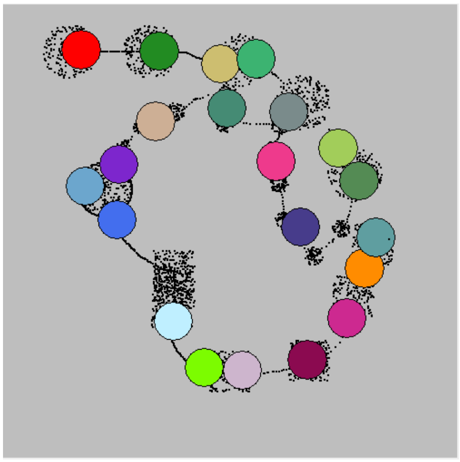

Figure 6: Incremental k-means shaped clustering

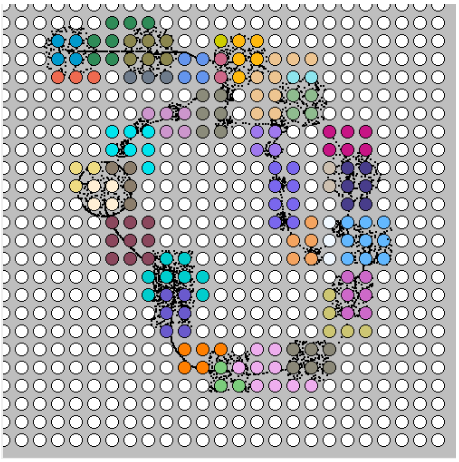

Figure 7: ART shaped clustering

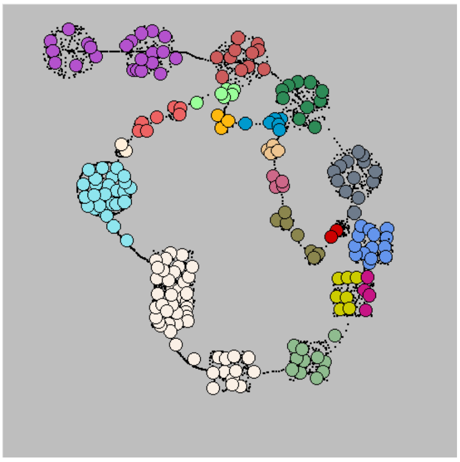
Figure 8: GNG shaped clustering

5.2. Set of random, predefined shapes

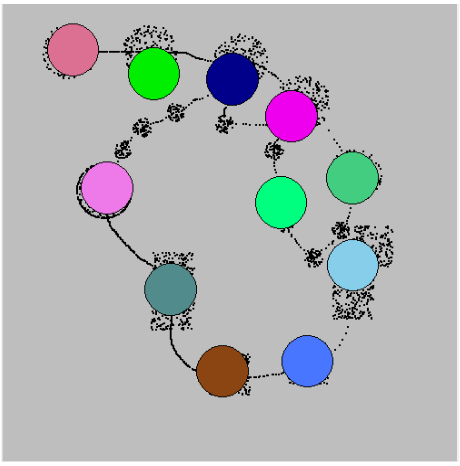
Figure 9: Online agglomerate random shaped clustering

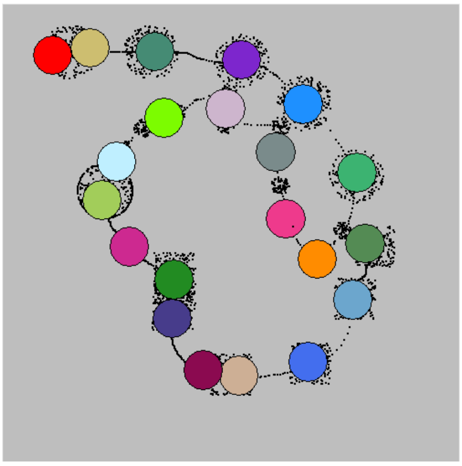
Figure 10: Incremental k-means random shaped clustering

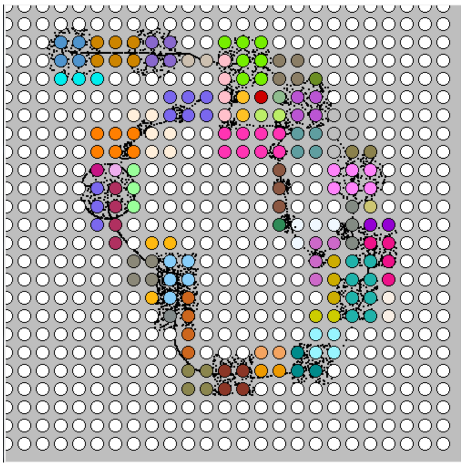
Figure 11: ART random shaped clustering

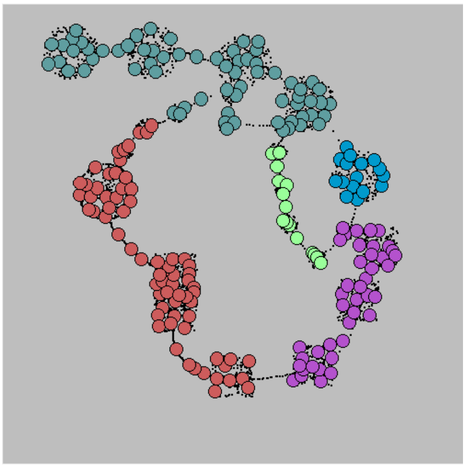
Figure 12: GNG random shaped clustering

5.3. Classified EEG data

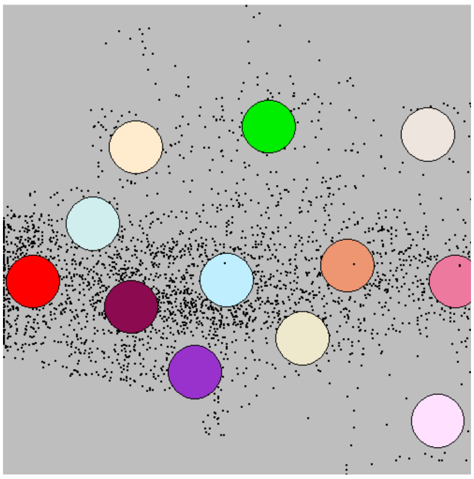
Figure 13: Online agglomerate EEG clustering

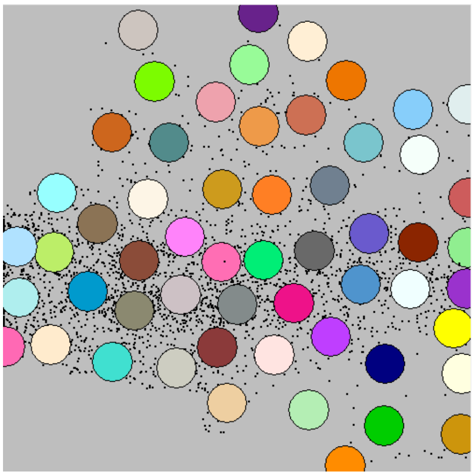

Figure 14: Incremental k-means EEG clustering

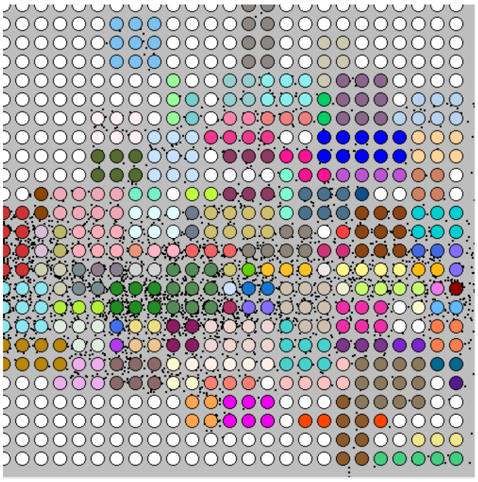

Figure 15: ART EEG clustering

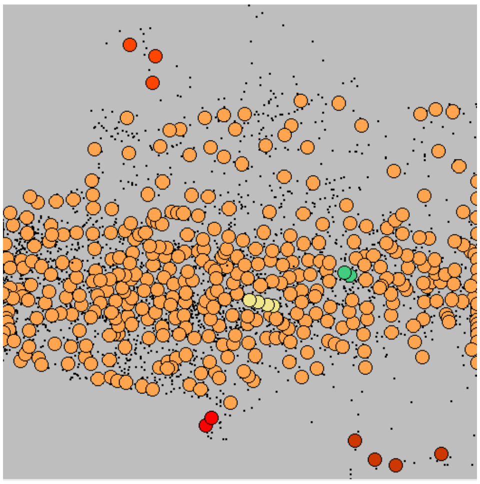

Figure 16: GNG EEG clustering

5.4. ECG and GSR data

Online agglomerate clustering did not complete this successfully within allotted time, and failed to deal with the problem of infinite length.

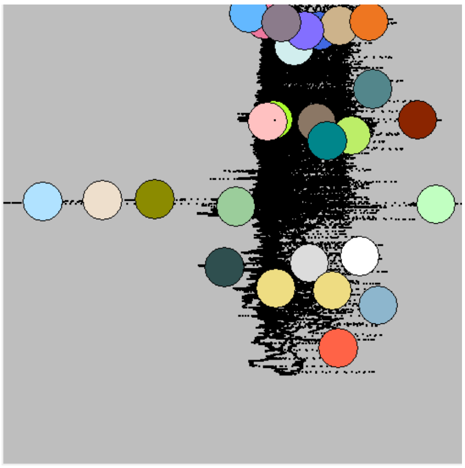

Figure 17: Incremental k-means ECG/GSR clustering

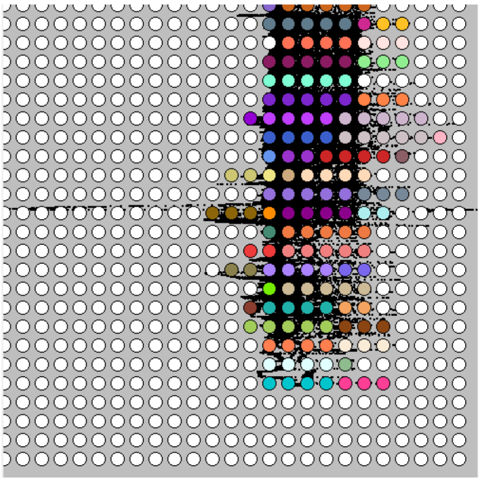

Figure 18: ART ECR/GSR clustering

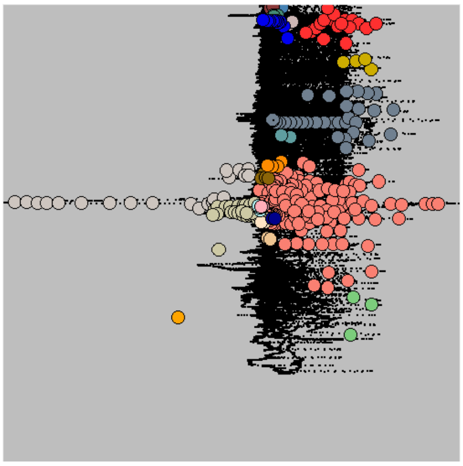

Figure 19: GNG ECR/GSR clustering

5.5. Time Analysis

All experiments were run, single-threaded and overnight with no interference on a 2.67 GHz processor with 4GB of RAM. Data was cached before clustered. Time was stopped for screen-drawing so as not to bias algorithms, such as ART and GNG, where data visualization is a lengthier. Python version 2.7 was used to handle garbage collection and other programming issues. Integrated DeveLopment Environment (IDLE), a part of Python, was run in a thread separate from processing. Given the high data processing rate of these algorithms, it was determined that the average time over multiple runs would not be required to be analyzed. However, these numbers should be taken as representative rather than guaranteed.

Table 1: Time analysis of various algorithms (in Hz)

	Online agglomerate clustering	Incre. k-means	ART	GNG
Ordered shapes	215	69672	11707	1396
Unordered Shapes	273	27243	4560	1403
EEG	189	15648	3803	801
ECG/GSR	-	45511	2035	472

6. DISCUSSION

6.1. Ordered Shapes

It is worth noting early that the size of the clusters in each of the first two algorithms are not a function of the true size, which is difficult for visualization purposes. However, the overall trend can be seen, so this is of little concern. In a time analysis, incremental k-means outperformed all other methods, with generally good responses from ART and GNG.

Agglomerate clustering does an excellent job of capturing the major data clusters, while missing the smaller clusters, and generally maintaining a fairly stable number of clusters. It hits the fifteen cluster limit, but at a generally acceptable time. Hitting this limit causes it to miss several of the smaller circles, and it classifies these as generally transitory states.

K-means, while correctly classifying many of the smaller clusters, and correctly classifying the larger clusters, misclassifies the edge-bound circle and two of the squares as a few different classes. It cannot merge these due to the deficiency of the algorithm.

ART does a successfully captures the meaning in the smaller clusters, but generally over-classifies the larger shapes. GNG, however, classifies the obvious shapes with very little error and acceptable time performance.

6.2. Randomized Shapes

In general, the performance of the incremental clustering algorithms did not degrade with the randomization of the data points. ART, however, narrows in the clusters that it is classifying, while GNG nearly completely fails to isolate individual classes of underlying structure. This behavior of the GNG algorithm is noted in other locations (Ancona, Ridella, Rovetta & Zunino, 1997).

6.3. EEG

The EEG dataset has no easily apparent patterns, but it does have locations of more data presentation instead of less. Online agglomerate clustering fails to find the patterns in the striated middle bands, as does k-means, and GNG. ART, however, classifies x-dimensional bands, which are likely to be the observed states. Additionally, outliers, which are likely to be noise, are classified into a classes of outliers in GNG and ART, but not in the online clustering algorithms.

6.4. GSR/ECG

The final dataset deals with a large amount of data to be processed, and is indicative of the worst-case, real-life situations where these algorithms may be used to classify trainee state. This data was collected at a sampling rate of 500 Hz, which puts a significant speed test on each of these algorithms. Online agglomerate clustering fails to finish classification in time. GNG also fails to complete in time, but it is worth noting that this is only by a slim margin of 472 Hz, rather than 500.

With additional speedups found in high performance computing, or simply a faster processor, it is possible that this algorithm would finish in time. However, it is worth noting that any real-life usage will likely be in more than two dimensions.

This is the second set of data with no readily determinable pattern. Incremental k-means expresses this as a series of unmerged clusters, while ART expresses this as roughly 20 striated bands. ART completely fails to capture the heartbeat, which is the most important part of the ECG data. Further experiments, which are not shown, show this to be true even at 5 times the imaging resolution.

7. CONCLUSION

The obvious conclusion, as demonstrated in the No Free Lunch Theorem (Wolpert & Macready, 1997), is that no one algorithm works best across all problems. The incremental clustering algorithms performed admirably across the domains of well-defined clustering, while performing quite poorly on real-world data. One of the important takeaways here is that cluster merging is too expensive of a proposition to allow in realtime.

Generally speaking, ART was able to classify well in realtime even when data patterns were not present or not apparent. The clusters formed by ART were of reasonable input size and classification. GNG was able to successfully classify provided that data was in the correct order and that a pattern was apparent, but was a bit slow on problems of high data input rate.

8. FUTURE WORK

Future work in this area will branch into two directions. The first direction will continue to look for algorithms that perform well on this type of data problem, or in the adaption of ART and GNG to classify fewer clusters, and operate with a quicker timeframe, respectively. Feature extraction of meaningful data will also be looked into, including the addition of time-stamp related data and its effect on formed clusters, which would appear to have an effect on other assortments of data (Beringer & Hüllermeier, 2006).

The second direction will look at the problems of how to model trainee state, state transitions, and future state prediction. Simultaneously, the author and others will look at how to present delayed reward to the machine learning algorithms which make decisions based on trainee states. It is expected that these studies will involve Markov Chain models, Markov Decision Processes, and Partially Observable Markov Decision Processes. Experiments of what constitutes a reward will also be conducted.

ACKNOWLEDGEMENTS:

The lead author would like to thank the Army's Simulation and Training Technology Center and University of Central Florida for continued support and development. Additionally, he would like to thank gromgull, D.S. Blank, Aaron Garrett, and Terry Stewart for parts of the software or software basis used in the above detailed algorithms.

REFERENCES:

Aggarwal, C., and Yu, P. S. Framework for clustering uncertain data streams. In *IEEE International Conference on Data Engineering*, 2008.

Ancona, F., Ridella, S., Rovetta, S., and Zunino, R., "On the importance of sorting in Neural Gas training of vector quantizers," in Proc. Int. Conf. Neural Networks'97, Houston, TX, June 1997, vol. 3, pp. 1804–1808.

Arroyo, I., Cooper, D.G., Burleson, W., Woolf, B.P., Muldner, K. and Christopherson, R. (2009a) 'Emotion sensors go to school', *Proceedings of the International Conference on Artificial Intelligence in Education*, IOS Press, July, pp. 17–24.

Beringer, J., Hüllermeier E., Online clustering of parallel data streams, Data Knowl. Eng. 58 (2) (2006) 180 – 204.

Carpenter, G.A., & Grossberg, S. (1995). Adaptive resonance theory (ART). In M. Arbib (Ed.), The handbook of brain theory and neural networks. (pp. 79-82). Cambridge, M.A.: MIT press.

Gao, Y., Raine, A., Venables, P.H., Dawson, M.E., Mednick, S.A., 2010. The development of skin conductance fear conditioning in children from ages 3 to 8 years. Dev. Sci. 13 (1), 201 –212.

Guedalia, I. D., London, M., Werman, M.. 1998. An on-line agglomerative clustering method for non-stationary data. Neural Computation. Citeseer.

Holmstrom, J. (2002). Growing neural gas. Unpublished Master's Thesis, Uppsala University.

Hore P, Hall LO, Goldgof DB. A fuzzy C means variant for clustering evolving data streams. IEEE-SMC, 2007

Jain AK (2008) Data clustering: 50 years beyond k-means. In: ECML PKDD '08: Proceedings of the 2008 European Conference on Machine Learning and Knowledge Discovery in Databases - Part I, Springer-Verlag, Berlin, Heidelberg, pp 3–4

Masud M., Chen Q., Khan L., Aggarwal C., Gao J., Han J., Thuraisingham B., Proceeding of the 2010 IEEE International Conference on Data Mining, IEEE Computer Society, 2010

Papadimitriou, S., Brockwell, A., Faloutsos, Ch.. Adaptive, unsupervised stream mining. The VLDB Journal, 13(3), 2004.

Wolpert, D.H., Macready, W.G. (1997), "No Free Lunch Theorems for Optimization," *IEEE Transactions on Evolutionary Computation* 1, 67.

Zito, T., Wilbert, N., Wiskott, L., Berkes, P. (2009). Modular toolkit for Data Processing (MDP): a Python data processing frame work, Front. Neuroinform. (2008)

AUTHORS BIOGRAPHIES

Keith Brawner is a researcher for the Learning in Intelligent Tutoring Environments (LITE) Lab within the U. S. Army Research Laboratory, Simulation and Training Technology Center (ARL STTC). He has five years of experience within U.S. Army and Navy acquisition, development, and research agencies. He holds a M.S. in Computer Engineering with a focus on Intelligent Systems and Machine Learning from the University of Central Florida, and is currently in pursuit of a doctoral degree in the same field. The focus of his current research is in machine learning, adaptive training, trainee modeling, and knowledge transfer. Dr. Avelino Gonzalez is his advisor.

Dr. Avelino Gonzalez received his B.S. and M.S. degrees from the University of Miami, Coral Gables, FL, in 1973 and 1974, respectively, and his Ph.D. degree from the University of Pittsburgh, Pittsburgh, PA, in 1979, all in electrical engineering. He is a Professor in the Department of Electrical Engineering and Computer Science at the University of Central Florida, Orlando, specializing in artificial intelligence and simulation.