An Experimental Investigation of Gladun's Theory of Automatic Classification

Brown, F. M., Jacobs, G., and Snider, J. S.
Artificial Intelligence Laboratory
University of Kansas
Lawrence, Kansas, 66045
(785)864-7376
brown@eecs.ukans.edu

Jackson, A. G. & LeClair, S. R.
Materials Directorate
Wright Laboratory
Wright Patterson AFB, Dayton, Ohio
(513)255-8787
leclaisr@ml.wpafb.af.mil

Abstract

Gladun's Pyramidal Net Classification Theory is a novel theory of automatic classification based on the idea of a hierarchical system of concepts formed through competitive concept formation procedures based on neural recognition metaphors. This paper addresses various observed phenomena of Gladun's methods by describing and analyzing the application of the methods to a number of classification problems. The problems include the prediction of the values of a parabola, circle, and sphere.

1. Introduction

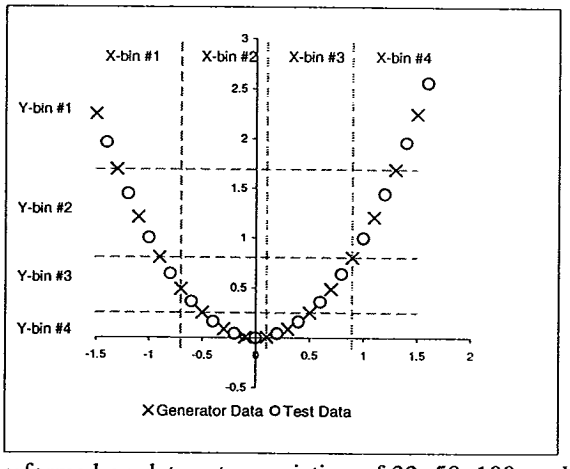
Gladun's Pyramidal Net Classification Theory [1,2,3,4,5,6,7,8,9,10,11,12,13] is a novel theory of automatic classification based on the idea of a hierarchical system of concepts formed through competitive concept formation procedures based on neural recognition metaphors. Our implementation of Gladun's theory works in several stages: creation of an input data set, partitioning of the data set into bins for the decision class (decision variable) and property values (attributes), preprocessing of the input data to eliminate redundancies, the Pyramidal Net Concept Formation process to create an Alpha Pyramidal Net, linking of input data to concepts, computation of classification values for each concept, the use of the training data set to choose appropriate concepts (called check nodes), and finally the generation of classification rules using the check nodes [16].

This paper addresses various observed phenomena of Gladun's methods by describing and analyzing the application of the above system to a number of classification problems. The first set of experiments (in section 2) deals with a parabola, the second set of experiments (in section 3) deals with a circle, the third set of experiments (in section 4) deals with a sphere. For each of these experiments, multiple input data sets were created using automated techniques and each was applied to the system with varying attribute bin selections. All of the experiments performed are presented in the appendices as cumulative data tables. References made to specific experiments are made throughout this paper. The experiments use the naming convention experiment set letter (j for parabola, c for circle, and s for sphere), followed by the total size of the data set, followed by a particular experiment letter. Thus, experiment j32a is a parabola experiment with 32 data points (16 generator points and 16 test points) designated letter a.

2. The Classification of Functional Phenomena: Parabola Experiments

The Data Sets

The first set of experiments attempted to predict the y values of a parabola $y = x^2$. The spreadsheets used generate the data sets for these experiments were set up such that every other data point is placed in the second half of the file so that it may be used to test the classification theory produced using the first half of the file (e.g. see figure). Experiments were



performed on data sets consisting of 32, 50, 100, and 1000 points. Perfect or near perfect classification results of 96% to 100% correctness were achieved for each size data set with 4 decision variable bins, when appropriate parameters were used. The decision variable was y and the other attributes available were x, quadrant, x-sign, and y-sign. Since the parabola data is entirely in the first two quadrants, the y-sign was the same for all data points and the quadrant and x-sign attributes were synonymous.

Appropriate Attribute(s)

Some attributes aid classification more than others. This often means that one or more attributes are needed for good classification, while others seem to make no difference whatsoever. In the case of the parabola experiments, it was found that the quadrant and y-sign attributes did not make a difference. However, the x attribute was found to be essential to good classification.

When the system was run on the 50 point data set, 92% correct classification was achieved when classifying y into 2 bins using only the x attribute, both with and without the use of the quadrant, x-sign, or y-sign attributes (see experiments j50a, j50e, j50g, and j50h).

More Decision Variable Bins

One of the more obvious points to make about classification is that it is harder to classify data into a larger number of decision variable bins. When only 2 decision variable bins are used, the system has no trouble coming up with 100% correct classification, as long as appropriate attributes are used. The difficulty of classification dramatically increases as the number of decision variable bins increases. There is always an upper limit on the number of decision variable bins that a given data set can reasonably classify. Smaller data sets, more complex data sets, and data sets with fewer appropriate attributes will have a smaller maximum number of decision variable bins.

Ratio of Attribute to Decision Variable Bins

In order to obtain good classification there needs to be a high enough ratio of attribute bins to decision variable bins. When this ratio is too low the algorithm does not have degrees of freedom with the bins available to perform good classification.

When the ratio of attribute bins to decision variable bins is below 2:1 the algorithm achieved around 50% correct classification, but when the ratio is above 2:1 the algorithm achieved better than 80% classification on all experimental runs with the parabola (except for j100b, which included the unnecessary quadrant and y-sign attributes; once those are discarded from the calculation the ratio drops to 1.5:1).

The picture above helps to illustrate the problem with a one-to-one ratio of bins. For the data points in any given x bin, there are 2 y bins that fit those data points. For instance, X-bin #1 has generator points in Y-bin #1 and Y-bin #2. Thus, the rules generated by the algorithm are forced to pick one or the other. This can be seen in experimental run j32c, where the following rules were produced, resulting in only 50% correct classification on 4 bins for y:

```
DETAILED RULES
CLASSES:  y=x2_*_0.25  y=x2_0.25_0.81
y=x2_0.81_1.69  y=x2_1.69_*
 0 0 2 2
(-> (and x_*_-0.7 x-sign_neg)
y=x2_0.81_1.69)
 0 0 2 2
(-> (and x_0.9_* x-sign_pos)
y=x2_0.81_1.69)
 2 2 0 0
(-> (and x-sign_pos (not x_0.9_*))
y=x2_*_0.25)
 2 2 0 0
(-> x_-0.7_0.1 y=x2_*_0.25)
```

Another illustration of the need for a better ratio of comes from the 1000 point parabola, runs j1kc and j1kd. In these two cases the decision variable, y, was given 4 bins and x-sign was given 2 bins. The only difference between the two runs was the number of bins for x, 4 bins in j1kc and 8 bins in j1kd. The resulting classifications were 50% correct for j1kc and 99.6% correct for j1kd.

Enough Data per Bin

Increasing the ratio of decision variable bins to attribute bins helps up to the point where there is not enough data in each bin to make the bin useful. When there is not enough data per bin, the algorithm may not make good decisions about bin relationships. In the 50 point parabola case, increasing the number of x bins from 8 to 16, and thus the ratio of x bins to y bins from 2:1 to 4:1, resulted in an increase in correctness from 84% to 92%. However, when the number of x bins was further increased to 24 bins, the percent correct dropped back to 84%. At 24 bins for x there could be only 1 data point in most of the bins, since there were only 50 data points in the data set and half of these are used for generating the P-net. Thus, a further increase in x bins would not make sense. When the number of data points in the set was then increased to 100, still with 24 bins for x and 4 for y, the correctness increased to 88%. With 100 data points the number of x bins was raised to 32 and 100% correct classification resulted. With 100 data points and 32 bins for x, 100% correct classification was also achieved with 8 bins for y.

Boundary Value Problems

Since bins are created from actual data points, their dividing points, or boundary values, always correspond to actual attribute values. Thus, when the data attributes are classified into bins some of the data sits on a bin boundary. The algorithm is deterministic in its classification, so each datum is put into a single bin. Sometimes a bin gets some data that might be better suited to an adjacent bin. Another manifestation of this problem is when two data points with attributes on bin boundaries are treated as identical points, so that one of them is

discarded and the other is chosen as the representative to make the rule. Since these points lie on a boundary between bins, it is possible that one of them should be in each adjacent bin. The decision variable attribute of these points is ignored while P-net nodes are created, even though it may reside in different decision variable bins. Thus, the information needed to distinguish between the two points is lost and incorrect classification results.

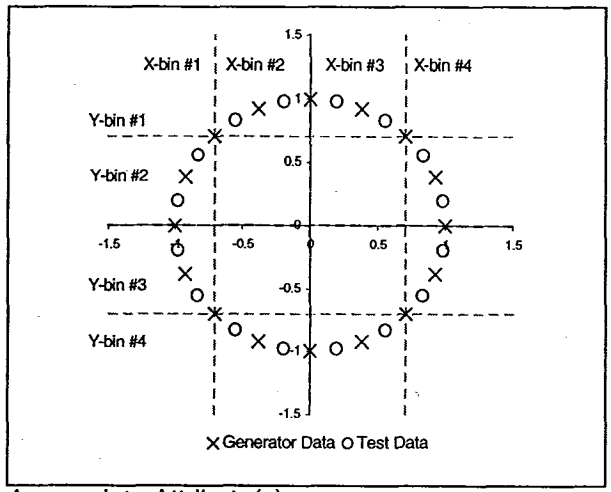
If this happens infrequently, the algorithm may still work out a near optimal classification. However, when this problem occurs repeatedly the algorithm may misclassify a large amount of data. The problem has a greater chance of disturbing the results when a large number of bins are created, since each bin introduces boundary values (albeit the outermost bins will only have one boundary value from the data set, since those bins are infinite). The problem may have a more pronounced harmful effect when there are already a small potential number of data points per bin (such as a small data set with a large number of bins), since this may result in some bins being completely ignored or misused.

The most commonly noticed effect of the boundary value problem is the slightly less than optimal classification of some data sets. This is very noticeable in the results of the parabola experiments; for instance experiments j50d, j100a, j100e, j1kb, j1kd, and j1kh all have over 95% correct classification and some of them differ very slightly in parameters. Many of these experiments were expected to produce 100% correct classifications, but due to boundary value problems they were all trivially sub-optimal.

3. The Classification of Relational Phenomena: Circle Experiments

The Data Sets

The second set of experiments was aimed at trying to predict the y values of a unit radius circle, $x^2 + y^2 = 1$. The spreadsheets used for these experiments were again set up such that every other data point is placed in the second half of the file so that it may be used to test the classification theory produced using the first half of the file (see figure). Experiments were performed on data sets consisting of 32, 64, 128, and 1024 points. Perfect classification results (100% correctness) were achieved for each size data set, when appropriate parameters were used. The number of decision variable bins used was 2, 4, 8, 16, and 32, not all of which were used on all data sets. The decision variable was y and the other attributes available were x, x^2, y^2, quadrant, and hemisphere. The y^2 attribute is not used often, since it is closely related to y and thus reduces the problem to classifying a different function.



Appropriate Attribute(s)

As with the parabola experiments, the circle experiments revealed certain attributes to be essential for good classification. In particular, the quadrant attribute was needed for all classifications and the x attribute was needed for runs with more than 2 decision variable bins.

In the 32 data point circle experiments with 2 decision variable bins it was seen that the omission of the quadrant attribute resulted in significantly lower correct classification and that it was the only attribute needed (see experiments c32a-d). However, when classifying y into 4 bins the quadrant attribute by itself resulted in only 50% correct classification (see experiment c32e). By adding the x attribute with a sufficient number of bins, classification rose to 87.5% (see experiments c32f and c32g).

It was also found that some attributes make no difference when added to the computation. Experiments c64g and c64h both resulted in 100% correct classification and c128f and c128g both achieved 87.5% correct classification. Experiments c64g and c128f included attributes y^2 and hemisphere, whereas the other two did not but were otherwise identical. Thus we see that the attributes y^2 and hemisphere do not contribute additional useful information to the process.

Ratio of Attribute to Decision Variable Bins

When there are 4 or more decision variable bins the ratio of attribute to decision variable bins must be above 1:1 to classify more than 50% correct and above 1.5:1 to classify more than 80% correct (see the cumulative circle experiment data).

The figure above helps to explain the problem with too few attribute bins. For the data points in any given x bin, there are 2 y bins that fit those data points. For instance, X-bin #1 has generator points in Y-bin #2 and Y-bin #3. Thus, the rules generated by the algorithm are forced to pick one or the other. The problem can be alleviated by introducing 4 bins for

the quadrant attribute. The quadrant then serves to distinguish between generator points in precisely the manner needed to correctly classify the data. This also explains why quadrant is a necessary attribute.

Enough Data per Bin

It was again found that even with a high ratio of attribute to decision variable bins, a lack of data points per bin results in poor classification. This is well illustrated by experiments c32i and c128a where the same attribute bins were used, but the 32 point data set (ratio of 4:1) simply did not have enough data for the number of bins used. When the number of data points was increased from 32 to 128 the correctness jumped from 56.25% to 100%. A similar result was noticed in experiments c128i and c1ka, where 32 attribute bins were used. The 128 point data set with a ratio of 1.63:1 resulted in 29.69% correct classification. The 1024 point data set with the same ratio resulted in 100% correct classification.

Boundary Problems

As seen with the parabola, boundary values again present a problem. Experiments c64e, c64f, and c64h show that we can compensate for the boundary value problem by introducing additional attributes to help distinguish between and thus prevent the merging of data points. Experiment c64e attempted to classify y into 16 bins using 16 bins for x and 4 bins for quadrant, resulting in just 50% correct classification. Experiment c64f added 4 bins for x^2 to those of c64e, resulting in 81.25% correct classification. By increasing the number of bins for x^2 to 8, c64h succeeded in achieving 100% correct classification.

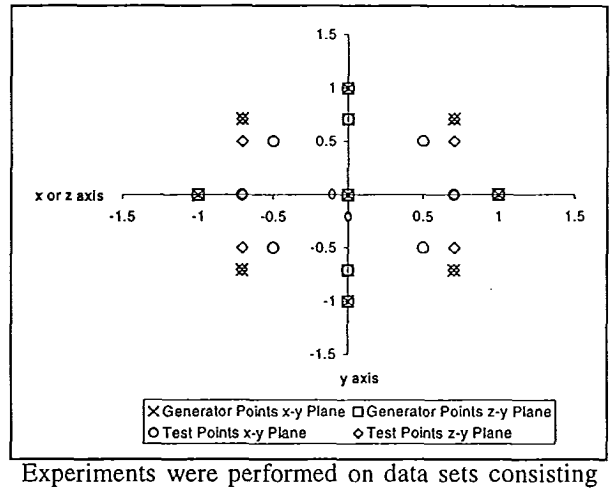
Classification is Harder with More Bins

Classifying the decision variable into a larger number of bins is again seen to be more difficult. However, by addressing the aforementioned problems 100% correct classification was achieved. In order to correctly classify y into 4 bins, more than 32 data points were needed, we used 64. This allowed the use of more attribute bins without sacrificing the number of data points per bin; thus addressing the problem of having a large enough ratio of attribute bins to decision variable bins, while maintaining enough data per bin, and compensating for the boundary value problem.

4. The Classification of Relational Phenomena with Multiple Independent Variables: Sphere Experiments

The Data Sets

The third set of experiments tried to predict the y values of a unit radius sphere, $x^2 + y^2 + z^2 = 1$. The spreadsheets used for these experiments were again set up such that every other data point is placed in the second half of the file so that it may be used to test the classification theory produced using the first half of the file. The data sets were constructed by sweeping a unit vector from the origin to a point on the sphere through various angles with respect to the x-axis and z-axis, thus producing points along longitudinal lines. The figure below illustrates the 32 point data set and was made by first projecting the sphere onto the x-y plane and then superimposing the sphere projected onto the z-y plane.



Experiments were performed on data sets consisting of 32, 64, 128, 1024, 2048, and 4096 points. Perfect classification results were achieved only for the 128 point data set with just 2 decision variable bins. Less than 75% correct classification was achieved for all other data sets and bin cardinalities less than 2048 points. The largest two data sets were able to achieve around 90% correct classification. Only 2, 4, and 8 decision variable bins were attempted. The decision variable was y and the other attributes available were x, x^2, y^2, z, z^2, x-y quadrant, x-y-z octant, x-sign, y-sign, z-sign, theta (angle from x-axis), and phi (angle from z-axis). Again, the y^2 attribute is not used often because of its close relation to y. Although sign information is inherently encoded into the quadrant and octant attributes, the independence of the sign attributes allowed them to contribute more to the classification in many experiments. It should be noted that the theta and phi attributes were used to create the data set and thus should viewed with skepticism as real attributes.

Appropriate Attribute(s)

As noticed with the parabola and circle, certain attributes contribute to correctness more than others. In the case of the sphere, the sign, x, z, and phi attributes proved most important. However, recall that the phi attribute should not be considered a true data attribute since it was used to construct the data sets, and in fact in a manner very fortuitous to its initial data bin assignments.

The best classification results for the sphere data sets with 4 decision variable bins included exactly x, x-

sign, y-sign, z, and z-sign. When other attributes were added or these attributes were removed, the resulting classification was either the same or worse in all such cases (e.g. experiments s2kb vs. s2kf, s2kh vs. s2kj, s128e vs. s128f vs. s128g).

Some Attributes Can be Misleading

Just as certain attributes can contribute to correctness more than others, an attribute can also have a negative impact on classification. In fact, sometimes an attribute will be misleading to the algorithm and cause a severe decrease in correctness.

The detrimental effect of some attributes was seen in experiments s128r and s128s, where the only difference between the two runs was the inclusion of 8 bins for the theta attribute. The introduction of this attribute resulted in the classification correctness dropping from 67.19% to 53.13%. Another manifestation of misleading attributes is when a bad bin distribution is attempted. This was seen in experiments where the sign attributes were changed from the usual 2 bins to 3 bins and all other attribute bins were held the same. The 2 bins per sign attribute experiment (s128f) resulted in one of the best runs for the 128 point data set at 70.31% correct classification; while the 3 bins per sign attribute experiment (s128h) correctly classified only 51.56% of the data.

Ratio of Attribute to Decision Variable Bins

For the sphere experiments with 2 decision variable bins a clear pattern of better results for a higher ratio was seen. In experiment s128a the x and z attributes were given 4 bins each and the octant attribute was given 8 bins, thus producing a bin ratio of 8:1 and resulting in 95.31% classification, with the remaining 4.69% being *incorrectly* classified. When the bins for the x and z attributes were increased to 8 each in experiment s128b (ratio of 12:1) the classification was still 95.31% correct, but the remaining points were simply *not classified*, rather than being misclassified. When the additional attributes of x^2, y^2 and z^2 were added with 8 bins each for experiment s128c (ratio of 26:1), 100% correct classification resulted.

When attempting to classify into 4 decision variable bins, the ratio of attribute to decision variable bins seemed to be far less important than for previous experiments. However, it is still observable that this ratio plays a role in classification for the sphere experiments. In fact, the best results obtained all show a ratio of 5:1 or greater, whereas the worse results frequently have a lower ratio (e.g. s2ke vs. s2kg and s128f vs. s128o).

The previous explanation of the problem with too few attribute bins for the circle experiments carries over for the sphere experiments, with the added problem of two rather than one independent dimension variable. For the data points in any given x bin, there are 2 y bins that fit those data points. Likewise for data points in any given z bin. Thus, the rules generated by the algorithm are forced to pick one or the other x bins and one or the other z bins. The problem can be alleviated by introducing 8 bins for the octant attribute, but the relationships are too complex for this to provide a complete fix (unlike the quadrant for the circle). The sign attributes sometimes work better than octant, though this is due to the nature of the classification algorithm needing more degrees of freedom to work, not an inherent advantage of the attributes themselves. In many cases the octant attribute or sign attributes are interchangeable with negligible differences in the resulting classification.

Enough Data per Bin

The need for enough data per bin proved to be an overwhelming concern for the sphere data sets. Due to the complexity of the attribute relationships, both a high ratio of attribute to decision variable bins and enough data per bin to deal boundary value problems were simultaneously needed. This forced a substantial increase in the size of the data set.

The attempts at classification on smaller data sets were not very successful. Trying to classify y into just 2 bins, the 32 point set achieved only 25% correct classification, while the 64 point data set reached only 81.25% (experiments s32a and s64a). Considering a 50% probability of randomly choosing the correct bin, these results are fairly poor. When the data points were increased to 128, the system was able to achieve 100% correct classification for the 2 decision variable case (see s128c).

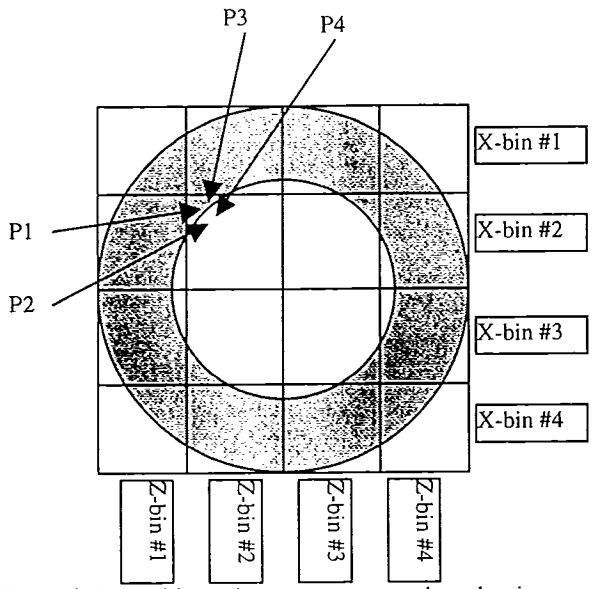
For the experiments with 4 decision variable bins only about 70% correctness was attained with the 128 point and 1024 point data sets (see s128f and s1kg). Although this is significantly better than the probability of random classification of 25%, it is still not as good as the previous experiment sets. By increasing the data set size to 2048 points, over 90% correct classification was accomplished (see s2kb). The 4096 point data set managed to classify 79.93% correct on the 8 decision variable bins problem (see s4kc). With an expected value via random classification of just 12.5%, this is a remarkable result. It is also impressive compared to the 44.53% correct classification obtained using the 1024 point data set (see s1kc).

Boundary Problems

The boundary value problems exhibited in the parabola and circle experiments are tame in comparison to those experienced by the sphere experiments. Aside from the usual problems of (1) data points residing on bin boundaries and thus needing arbitrary and occasionally wrong bin

decisions and (2) multiple data points being merged due to identical attribute bin classing because they shared a boundary value for those attributes, the sphere data sets face the problem of large numbers of points sharing boundary values with the independent x and z attributes. This new complication allows a significant portion of the data set to fall prey to the previous two boundary value problems.

Consider the 4 bin case for the decision variable y. Any one of the y bins orthogonally projected onto the x-z plane will intersect some equal number of both x bins and z bins (refer to the figure). If there are too



few of these bins, then any two points having y values near the boundary of the projected y bin (the gray area of the figure), but one inside (point P1) and the other outside (point P2), could easily have the same x bin value (X-bin #2) and z bin value (Z-bin #2). This may happen several times (e.g. points P3 and P4). The result is that all four points are merged due to them sharing identical attribute bin classifications. Thus, either P1 and P3 are misclassified or P2 and P4 are misclassified. To make matters worse, P1 and P3 may actually belong to different y bins, from above and below the x-z plane, which project to the same area of the x-z plane and likewise for P2 and P4. In such a situation four points that should be classified into four different y bins are merged into a single point prior to rule generation. This may result in very poor classification (only 25% of the points in our example will be correctly classified).

More Data to Classify Can Make Classification Harder

Although increasing the size of the data set tends to aid classification in general, having more data to classify sometimes results in slightly worse classification. The trade off between having more data to classify and helping the boundary value problem, enough data per bin problem, and subsequently the ratio of attribute to decision variable bins problem usually leans toward the "more is better" size data sets.

The most obvious indication of the problem with more data comes from the 2k versus 4k experiments. The best classification for the 2k data set with 4 decision variable bins was 91.02% correctness, whereas the 4k data set resulted in 89.11% correctness for the best experiment run. Although this difference is minor, especially considering an expected random classification success rate of only 25%, the difference is still noticed and attributable to the larger amount of data (see s2kb and s4kb).

Sphere Results

When the following conditions hold:

- the sign attributes have 2 bins,
- the x and z attributes have at least 4 bins,
- there are at least 8 data points per bin of the attribute with the most bins, and
- the ratio of attribute bins to decision variable bins is 5:1 or greater,

at least 67% correct classification was achieved for the sphere experiments. When the above factors hold but the ratio of attribute bins to decision variable bins is 8:1 or greater, at least 85% correct classification was achieved. When the above factors hold but number of data points per attribute with the most bins is less than 8, no better than 67% correct classification was achieved. When the above factors hold but the ratio of attribute bins to decision variable bins is less than 8:1, only one experiment attained above 70% correct classification.

Of the 16 experiments with 80% or better classification:

- only the 64 and 128 point data sets with 2 decision variable bins and two of the 2048 data sets do not use the sign attributes, but all 6 of these use the x-y-z octant attribute,
- all use the x and z attributes with at least 4 bins,
- only the 64 point data set with 2 decision variable bins has fewer than 8 data points per bin for the attribute with the most bins, and
- all have a ratio of attribute bins to decision variable bins of 5:1 or greater.

Of the 16 experiments with 60% or worse classification:

- 7 do not use the sign attributes,
- 5 do not use the x and z attributes,
- 8 have fewer than 8 data points per bin for the attribute with the most bins, and
- 7 have a ratio of attribute bins to decision variable bins less than 8:1.

5. Conclusions

Gladun's theory of classification produces excellent results on simpler problems and much better than chance results on more complicated problems.

List of Potential Problems with Classification

Several problems can arise and cause classification complications; however, most of these problems can be dealt with in some manner. What follows is a set of guidelines to help deal with these problems.

1. Perhaps the most important item to classification is to use appropriate attributes. When a necessary attribute is left out, poor classification will result regardless of other considerations. Likewise, the "shotgun approach" of including every attribute available may result in poor classification due to inclusion of a misleading attribute.
2. Although obvious, it is still important to note that increasing the number of decision variable bins will invariably lead to more difficult classification.
3. In order to provide the algorithm with enough latitude in its attribute bin decisions, it is imperative to provide enough attribute bins with respect to the number of decision variable bins. The necessary ratio between the two may vary from problem to problem, but there will certainly need to be more attribute bins than decision variable bins for any interesting data set.
4. Simply increasing the number of attribute bins in order to increase the aforesaid ratio will not help if there is too little data available for each bin. Having bins with only 1 data point in them nearly voids the entire purpose of having bins to begin with and increases the inherent problem with boundary values. When too little data is supplied, the system must make arbitrary choices about data bins, whereas more data allows the "preponderance of data" effect to be applied, where the choices are made based on the majority of values fitting a certain concept. This also reduces the relative number of boundary values compared to the whole of the data set, so that when an incorrect bin decision is made for a boundary value, the overall impact is minimized.

More Research and System Changes Needed

The above investigative results imply (not surprisingly) that the decisions made about bins for data are of great importance to correct classification. Since the partitioning of data into bins is so important, further research into the area is warranted. Several decisions about the bin related portions of the Gladun P-Net system were made before a good understanding of their impact could be gleaned. The bin creation technique and subsequent use of bins and check nodes in the algorithm might provide better results if they were more sophisticated.

Acknowledgements

This research was supported by the Materials Directorate of Wright Laboratory of the US Air Force under the Advanced Materials Processing Initiatives Contract # F33615-96-D-5835, delivery order 1109.

References

[1]Brown, Frank M., Alternative Design Deduction System for Concept Formation, Final Report F33615-90-C-5944 for USAF Materials Laboratory at Wright Research, AI Lab, University of Kansas, 1994.

[2]Gladun, Viktor P., "Growing Semantic Networks in Adaptive Problem Solving Systems, V.M. Glushkov Institute of Cybernetics, Academy of Sciences of the Ukrainian SSR.

[3]Gladun, V.P., *Heuristic Search in Complex Environments,* Evristichskiy Poisk v Slozhnykh Sredakh, Nauka Dumka Publishing House, 1977.

[4]Gladun, V.P., *Protsessy Formirovaniya Novykh Znaniy,* (i.e. *Processes of the Formation of New Knowledge*), Sofiya 1994, pp. 2-163, 166-192.

[5]Gladun, Viktor P., *Planirovaniye Reshenij* (i.e. *Planning Solutions*), Akademiya Nauk Ukrainskoy SSR Institut Kibernetiki im V.M. Glushkova, Nauka Dumka Publishing House, Kiev, 1987.

[6]Gladun, V.P. and Rabinovich, Z.L., "Formation of the World Model in Artificial Intelligence Systems", Machine Intelligence 9, Ellis Horwood Ltd., Chichester 19809, pp. 299-309.

[7]Gladun, V.P., Vashenko, N.D., Kiselyova, N.N., Javorsky, A.L., "Development of Computer Methods for Prediction of New Materials Having Predefined Properties", SPC-95-4016, 1995.

[8]Jackson, A.G., "Discovery Methods for Materials Process Design" Internal Report, Materials Lab., July, 1994.

[9]Kiselyova, N.N., "Principles of the Computer Design of New Inorganic Compounds", A.A. Baikev Institute of Metallurgy, Russian Academy of Sciences 117911 GSP-1, Moscow, Leninski Prospect, 19.

[10]LeClair, Steven R. and Jackson, A.G., "Window-on Europe", Materials Process Design (WL/MLIM), Materials Directorate, Wright Laboratory, Wright Patterson Air Force Base, Ohio 45433-6533, 1994.

[11]Vaschenko, N., "Operations on Concepts and their Realization when Knowledge Formation", Gorsistemotechnika Scientific Industrial Centre, Krasnoarmejskaya, 23b. Kiev, 252004, USSR.

[12]Vaschenko, N. "CONFOR: A Tool for Data Generalization and Prediction of New Materials", Inst. of Applied Informatics, Ukrainian Assoc. of Developers and Users of Intelligent Systems, 1995.

[13]Brown, Frank M., et al., "A Description and Analysis of Gladun's Pyramidal Net Classification Theory", AI Lab Report 98-2, AI Lab, University of Kansas, Lawrence, Kansas 66045, 1998.

Experiment Data (non-referenced experiments omitted to save space)

Experiment Name	Data Points	Dec Var. Bins	Attribute Bins				% Correct	# Data Points per Dec. Var. Bin	# Data Points per Bin of Attribute with Most Bins	Ratio of Attribute Bins to Dec. Var. Bins
		y	x	quad	x-sign	y-sign				
j32a	32	2	4	0	2	0	100.00	8.00	4.00	3.00
j32c	32	4	4	0	2	0	50.00	4.00	4.00	1.50
j50a	50	2	4	4	2	2	92.00	12.50	6.25	6.00
j50b	50	4	8	0	2	0	84.00	6.25	3.13	2.50
j50c	50	4	16	0	2	0	92.00	6.25	1.56	4.50
j50d	50	2	8	0	2	0	96.00	12.50	3.13	5.00
j50e	50	2	4	0	2	0	92.00	12.50	6.25	3.00
j50g	50	2	4	4	0	0	92.00	12.50	6.25	2.00
J50h	50	2	4	0	0	0	92.00	12.50	6.25	2.00
j100a	100	2	4	4	2	2	98.00	25.00	12.50	6.00
j100b	100	4	4	4	2	2	52.00	12.50	12.50	3.00
j100e	100	2	8	0	2	0	96.00	25.00	6.25	5.00
j100f	100	4	24	0	2	0	88.00	12.50	2.08	6.50
j100g	100	4	32	0	2	0	100.00	12.50	1.56	8.50
j100h	100	8	32	0	2	0	100.00	6.25	1.56	4.25
j1kb	1000	2	8	0	2	0	99.60	250.00	62.50	5.00
j1kc	1000	4	4	0	2	0	50.00	125.00	125.00	1.50
j1kd	1000	4	8	0	2	0	99.60	125.00	62.50	2.50
j1kh	1000	8	16	0	2	0	95.20	62.50	31.25	2.25

Experiment Name	Data Points	Dec. Var. Bins	Attribute Bins						% Correct	# Data Points per Dec. Var. Bin	# Data Points per Bin of Attribute with Most Bins	Ratio of Attribute Bins to Dec. Var. Bins
		y	x	x2	y2	quad	hemi	axis?				
c32a	32	2	4	0	0	4	0	0	100.00	8.00	4.00	4.00
c32b	32	2	2	0	0	2	0	0	100.00	8.00	8.00	2.00
c32c	32	2	2	0	0	0	0	0	50.00	8.00	8.00	1.00
c32d	32	2	0	0	0	2	0	0	100.00	8.00	8.00	1.00
c32e	32	4	0	0	0	4	0	0	50.00	4.00	4.00	1.00
c32f	32	4	2	0	0	2	0	0	50.00	4.00	8.00	1.00
c32g	32	4	4	0	0	4	0	0	87.50	4.00	4.00	2.00
c32h	32	4	8	0	0	4	0	0	87.50	4.00	2.00	3.00
c32i	32	4	8	4	0	4	0	0	56.25	4.00	2.00	4.00
c64a	64	4	4	0	0	4	0	0	100.00	8.00	8.00	2.00
c64e	64	8	4	0	0	4	0	0	50.00	4.00	8.00	1.00
c64f	64	8	8	0	0	4	0	0	81.25	4.00	4.00	1.50
c64g	64	8	8	4	4	4	2	0	100.00	4.00	4.00	2.75
c64h	64	8	8	4	0	4	0	0	100.00	4.00	4.00	2.00
c128a	128	8	8	4	0	4	0	0	100.00	8.00	8.00	2.00
c128f	128	16	16	4	0	4	0	0	87.50	4.00	4.00	1.50
c128g	128	16	16	4	4	4	2	0	87.50	4.00	4.00	1.88
c128i	128	32	32	16	0	4	0	0	29.69	2.00	2.00	1.63
c1ka	1024	32	32	16	0	4	0	0	100.00	16.00	16.00	1.63

Experiment Name	Data Points	Dec. Var. Bins	Attribute Bins												% Correct	Data Points per Dec. Var. Bin	Data Points per Bin of Attribute with Most Bins	Ratio of Attribute to Dec. Var. Bins
		y	x	x2	x-sign	y2	y-sign	z	z2	z-sign	[redacted]		xy quad	xyz oct				
s32a	32	2	4	0	0	0	0	4	0	0	[redacted]		0	8	25.00	8.00	2.00	8.00
s64a	64	2	4	0	0	0	0	4	0	0	[redacted]		0	8	81.25	16.00	4.00	8.00
s128a	128	2	4	0	0	0	0	4	0	0	[redacted]		0	8	95.31	32.00	8.00	8.00
s128b	128	2	8	0	0	0	0	8	0	0	[redacted]		0	8	95.31	32.00	8.00	12.00
s128c	128	2	8	8	0	8	0	8	8	0	[redacted]		4	8	100.00	32.00	8.00	26.00
s128e	128	4	8	0	2	0	2	8	0	2	[redacted]		0	8	70.31	16.00	8.00	7.50
s128f	128	4	8	0	2	0	2	8	0	2	[redacted]		0	0	70.31	16.00	8.00	5.50
s128g	128	4	8	0	0	0	0	8	0	0	[redacted]		0	8	65.63	16.00	8.00	6.00
s128h	128	4	8	0	3	0	3	8	0	3	[redacted]		0	0	51.56	16.00	8.00	6.25
s128o	128	4	8	0	0	0	2	8	0	0	[redacted]		0	0	65.63	16.00	8.00	4.50
s128r	128	4	16	0	2	0	2	8	0	2	[redacted]		0	0	67.19	16.00	4.00	7.50
s128s	128	4	16	0	2	0	2	8	0	2	[redacted]		0	0	53.13	16.00	4.00	9.50
s1kc	1024	8	16	8	0	8	0	16	8	0	[redacted]		4	8	44.53	64.00	32.00	8.50
s1kg	1024	4	16	0	2	0	2	16	0	2	[redacted]		0	0	70.12	128.00	32.00	9.50
s2kb	2048	4	16	0	2	0	2	16	0	2	[redacted]		0	0	91.02	256.00	64.00	9.50
s2ke	2048	4	16	0	2	0	2	16	0	2	[redacted]		0	8	91.02	256.00	64.00	11.50
s2kf	2048	4	16	0	0	0	0	16	0	0	[redacted]		0	8	89.26	256.00	64.00	10.00
s2kg	2048	4	0	0	2	0	2	0	0	2	[redacted]		0	0	53.13	256.00	512.00	1.50
s2kh	2048	4	12	8	2	8	2	12	8	2	[redacted]		0	0	85.45	256.00	85.33	13.50
s2kj	2048	4	12	0	0	0	0	12	0	0	[redacted]		0	8	90.82	256.00	85.33	8.00
s4kb	4096	4	32	0	2	0	2	32	0	2	[redacted]		0	0	89.11	512.00	64.00	17.50
s4kc	4096	8	32	0	2	0	2	32	0	2	[redacted]		0	0	79.93	256.00	64.00	8.75