ence between the two methods (unless, as was apparently the case in the original Bledsoe-Browning work, a correlation method that happens to favor one over the other is chosen to assess similarity), but with five alphabets the 2-tuple is clearly superior to the 1-tuple.

There is a great need for stringent tests and comparative studies of different pattern recognition methods. But an experiment should make explicit what are the factors being varied, and lead to unambiguous statements as to the sources of effects demonstrated. In the present case, it would seem that Highleyman and Kamentsky have demonstrated the limitations of the basic 2-tuple method. But they have also demonstrated ways whereby it can be strikingly improved. The most important conclusion to be drawn from their replication would seem to be that if the 1-tuple method can be made to work so well, so easily, then it is to the larger n-tuple methods, which are guaranteed to work even better, that these improvements should be made.

LEONARD UHR
Mental Health Res. Inst.
University of Michigan
Ann Arbor, Mich.

Further Comments on the N-tuple Pattern Recognition Method*

The primary purpose of our original letter[1] was to dispel a false conclusion to which a reader might be led by Bledsoe and Browning's paper,[2] *i.e.*, that a machine based on $n=2$ is sufficient for the recognition of handprinting with an accuracy of 80 per cent or so. We considered these results to be somewhat misleading because their limited data source was not described in the paper. We are happy to note the greatly improved results which Bledsoe and Browning obtained with higher n when operating on our data, since we do feel that their method has merit when applied properly. The value of n required is quite important, however, since the complexity of the resulting machine, as measured by the number of memory cells required, increases almost exponentially with n.

With regard to Dr. Uhr's comments, I would like to make the following observations.

1) We chose the correlation method because we felt that it was based upon an easily understood technique. Such a technique would indicate to some extent the variability of the data to which it was applied.

2) The correlation technique which we used is not equivalent to Bledsoe and Browning's method for $n=1$, in which the memory matrix is comprised of the probabilities of occurrence of the various states. The difference lies in an appropriate normalization of the probabilities in the correlation technique such that the sum of the squares of the probabilities in a particular matrix is unity. Bledsoe and Browning simply added the unnormalized probabilities. It is a simple matter to construct examples for $n=1$ which show the need for proper normalization. For example, consider two patterns represented by a two-element matrix, as in Fig. 1(a). Assume that the noise characteristics are such that the unnormalized probability matrices are as shown in Fig. 1(b). Obviously, using these matrices, an ideal pattern A will always be classified as pattern B. However, if both matrices are normalized as described previously, shown in Fig. 1(c), the ideal patterns are classified correctly.

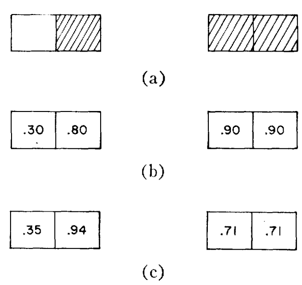

Fig. 1—Two-element matrix. (a) Pattern A, left, and pattern B, right; (b) unnormalized probability matrices; (c) normalized probability matrices.

3) We feel that the use of probabilities rather than binary weights would improve the method of Bledsoe and Browning; this is a point which they also made in their paper. In fact, we attempted recognition using 2-tuples where the memory matrix consisted of the (unnormalized) probabilities of state occurrences based on 50 samples of each hand-printed character (the same data as were used to construct the probability matrices for the correlation test). The recognition rate was improved from 19.7 per cent with the binary matrix to 30.7 per cent with the probability matrix. However, it can be argued that the need for proper normalization (as discussed above) is also existent for $n>1$. The problem of whether a meaningful normalization exists for these cases is yet to be studied. The normalization argument, incidentally, holds also for a matrix composed of binary weights.

4) Dr. Uhr's comment that the correlation technique retained more information than the method of Bledsoe and Browning in the case of a binary matrix is a good point. However, in the above experiment, the 2-tuple method retained as much information (in fact, more information, since the probabilities of pair states were retained) as the correlation method. Yet it still resulted in significantly poorer performance (30.7 per cent recognition rate vs 77.2 per cent), perhaps because of the lack of an appropriate normalization. We had also tried other random arrangements of pairs, with essentially the same results.

W. H. HIGHLEYMAN
Bell Telephone Labs., Inc.
Murray Hill, N. J.

* Received by the PGEC, November 30, 1960.

[1] W. H. Highleyman and L. A. Kamentsky, "Comments on a character recognition method of Bledsoe and Browning," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-9, p. 263; June, 1960.

[2] W. W. Bledsoe and I. Browning, "Pattern recognition and reading by machine," *Proc. EJCC*, pp. 225–232; December, 1959.

Computer Model of Gambling and Bluffing*

I wish to outline a project as yet not complete, which may be of some interest.

The machine simulation of human behavior in the mental states of uncertainty, such as estimation, prediction, choice, risk-taking, decision-making, makes more comprehensive these difficult conceptual and logical problems for the social scientist, psychologist, military strategist, etc.

Interesting studies can be pursued with digital computers on the playing of games.[1-8] An important subclass of games is the one in which the players make probability judgements, and can have hidden plans, etc.—in contrast to the games in which the information on the previous history and present position is perfect. Fairly exact experimentation would be possible with a poker-playing machine since here a human opponent's motivated responses are primarily controlled by simple numerical properties of the stimulus situation. Such a program may serve as a model of human gambling and bluffing in business competition, critical military situations, etc., by describing the objective vs subjective probability scales of conservative, mathematically fair (if any), and extravagant players. We could explain, for example, why and how gamblers characteristically overvalue long shots (low probability of high winnings) and undervalue short shots (high probability of low winnings).

A sketchy flow-chart of a poker program under construction can be seen on Fig. 1. The game is a variant of Draw-Poker, known as "open on anything." For the sake of simplicity, the step of paying the ante is left out; moreover, the opponent always makes the first bid. The steps are as follows:

1) Deal 5 cards for each, the machine and the opponent.

2) Calculate[9] the optimum number of cards to be exchanged by the machine at the second dealing n_{opt}; moreover, calculate the expected value of the probability of the machine's winning after the second dealing $E(p_2)=p_1$.

3) The opponent has bid M chips.[10]

* Received by the PGEC, October 10, 1960.

[1] A. Bernstein, *et al.*, "A chess playing program for the IBM 704," *Proc. WJCC*, Los Angeles, Calif., pp. 157–159; May, 1958.

[2] N. V. Findler, "Some remarks on the game 'Dama' which can be played on a digital computer," *Computer J.*, vol. 3, pp. 40–44; April, 1960.

[3] N. V. Findler, "Programming games," [Pt. (a) of Paper BI 3.3], Summarized Proc. of the First Conf. on Automatic Computing and Data Processing, Australia; May, 1960.

[4] J. Kister, *et al.*, "Experiments in chess," *J. Assoc. Computing Mach.*, vol. 4, pp. 174–177; April, 1957.

[5] A. Newell, "The chess machine," *Proc. WJCC*, Los Angeles, Calif., pp. 101–108; March, 1955.

[6] A. Newell, *et al.*, "Chess-playing programs and the problem of complexity," *IBM J. Res. & Dev.*, vol. 2, pp. 320–335; October, 1958.

[7] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM J. Res. & Dev.*, vol. 3, pp. 211–229; July, 1959.

[8] C. E. Shannon, "Programming a computer for playing chess," *Phil. Mag.* (7), vol. 41, pp. 256–275; March, 1950.

[9] Since, in the general case, when the whole stock of cards is played off before a new shuffling takes place, tabulated probabilities are obviously awkward and cumbersome, the Monte Carlo technique is to be used with the steps 2 and 8.

[10] The notations M and $M+N$ always represent the current value of chips in the pot, regardless of how many bidding cycles have lead to it.