J. E. Dammann

An Experiment in Cluster Detection

This study was designed as a practical test of the cluster finding and coding method described in an earlier paper.[1] In order to test such a method it is necessary to use data in which the clusters are known. The clusters for this experiment were generated by taking measurements on sustained spoken vowels. The utterance of each vowel was sampled at a 20-millisecond rate through 15 filters, and the outputs of the filters were quantized into 13 levels. Each time sample was taken as a pattern, thus producing a cluster of patterns from an utterance of one vowel. Eight vowels were used: |i|, |I|, |ɛ|, |*a*|, |ɔ|, |ɜ˙|, |o| and |u|. These are the vowel sounds found in the words "beet," "bit," "bet," "pot," "brought," "Bert," "boat" and "boot." From the eight classes a total of 250 patterns were generated. Figure 1 shows a typical sample from the data of each class in column 2 and two additional samples in columns 1 and 3 to give an indication of the divergence that occurred among the samples.

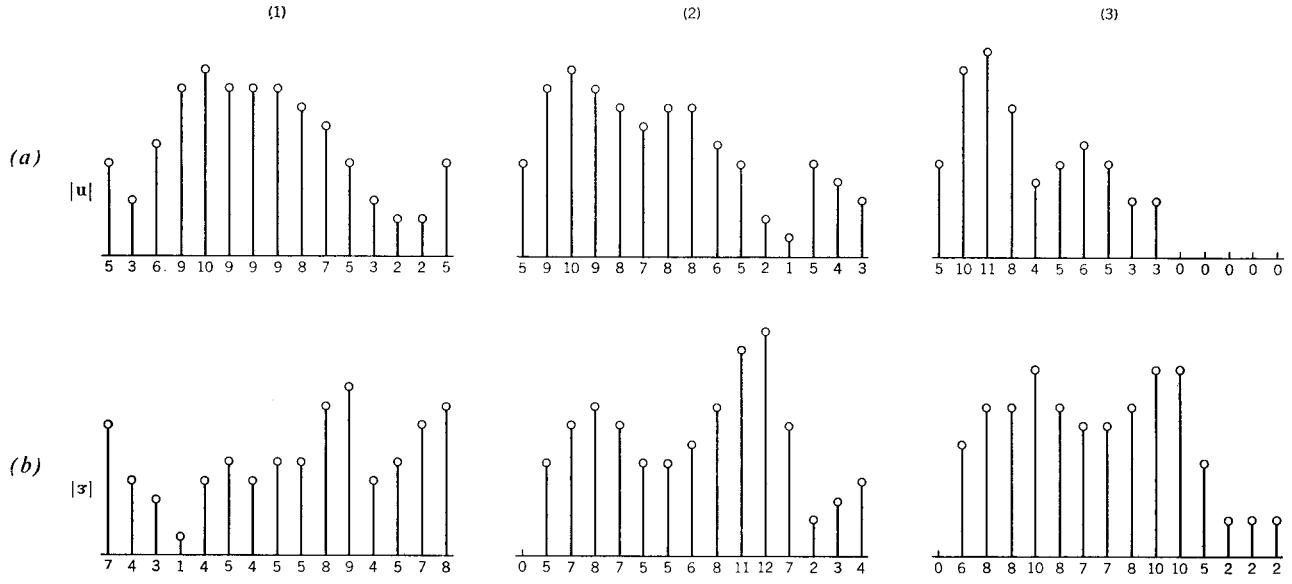
The set of 250 patterns, viewed as samples of an unknown space for purposes of this experiment, was subjected to a series of analyses to determine whether clusters existed and whether the method would separate them.

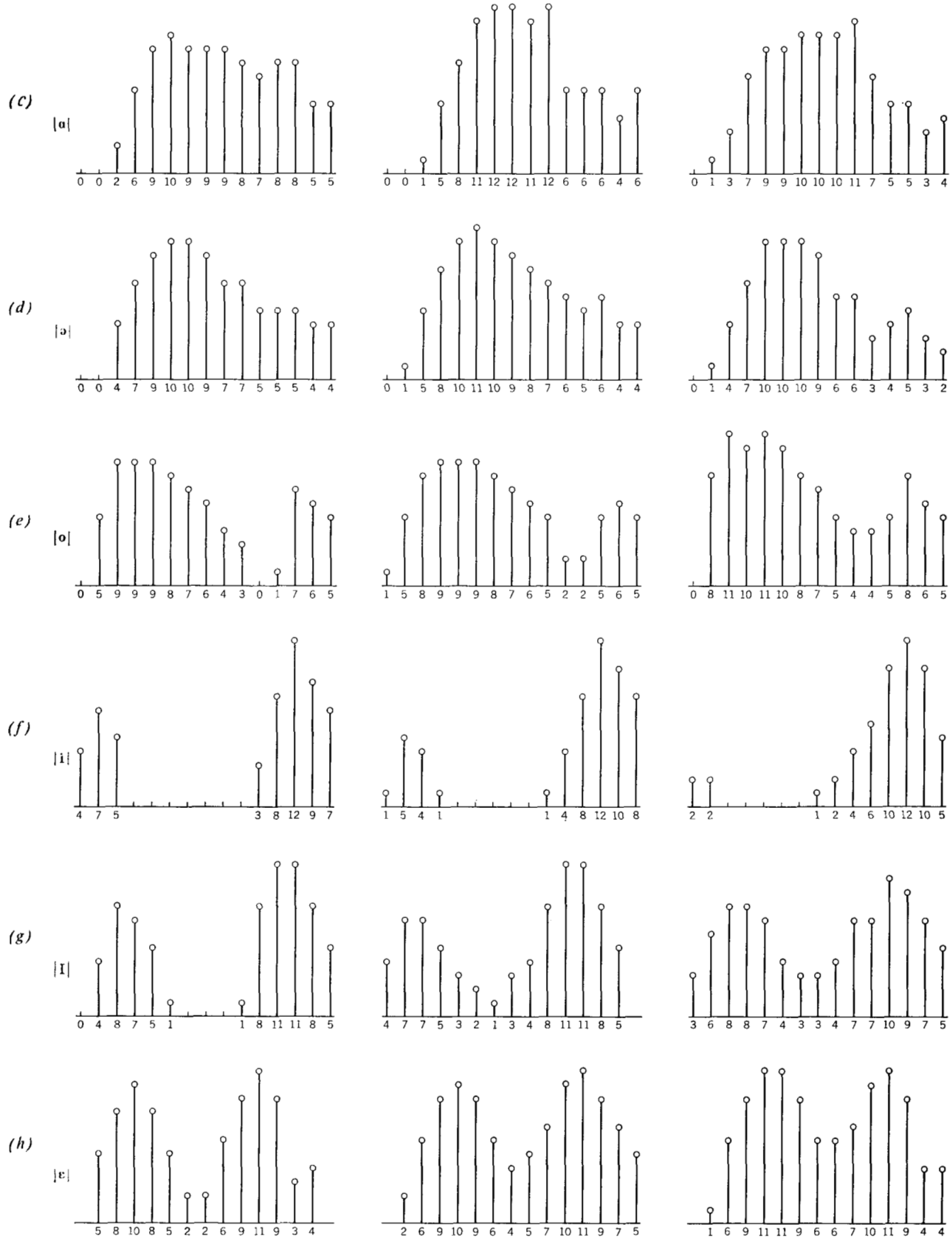
The first *S*-line distribution, formed by using all the patterns (level 1), is shown in Fig. 2.* A wide gap in the distribution will be noticed on the right side in the vicinity of the short bar intersecting the base line. Since this gap appeared to be the most significant in the distribution, a cut was made in it which is indicated by the vertical bar. Each side of the distribution was then analyzed separately in the second level of analysis. The results of these analyses are shown in Fig. 3.

Figure 3a shows the *S*-line for the small cluster isolated from the rest of the samples by the first level. It is not obvious from this distribution that only one class is represented, although the clustering of these samples on the first *S*-line is suggestive. It is a weakness of the present

* To aid in following the sequential separation, the approximate positions of the clusters throughout the analysis in Figs. 2-7 are identified by phonetic symbols under the *S*-line. A bar under a symbol at a particular level indicates that the associated class is first isolated at that level.

Figure 1 (a)-(h) Samples of the data taken from the eight classes of vowel sounds. A typical sample for each class appears in column 2; variations are shown in columns 1 and 3.




(c)
|a|
0 0 2 6 9 10 9 9 9 8 7 8 8 5 5
0 0 1 5 8 11 12 12 11 12 6 6 6 4 6
0 1 3 7 9 9 10 10 10 11 7 5 5 3 4
(d)
|ɔ|
0 0 4 7 9 10 10 9 7 7 5 5 5 4 4
0 1 5 8 10 11 10 9 8 7 6 5 6 4 4
0 1 4 7 10 10 10 9 6 6 3 4 5 3 2
(e)
|o|
0 5 9 9 9 8 7 6 4 3 0 1 7 6 5
1 5 8 9 9 9 8 7 6 5 2 2 5 6 5
0 8 11 10 11 10 8 7 5 4 4 5 8 6 5
(f)
|i|
4 7 5 3 8 12 9 7
1 5 4 1 1 4 8 12 10 8
2 2 1 2 4 6 10 12 10 5
(g)
|I|
0 4 8 7 5 1 1 8 11 11 8 5
4 7 7 5 3 2 1 3 4 8 11 11 8 5
3 6 8 8 7 4 3 3 4 7 7 10 9 7 5
(h)
|ɛ|
5 8 10 8 5 2 2 6 9 11 9 3 4
2 6 9 10 9 6 4 5 7 10 11 9 7 5
1 6 9 11 11 9 6 6 7 10 11 9 4 4

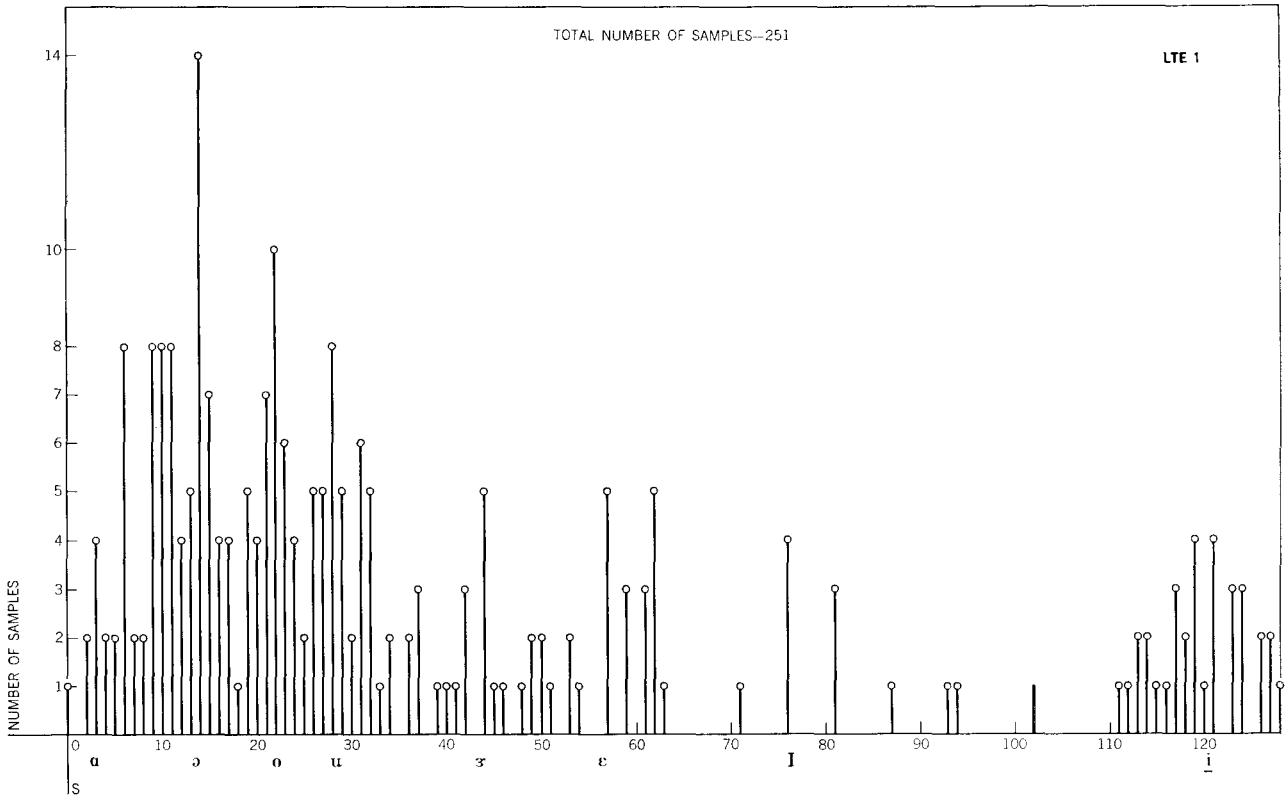


Figure 2 Level 1 of the *S*-line analysis.

experiment that not enough samples could be included to form a well-populated distribution when a class was isolated. A well-populated distribution at the level of Fig. 3a might have indicated more strongly that no sharp clustering is to be expected within the part of the sample space occupied by these points. A reference to the original data shows that 33 spectra generated by a speaker uttering a steady state |i| were included in the experiment and that these are in the 33 samples isolated by level 1.

Figure 3b shows a more complex situation than does Fig. 2. There is no gap which appears to be significant although the distribution is strongly bimodal. The separation of the *S*-line in this case is made by taking overlapping sections. All points to the left of the rightmost bar were taken as one group and the points to the right of the leftmost bar were taken as the other group. The area of overlap, the dip between the two modes, is the area between the two bars.

The third level of analysis and the distributions resulting from the two parts of Fig. 3b are shown in Fig. 4. Figure 4a shows a significant gap. The two parts of this distribution were separated as indicated by the bar. Figure 4b shows a wide gap with three samples near its center. These three samples were overlapped, with the bars placed in the widest gaps which occurred on either side.

Figure 5 shows the four distributions of level 4. Figure 5a shows a sparsely populated distribution of 31 samples, and the same considerations apply to considering it a terminal distribution as were discussed concerning the |i| samples. Of 32 samples from the original |*a*| class, 30 appear in Fig. 4a. One sample labeled |o| also occurs. This particular sample of |o| will be discussed subsequently.

Figure 5b exhibits wide separation with only six samples occurring to the left of the gap. Four of these samples are from the overlap of the distribution of Fig. 3b. All of these points subsequently group with a cluster derived from the upper part of Fig. 3b and these six points are therefore assigned to that cluster. The cluster of 27 points on the right end of the scale appear to be a cluster but are analyzed again in level 5. Separations are made in the largest gaps of Figs. 5c and 5d so that the analysis may proceed to that level.

Figure 6a shows 24 of the 25 original samples of |ɔ|.

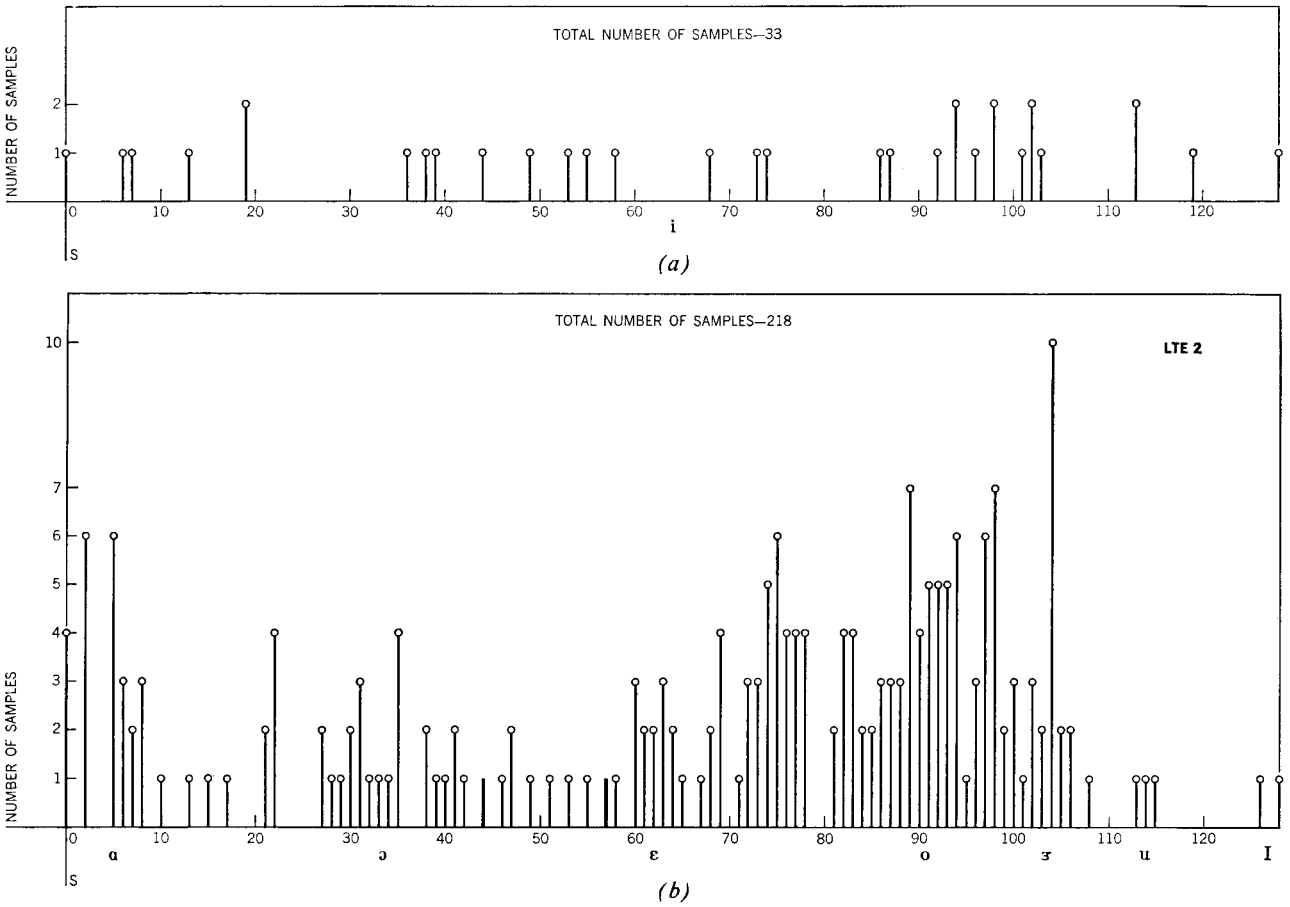


Figure 3 Level 2.

Though sparse, the distribution shows a certain bimodality. A well populated distribution might or might not show a bimodality indicating subclasses within this class. The distribution of Fig. 6b is divided at its largest gap. Figure 6c contains 33 of the original 34 samples of |ɜ˙|. It shows two renegade samples at the far left. One of these is the spectrum shown on the left in Fig. 1. The other exhibits an even stronger component in the two highest filters compared to the typical sample shown in the middle of Fig. 1.

The group of five samples at the left of Fig. 6d consists of three samples from the overlap of Fig. 3b (which group with the |ɔ|'s of Fig. 6a), one |ɜ˙| from the overlap of Fig. 4b (which groups with the |ɜ˙|'s of Fig. 6c) and one |ɔ| which was on the "wrong" side of Fig. 3b and not in the overlap region. The points on the right of Fig. 6d represent 30 of the original 31 |o|'s. It also contains three |u|'s which fell on the |o| side of the separation in Fig. 5d. The right side of Fig. 6d was not re-analyzed. The distribution of Fig. 6e is composed of 39 of the original 42 |u|'s. It also contains one sample of |ɜ˙|.

Figure 7a (level 6) shows 22 of the original 25 |ɛ|'s. The overlap points from Fig. 3b (also found in Fig. 5b) map to this distribution. Figure 7b shows 28 of the original 28 |I|'s. It also contains the |o| sample and the |ɛ| sample from the overlap of Fig. 4b, but these samples are not isolated in the distribution. It also contains the same sample of |ɜ˙| as found in Fig. 6e. This sample is isolated at the high end of the distribution.

From these results it is seen that five levels of analysis were necessary to separate the eight classes. Seven linear threshold elements are required to separate these classes sequentially in the manner developed by the analysis, if the provision is made that a pattern recognized simultaneously as |ɛ| and |ɔ| is |ɛ|. Furthermore, the analysis suggests an output code for parallel operation of seven elements and identifies aberrant samples relative to this code. If the output 1 is assigned to sums above the threshold (the right side of the *S*-line) and 0 to sums below the threshold, the suggested parallel code is as shown in Table 1. Comparison of Table 1 with the threshold element functions (LTE 1, etc.) in Figs. 2–7 will indicate the correspondence. The weights as derived from the analysis could not be used with this code, but conventional

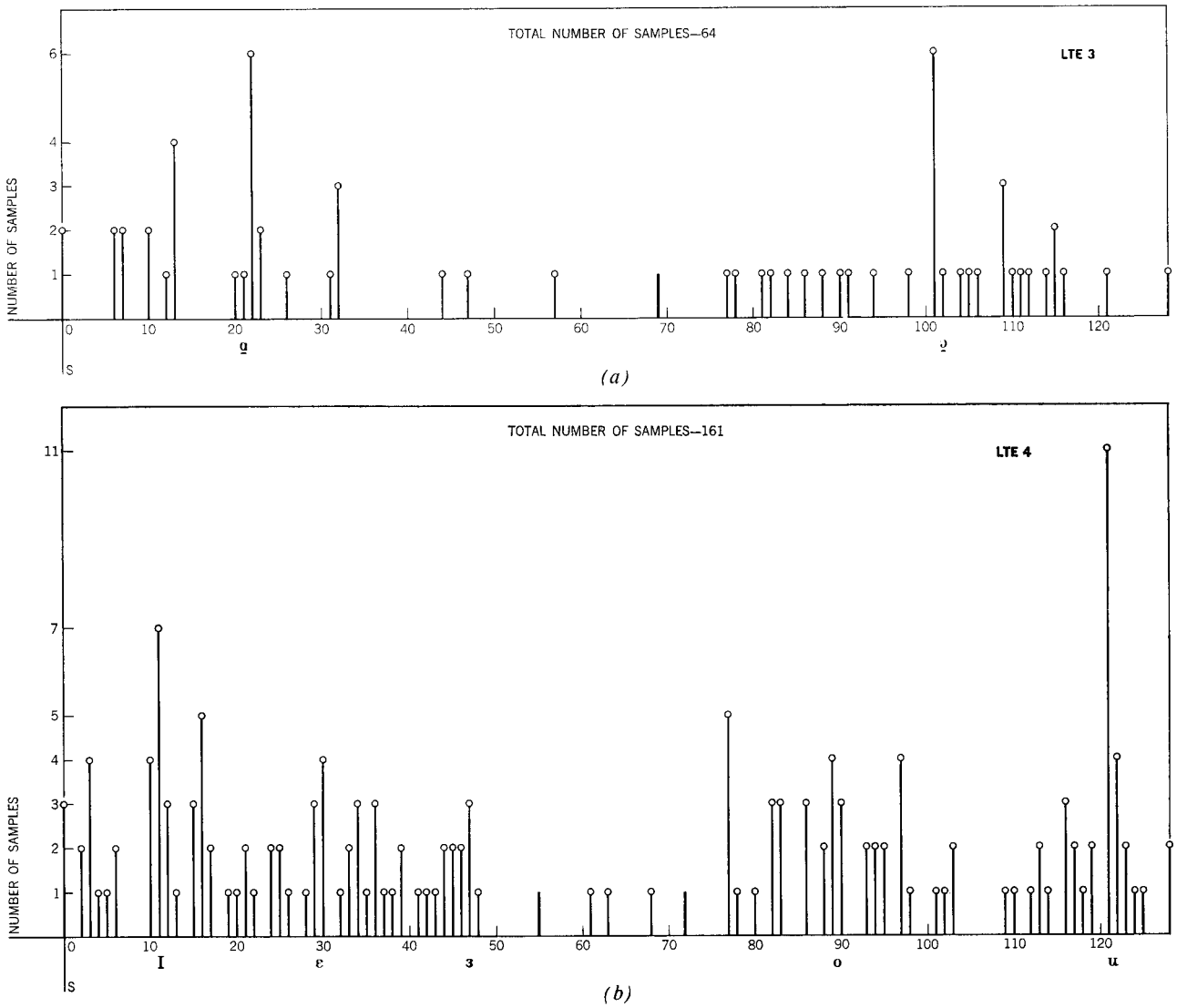


Figure 4 Level 3.

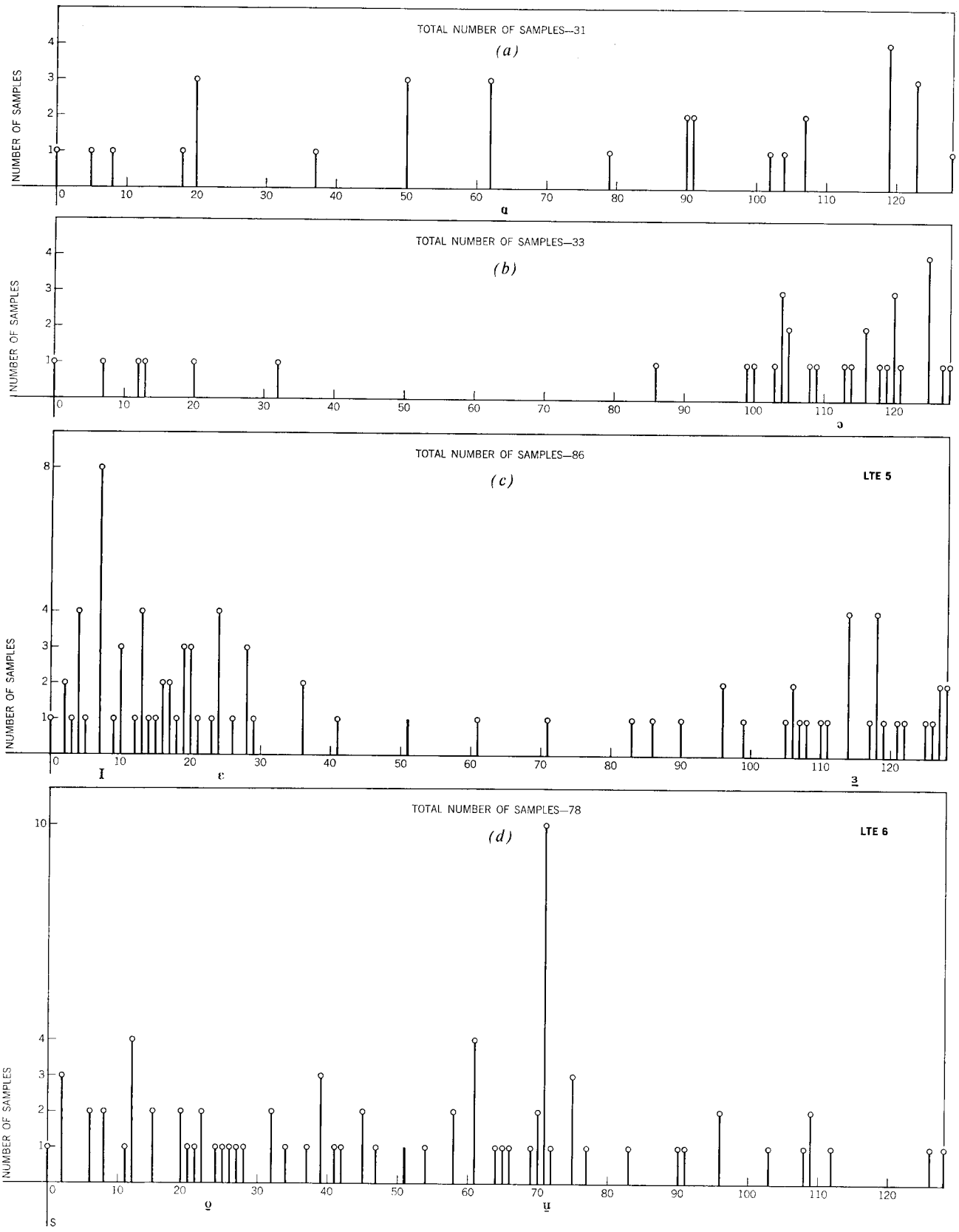
Figure 5 (Facing page) Level 4.

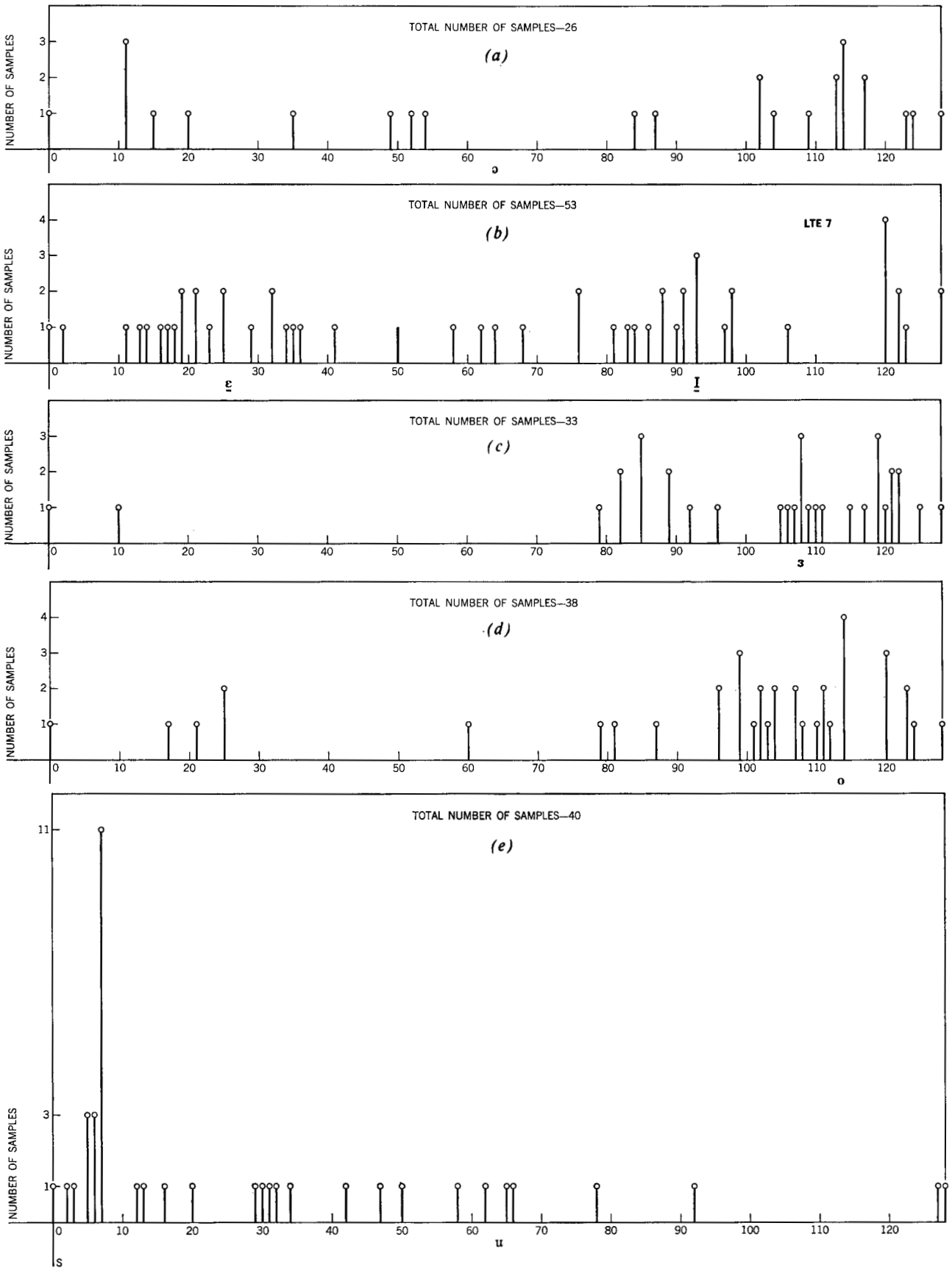
Table 1 A possible output code for the detected classes.

	Linear Threshold Elements						
Class	1	2	3	4	5	6	7
|i|	1	x	x	x	x	x	x
|I|	0	1	x	0	0	x	1
|ɛ|	0	1	x	0	0	x	0
|ɜ˙|	0	1	x	0	1	x	x
|a|	0	0	0	x	x	x	x
|ɔ|	0	0	1	x	x	x	x
|o|	0	1	x	1	x	0	x
|u|	0	1	x	1	x	1	x

training of adaptive threshold elements to this code should lead to good results.

An examination of the samples which appeared in "wrong" classes should be of value in evaluating the results of the experiment. For instance, one sample of |ɜ˙| did not find its way to the |ɜ˙| class at all. Instead it is found with |I|'s and the |u|'s. This spectrum is (a) in Fig. 8. It was the first sample of the |ɜ˙| utterance. One sample of |o| did not remain with the |o| cluster but was instead diverted to the |a| class. It was the first sample of an utterance of |o| and is shown in Fig. 8b. Another sample of |o| appeared with the |I|'s although it also remained with the |o| cluster. It was also the first sample of an |o| utterance. The first sample of the |ɛ| utterance


TOTAL NUMBER OF SAMPLES—31
(a)
NUMBER OF SAMPLES
TOTAL NUMBER OF SAMPLES—33
(b)
NUMBER OF SAMPLES
TOTAL NUMBER OF SAMPLES—86
(c)
LTE 5
NUMBER OF SAMPLES
TOTAL NUMBER OF SAMPLES—78
(d)
LTE 6
NUMBER OF SAMPLES


TOTAL NUMBER OF SAMPLES—26
(a)
NUMBER OF SAMPLES
ɔ
TOTAL NUMBER OF SAMPLES—53
(b)
LTE 7
NUMBER OF SAMPLES
ɛ
I
TOTAL NUMBER OF SAMPLES—33
(c)
NUMBER OF SAMPLES
ɜ
TOTAL NUMBER OF SAMPLES—38
(d)
NUMBER OF SAMPLES
o
TOTAL NUMBER OF SAMPLES—40
(e)
NUMBER OF SAMPLES
u
s

was classified as an |I| by level 5. This spectrum is shown in Fig. 8c. Samples occurring at the beginning of utterances are often found to be noisy because of the onset of vocal vibrations.

Thirteen samples altogether, 3 |u|, 3 |ɛ|, 2 |o|, 2 |.|, 2 |a| and 1 |ɔ|, were not classified unambiguously, that is, they did not end up in the proper class or were also represented with other classes. Two of these samples (one |o| and one |ɜ˙|) can be attributed to unresolved overlapping.

Percentage statements of accuracy (94.8 in this case) mean little in an experiment such as this, and the best evaluation is a subjective one performed by readers experienced in working with complex patterns. In a case such as the present experiment where the patterns emerge from the real world rather than from a pattern-generating rule, there can be no assurance that any particular sample belongs to its parent class in accordance with any reasonable objective criterion. It is probable, for example, that most observers will agree with the analysis in this experiment that indicated the |o|-generated pattern of Fig. 8b resembles the |a| class more than any other class.

In a typical adaptive system approach to this data the system would have been required to learn to identify the sample of Fig. 8b as |o| and to distinguish it from |a|, thus distorting the criteria for |o| -ness and |a| -ness.

In determining unknown classes the experimenter is faced with the task of sequentially separating the sample space into groups of patterns. The problems of where to make these separations and how long to continue to make

Figure 6 (page 86) Level 5.

Figure 7 (below) Level 6.

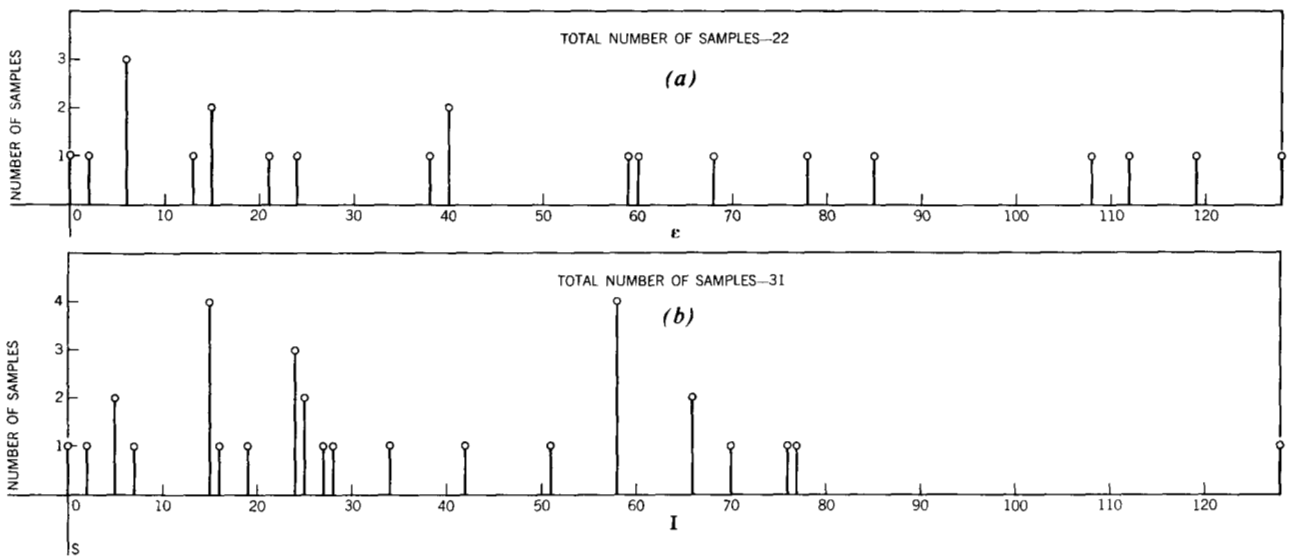
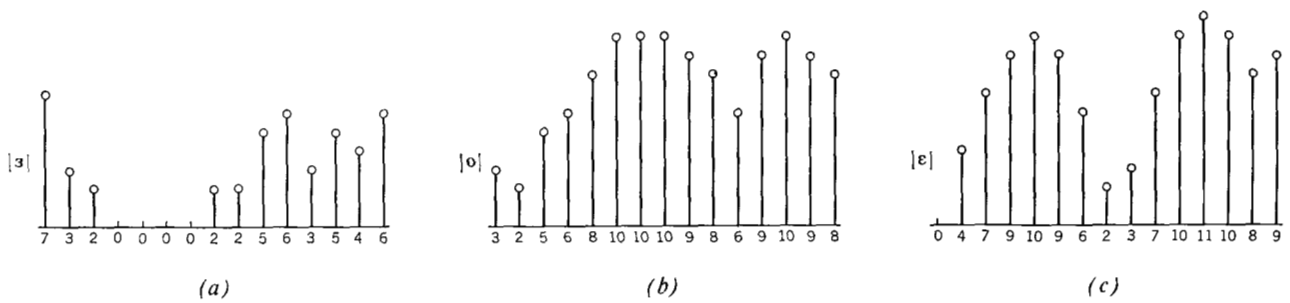


Figure 8 Three examples of "incorrectly" classified data.



separations both arise. The experiment described here is more successful in demonstrating the value of the *S*-line method for the first of these two problems than for the second. However, those groups of items which may be considered classes for one purpose may be subclasses for another. That is, the experimenter is ordinarily forced to observe the results of separations and make value judgments as to whether the points thus obtained should be considered to make up a class at the level of abstraction in which he is interested. In making such decisions, the primary value of the *S*-line method is that it will give an indication, for a selected group of points, of the extent to which they can be divided.

Within the limitations imposed by the number of points within the clusters of the sample space, this experiment resulted in the satisfactory separation of eight unknown clusters of complex patterns in an arbitrary sample space, identified aberrant samples, and suggested an output code.

Reference

1. "A Technique for Determining and Coding Subclasses in Pattern Recognition Problems," R. L. Mattson and J. E. Dammann, *IBM Journal* 9, 294 (1965).

Received August 26, 1965.