

Correspondence

Fuzzy Sets and Decisionmaking Approaches in Vowel and Speaker Recognition

SANKAR K. PAL AND DWIJESH DUTTA MAJUMDER

Abstract—Some applications based on the theory of fuzzy sets in problems of computer recognition of vowels and identifying the person from his spoken words using only the first three formants (F_1 , F_2 , and F_3) of the unknown utterance are presented. Two decision algorithmic methods using weighted-distance functions and property sets are developed and implemented with the optimum size of the training set on a large number of Telugu (an important Indian language) speech sounds with a recognition score of 82 percent for vowels and 97 percent for the speaker.

I. INTRODUCTION

The problem of pattern recognition in general and that of speech recognition in particular are increasingly drawing the attention of scientific workers because of the potentiality for man-machine communication. The concept of pattern classification may be viewed as a partition of feature space or a mapping from feature space to decision space. There are various mathematical tools suggested by different authors [1]–[3] leading to the machine recognition of patterns. But more often than not, situations in the field of natural and social sciences are too complex for precise mathematical analysis. Classes of objects in that field do not have well-defined criteria of membership. To demonstrate such complexly behaved systems, the concepts of fuzzy sets and the subsequent developments in decision process [4]–[6] could be applied to a reasonable extent. This concept is approximate but provides an effective and more flexible basis for analysis of systems which are not precisely defined.

The problem of speech recognition has been dealt with by several researchers [7]–[12], [26] using time, frequency, or time-frequency (spectrograph) domain analysis. Use of computers in the speech recognition problem was first made by Forgie and Forgie [8] who recognized ten vowels with 93 percent accuracy. Denis and Mathews [12] indicated the advantages of computers for solving many of the problems encountered in speech research. The feasibility of automatic speaker identification has been demonstrated by many authors using various speech characteristics such as spectral data from filter banks [13], nasals [14], pitch contours [15], pitch intensity and formants [16], linear predictor coefficient (LPC) analysis [17], and zero crossings and amplitude measurement [18]. The method of formant analysis for speaker recognition was first adopted by Doddington [16] who had used Schafer and Rabiner's technique [11] to convert speech into pitch, intensity, and formant values. It has been found that higher order formants (4th and 5th) are more speaker dependent, but it is generally difficult to extract them. In spite of measurement difficulty and larger processing time, formants have potential application in speech and speaker recognition because of their remarkable inter-repetition stability [19] and their close relation to the phonetic concepts of segmentation and equivalence.

The present correspondence shows an application of fuzzy set theory in the field of artificial intelligence for machine recognition

of speech and informants using the first three formants only. The problem has a far-reaching significance in man-machine communication problems where simultaneous identification of speaker and speech sound is important. This is a part of the investigations [20]–[25] on man-machine communication research, a program under development in the ECS Laboratory, Indian Statistical Institute. Two methods for classification analysis on the basis of fuzzy algorithms are, first of all, described in this correspondence, and these are implemented to develop an automatic system to answer sequentially, with a minimum number of errors, the question, "What is the vowel contained in and who is the informant of an unknown utterance in a large number of spoken CNC (consonant-vowel nucleus-consonant) words?" The first part of the question requires the computation of weighted distance functions used in estimating the membership value for each vowel class. Second, a recognition method based on evaluating property sets and finding the similarity vectors corresponding to different classes were considered for identifying speakers. The Honeywell-400 electronic computer was used as the data processing system.

This program was carried out on a set of Telugu (one of the major Indian languages) words containing about 900 commonly used speech units for 10 vowels in CNC combination and uttered by three informants in the age group of 28–30 years. It is expected that variations in sex and age will lead to a better recognition score. As it is believed that formants have potential application in speech and speaker recognition because of their remarkable inter-repetition stability and their close relation to the phonetic concepts of segmentation and equivalence, the first three formants F_1 , F_2 , and F_3 only were considered as characterizing features from a spectrum analysis made on a Kay Sonagraph (model no. 7029A). It is to be noted that higher formants (4th and 5th) and fundamental voice frequency F_0 are found to be very speaker dependent and are not considered essential as recognition criteria. Since F_3 is more speaker dependent as compared to F_1 and F_2 , parameters selected for informant recognition are (F_3/F_1) or $(F_3 - F_1)$, (F_3/F_2) or $(F_3 - F_2)$, and F_3 . Since each of these is a function of F_3 , the inverse of standard deviation of the F_3 features is used as the weighting coefficient. The second part of the experiment consists in the study of the variation of the machine's performance in identifying a speaker with different sizes of training samples used in learning. Results are described in tabular form, where each score mentioned is the average value of three observations performed on a specified training set.

II. CLASSIFICATION ANALYSIS USING FUZZY SETS

A. Fuzzy Sets

A fuzzy set (A) in space of points $X = \{x\}$ is a class of events with a continuum of grades of membership and is characterized by a membership function $\mu_A(x)$ which associates with each point in X a real number in the interval $[0,1]$ with the value of $\mu_A(x)$ at x representing the grade of membership of x in A . Formally, a fuzzy set A with its finite number of supports x_1, x_2, \dots, x_n is defined as

$$A = \{(\mu_A(x_i), x_i)\} \quad (1)$$

or

$$A = \bigcup_i \mu_i/x_i, \quad i = 1, 2, \dots, n. \quad (2)$$

Manuscript received May 3, 1976; revised September 20, 1976, and March 8, 1977. This work was supported in part by the Council of Scientific and Industrial Research, Government of India.

The authors are with the Department of Electronics and Communication Sciences, Indian Statistical Institute, Calcutta 700 035, India.

The characteristic function denotes the degree to which an event x_i may be a member of A , and as it approaches unity, the grade of membership of x_i in A becomes higher.

B. Decisional Algorithms

Consider an unknown pattern

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ \vdots \\ x_N \end{bmatrix}$$

where x_n denotes the measured n th feature of the event, represented by a point in the multidimensional vector space Ω_X , consisting of m ill-defined pattern classes $C_1, C_2, \dots, C_j, \dots, C_m$. Let $R_1, R_2, \dots, R_j, \dots, R_m$ be the reference vectors where R_j associated with C_j contains h_j number of prototypes such that

$$R_j \ni R_j^{(l)}, \quad l = 1, 2, \dots, h_j$$

The pattern X can then be assigned to be a member of that class to which it shows maximum similarity as measured by the algorithms described below.

Method I: We define a membership function $\mu_j(X)$ associated with pattern X for the j th class as

$$\mu_j(X) = [1 + (d(X, R_j)/E)^F]^{-1.0} \quad (3)$$

where E is an arbitrary positive constant, F is any integer, $d(X, R_j)$ is the distance between X and R_j , and

$$d(X, R_j) = \bigwedge_l \left[\sum_n (W_{jn}^{(l)}(x_n - R_{jn}^{(l)}))^2 \right]^{0.5} \quad (4)$$

where \bigwedge denotes minimum. The above expression represents the minimum value of the weighted distances of an unknown pattern X from all its expected values in class C_j and $W_{jn}^{(l)}$ ($|W_{jn}^{(l)}| \leq 1$) corresponds to the l th prototype in C_j and denotes the magnitude of the weighting coefficient along the n th coordinate.

Constants E and F in (3) have the effect of altering the fuzziness of a set [5], [6], [22], [23]. The μ -value defining the grade of membership of X in C_j as shown by this equation is unity for $d(X, R_j) = 0$, zero for $d(X, R_j) = \infty$, and increases with decreasing the value of $d(X, R_j)$. Thus an unknown pattern X is recognized to be a member of the k th class if

$$\mu_k(X) = \bigvee_j \{ \mu_j(X) \}$$

where \bigvee denotes maximum and $j, k = 1, 2, \dots, m$.

Method II: Let $p_1, p_2, \dots, p_m, \dots, p_N$ be the N properties each of which represents some aspects of the unknown pattern X and has value only in the interval $[0, 1]$ such that

$$X = \{p_1, p_2, \dots, p_m, \dots, p_N\} \quad (5)$$

where

$$p_n = \left(1 + \left| \frac{\bar{x}_n - x_n}{E} \right|^F \right)^{-1.0} \quad (6)$$

and \bar{x}_n is n th reference constant determined from representative events of all the classes. Constants E and F have the same effect as in the previous method in affecting the fuzziness of a set.

If there are h number of prototypes in a class C_j , each reference point may then be represented as

$$R_j^{(l)} = \{p_{1j}^{(l)}, p_{2j}^{(l)}, \dots, p_{mj}^{(l)}, \dots, p_{Nj}^{(l)}\} \quad (7)$$

where $p_{nj}^{(l)}$ denotes the degree to which property p_n is possessed by the l th prototype in C_j . Then the similarity vector $S_j(X)$ for the

pattern X with respect to the j th class has the form

$$S_j(X) = \{s_{1j}, s_{2j}, \dots, s_{mj}, \dots, s_{Nj}\} \quad (8)$$

$$s_{nj} = \frac{1}{h} \sum_l s_{nj}^{(l)} \quad (9)$$

$$s_{nj}^{(l)} = (1 + W |1 - (p_n/p_{nj}^{(l)})|)^{-2Z} \quad (10)$$

where the numerical value of s_{nj} denotes the grade of similarity of the n th property with that of C_j . W is any positive constant dependent on each of the properties, and Z is an arbitrary integer. With the knowledge of all the similarity vectors one can decide $X \in C_k$ if

$$|S_j(X)| < |S_k(X)|, \quad k, j = 1, 2, \dots, m; k \neq j.$$

III. METHOD OF SEGMENTATION AND FEATURE EXTRACTION

A number of discrete phonetically balanced (PB) speech samples for all the vowels in CNC form were selected from the Telugu vocabulary. CNC combination is taken because the form of consonants connected to a vowel is responsible for influencing the role and quality of vowels. These speech units were recorded by five informants on an AKAI tape recorder. By listening experiments among ten listeners, only 871 samples uttered by three male informants in the age group of 28–30 years were chosen. Spectrum analyses were carried through a very standard audio frequency spectrum analyzer, Kay Sonagraph (model no. 7029A), which yields a permanent record of the formant frequencies F_1, F_2 , and F_3 in the range 5 Hz–16 kHz. Sonagraph was operated in the normal mode in the band 80 Hz–8 kHz with a wide BPF (bandwidth 300 Hz).

A. Segmentation Procedure

The segmentation procedure should satisfactorily solve the problems of determination of vowel boundary in relation to stops, fricatives, affricates, laterals in voiced/unvoiced, aspirated/unaspirated as well as their combined manners. The boundaries in these different situations need to be defined first.

1) For the stop consonants in the final position the start of occlusion period of the consonant will indicate the termination of the vowel. In some particular cases the vowel formants extend to a certain limit into the occluded part of the stop consonant, and for these cases the terminal point of the vowel includes this extended part. For unaspirated stops in the initial position, a gap very often separates the energy of plosion and the adjacent vowel formants. The starting point of the formants excludes this gap. For aspirated stops in the initial position, the onset of F_0 is taken to be the boundary.

2) The fricatives with a wide-band continuous energy spectrum of low intensity can be easily separated from the narrow band of much stronger intensity of vowel formants. The line separating these two distinctly separate spectral distributions is the vowel boundary for fricatives in the initial and also in final position. The line of demarcation for affricates in initial and final position is same as in fricative and stop consonants, respectively.

3) The segmentation problem becomes more complex for the liquid and vowel combinations. The liquids possess a formant-like structure very similar to vowel formants and thus create real confusion. But careful observation will reveal that the formant structure of liquids is less intense with a much lesser degree of transition. These characteristic differences are used for determining the vowel boundary in this case.

TABLE I
VOWEL SPEAKER MATRIX SHOWING NUMBER OF
EVENTS IN EACH CLASS

	X	Y	Z
/i/	43	61	68
/e/	60	68	79
/ə/	20	24	28
/a:/	29	19	41
/o/	39	35	56
/u/	41	59	51

B. Measurement Procedure

Formants F_1 , F_2 , and F_3 were obtained manually at the steady state of the vowels. The steady state of the vowel is that part on the record in which all formants lie parallel to the time axis. The transition is depicted by the inclined formant patterns. The exact point of inflection is difficult to locate in the records. This can be done very satisfactorily by tracing the central line for each formant band. Once these points are located for all available formants, the steady state of the vowel is taken to be the shortest horizontal span for all the formants.

In view of the large amount of data to be handled, the formant frequencies have been measured from the base line with a specially constructed scale. A rechecking on 5 percent of the samples revealed that formant frequencies have been recorded within an accuracy of 10 Hz. In a few cases, for particularly fast informants, it has been noticed that the vowel hardly reaches a stable state. In such cases the congruences of the on- and off-glides have been taken as the steady state.

In order to make the method of segmentation and formant extraction automatic, the maxima in the rate of change of spectral composition determined by a running summation of the absolute values of the rate of intensity change in a number of successive bands may be used, in general, as the criteria for the automatic method of determining segment boundaries. The method for automatic formant frequency extraction has been reported in a previous communication [20].

The respective features thus constitute a three-dimensional feature vector space Ω_F where each utterance of a speaker may be treated as an event from a population and each dimension represents an invariant characteristic of that event. Ten Telugu vowels (/ə/, /a:/, /i:/, /i:/, /u:/, /u:/, /e:/, /e:/, /o:/, /o:/) including shorter and longer categories were divided into six groups which contain vowels differing only in phonetic feature. Therefore, the multidimensional vector space is partitioned into 18 pattern classes (6 vowels \times 3 speakers), and each point in Ω_F thus associates three measured features corresponding to a vowel uttered by one of the three informants. The number of samples belonging to each group is tabulated through a vowel-speaker (X , Y , and Z , let us say) matrix (Table I).

IV. RECOGNITION PROCEDURE

The block diagram of the recognition program based on the automatic spoken word recognition model [20] developed at ISI is shown in Fig. 1, in which an appropriate serial answer of the question, "What is the vowel and who is its speaker?" is given. The system first of all searches for the vowel class irrespective of speakers, and then with the information of the vowel uttered, the specific informant is identified.

Prototype points chosen for vowel identification are the average of the coordinate values corresponding to the entire set of samples in a particular class. The features with increasing variance

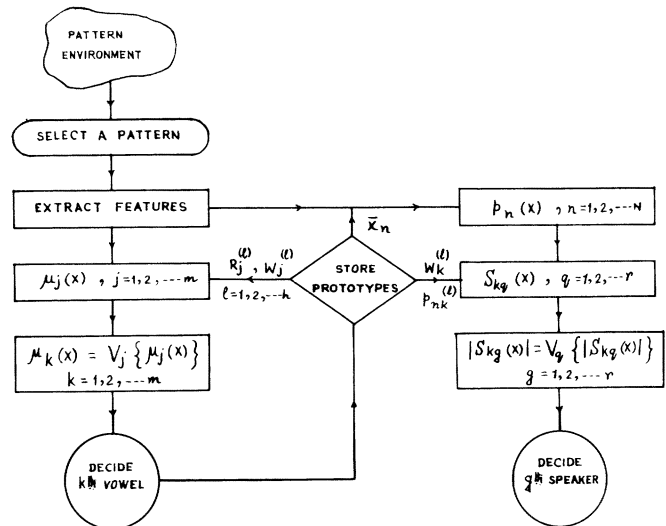


Fig. 1. Sequential recognition of vowel and its informant.

have been weighted with decreasing values of the weighting coefficients, and the inverse of the standard deviation of the formants as weighting coefficients were studied. Although shorter and longer types of vowels /i/, /u/, /e/, and /o/ are treated as the same group, they were given individual reference vectors and weighting coefficients computed over their respective set of events. Thus in our experiment of vowel sound recognition, $m = 6$, $N = 3$, $h = 1$, for /ə/, and /a:/, and $h = 2$ for /i/, /o/, /e/, and /u/. Computing membership values with respect to all the vowel classes using (3), an unknown vowel speech pattern is assigned to the k th class ($k = 1, 2, \dots, 6$) associated with maximum μ -value.

Whenever a vowel class of the unknown utterance is determined, then the system is engaged in finding its informants with the knowledge of stored prototypes for three speaker subclasses ($r = 3$) in that recognized vowel region. Since F_3 as compared to F_1 and F_2 bears more significant information about the speaker, the recognition features selected are (F_3/F_1) , (F_3/F_2) , and F_3 . Experiments were also done replacing the first two parameters by $(F_3 - F_1)$ and $(F_3 - F_2)$, respectively. Properties corresponding to each of the features were computed with $E = 100$, $F = 2$, and using (11):

$$\bar{x}_n = \sqrt{\frac{1}{h}} \sum_{q=1}^h x_n^{(q)}, \quad n, q = 1, 2, 3. \quad (11)$$

Similarity of the pattern with all the classes of informants was measured for $Z = 1$ by (9) and (10). Constant W was considered to be the inverse of the standard deviation of the recognition parameters, and in a part of the experiment (only for vowel /i/), similarity vectors were measured with $W = 1$ to investigate the influence of phase weights attached with the features. A subregion C_g ($g = 1, 2, 3$) possessing maximum closeness as measured by the magnitude of the similarity vectors is decided by the machine to be the corresponding speaker of the utterance.

To study the effect of training sets used in learning on a classified set, the above method of speaker identification was repeated thrice for each of the different sample sizes, viz., 1, 3, 5, 7, 10, and 15. This part of the experiment was carried out only with formants of the vowel /i/. Here samples were randomly drawn from each group of informants by which reference constants \bar{x}_n and W corresponding to the size of training samples were estimated. In a few cases, where the standard deviation of the magnitudes of a coordinate in a training set was zero, the corresponding W value was set at unity. Although it does not satisfy

	RECOGNISED AS						
	/i/	/e/	/ə/	/a:/	/o/	/u/	
S P O K E R	/i/	155	17				
	/e/	25	154	20	8		
	/ə/		9	52	8	2	1
	/a:/			17	70	2	
	/o/		1	7		138	34
	/u/			1		6	144

CORRECT RATE = 82%

Fig. 2. Confusion matrix illustrating machine's performance on vowel recognition.

(4), it is still logical in the sense that an attribute occurring with identical magnitude in all members of a training set is an all-important feature of the set, and hence its contribution in the discriminant function needs not be reduced.

V. EXPERIMENTAL RESULTS

The results of vowel recognition are explained by a confusion matrix in Fig. 2, where the figure in a cell represents the number of instances in which the same decision was made by the machine, and the diagonal elements thus indicate the number of utterances correctly identified. Overall recognition is about 82 percent, and it was found that the second maximum membership for classes /i/, /e/, /ə/, /a:/, /o/, and /u/, as expected from their phonetical order, correspond to vowel regions /e/, /i/, /a:/, /ə/, /u/, and /o/, respectively. In Fig. 2, confusion in machine recognition of a vowel is seen to be limited only to neighboring classes constituting a vowel triangle. This is in agreement with other experiments [22]–[24], [26].

Typical formant frequencies for the vowel /i/ uttered by three male informants in the age group of 28–30 years are shown in Table II. Table III illustrates the correct rate of decision rendered by the machine in identifying formants for each of the vowel speech sounds.

Scores shown are the average value of three observations only when the machine was trained by a set of five utterances for each speaker. With the fixation of appropriate phase weights which ensure the correct representation of feature-importance in classification, overall recognition accuracy is much improved (≈ 12 percent) compared to the case of no coefficients ($W = 1$). Whether parameters $(F_3 - F_1)$ and $(F_3 - F_2)$ are taken instead of (F_3/F_1) and (F_3/F_2) , respectively, the computer decision is altered very little.

Finally, variations of error rate (percent) with the set of training samples is graphically shown in Fig. 3, which demonstrates improvement in the machine's performance as the number of known labeled samples is increased. The curve is drawn using only the patterns of the vowel /i/. Sample sizes containing more than 8–10 utterances used in learning do not reduce the percentage error significantly. It could therefore be stated that after an optimum size of training patterns is achieved to provide good representation and weighting coefficients which characterize the classes, variation of recognition scores with training sets becomes insignificant.

TABLE II
TYPICAL FORMANTS FOR TELUGU VOWEL /i/
UTTERED BY THREE SPEAKERS

Informant	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
X	350	2350	2800
Y	350	2250	2700
Z	300	2350	3000

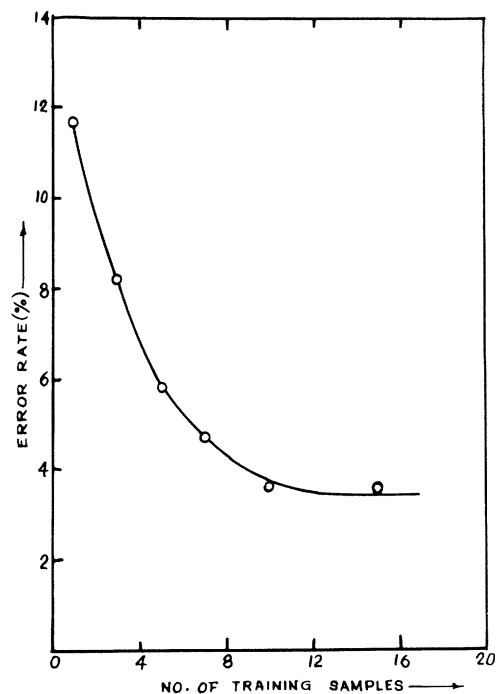


Fig. 3. Effect of training set on speaker identification score.

VI. CONCLUSIONS

Classification analysis using the concept of fuzzy sets is studied for machine recognition of informant and speech sounds for a large number of utterances. Accuracy of vowel sound recognition is about 82 percent when the decision of the machine was based only on the highest membership values. By incorporating a second choice under the control of a supervisory learning scheme, supposed to be based on linguistic constraints, the above score can be improved by 15 percent [21]. Knowledge of the weighting coefficients and reference vectors used is also available from any size of the training samples containing more than 12–16 utterance, without affecting the overall score [21].

Although it is well known that the fundamental voice frequency F_0 and higher formants (F_4 and F_5) are more speaker dependent, computer decision on the identification of informants in the age group of 28–30 years using property sets and the first three formants only is found to be satisfactory. It may be remarked that informants widely varying in age and sex are expected to result in a large discrimination in F_3 , leading to a better identification. At best, for 97 percent of the speech sounds, the machine was seen to render correct recognition of the speaker by comparing the magnitudes of similarity vectors. This is expected to be extended to a large number of informants varying in sex and age. It can be also stated that after an optimum number of learning samples (8–10 only) sufficient to characterize the representative points of a class, the size of the sample space can be extended enormously without disturbing the overall recognition score.

TABLE III
ACCURACY RATE (PERCENT) (AVERAGED OVER THREE OBSERVATIONS) OF SPEAKER
IDENTIFICATION FOR EACH VOWEL FORMANT

Speaker	Recognition Score (percent) for $h = 5$							
	/i/	/i/*	/i/**	/e/	/ə/	/a:/	/o/	/u/
X	89.47	73.69	84.21	92.60	100.00	82.76	95.84	91.30
Y	90.90	69.67	93.94	90.63	83.40	78.95	94.45	90.63
Z	100.00	100.00	100.00	100.00	96.43	95.13	83.34	92.00
TOTAL:	94.12	82.36	94.12	94.85	93.06	87.64	90.63	91.25

* Without weighting coefficient.

** Replacing (F_3/F_1) and (F_3/F_2) by (F_3-F_1) and (F_3-F_2) , respectively.

The paper therefore shows the possibility of recognizing both vowels and speakers with the help of the same features measured on a CNC word. Results of speaker identification were tabulated with the assumption that the vowel in the utterance is correctly recognized. If the immediate knowledge of the vowel is used (for sequential processing) then some but not all of the 18 percent error for vowel recognition will pass to speaker identification, since some samples were found to have error both in identifying the vowel and the speaker. However, our decision algorithm does not preclude the possibility of correct speaker identification from the misrecognized vowel features.

ACKNOWLEDGMENT

The authors wish to express their appreciation to Prof. C. R. Rao, F.R.S., Director, Indian Statistical Institute, for his interest in the work and acknowledge with thanks the help rendered by A. K. Datta, N. R. Ganguli, and B. Mukherjee in processing the spectrographic data. S. K. Pal, one of the authors, is also indebted to the Council of Scientific and Industrial Research, Government of India, for the award of a fellowship for this project.

REFERENCES

- [1] G. S. Sebestyen, *Decision Making Processes in Pattern Recognition*. New York: Macmillan, 1962.
- [2] K. S. Fu, *Sequential Methods in Pattern Recognition and Machine Learning*. London: Academic, 1968.
- [3] —, *Syntactic Methods in Pattern Recognition*. London: Academic, 1974.
- [4] L. A. Zadeh, "Fuzzy sets," *Inform. Control*, vol. 8, pp. 338-353, 1963.
- [5] —, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 28-44, Jan. 1973.
- [6] L. A. Zadeh, K. S. Fu, K. Tanaka, and M. Shimura, *Fuzzy Sets and Their Application to Cognitive and Decision Processes*. London: Academic, 1975.
- [7] T. Sakai and S. Doshita, "The automatic speech recognition system for conversational sound," *IEEE Trans. Electronic Computers*, vol. EC-12, pp. 835-846, Dec. 1963.
- [8] J. W. Forgie and C. D. Forgie, "Results obtained from a vowel recognition computer program," *J. Acoust. Soc. Am.*, vol. 31, pp. 1480-1489, 1959.
- [9] D. R. Reddy, "Segmentation of speech sounds," *J. Acoust. Soc. Am.*, vol. 40, pp. 307-312, 1966.
- [10] R. Jacobson, C. G. N. Fant, and M. Halle, "Preliminaries to speech analysis," MIT Acoust. Lab. Rep., no. 13, May 1952.
- [11] R. W. Shaffer and L. R. Rabiner, "System for automatic formant analysis of voiced speech," *J. Acoust. Soc. Am.*, vol. 47, pp. 634-648, 1970.
- [12] P. B. Denes and M. V. Mathews, "Spoken digit recognition using time-frequency pattern matching," *J. Acoust. Soc. Am.*, vol. 32, pp. 1450-1455, 1960.
- [13] S. Pruzensky, "Pattern matching procedure for automatic talker recognition," *J. Acoust. Soc. Am.*, vol. 35, pp. 354-358, 1963.
- [14] J. W. Glenn and N. Kleiner, "Speaker identification based on nasal phonation," *J. Acoust. Soc. Am.*, vol. 43, pp. 368-372, 1968.
- [15] B. S. Atal, "Automatic speaker recognition based on pitch contours," *J. Acoust. Soc. Am.*, vol. 52, pp. 1687-1697, 1972.
- [16] G. R. Doddington, "A method of speaker verification," *J. Acoust. Soc. Am.*, vol. 49, pp. 139, 1971.
- [17] A. E. Rosenberg and M. R. Sambur, "New technique for automatic speaker verification," *IEEE Trans. Acoust. Speech and Sig. Processing*, vol. ASSP-23, p. 169, Apr. 1975.

- [18] D. A. Wasson and R. W. Donaldson, "Speech amplitude and zero crossing for automated identification of human speakers," *IEEE Trans. Acoust. Speech and Sig. Processing*, vol. ASSP-23, pp. 390-392, Aug. 1975.
- [19] D. J. Broad, "Formants in automatic speech recognition," *Int. J. Man-Machine Studies*, vol. 4, p. 411, 1972.
- [20] D. D. Majumder and A. K. Datta, "A model for spoken word recognition," *Proc. Automazione Estrumentazione*, Milan, Italy, pp. 249-259, 1968.
- [21] D. D. Majumder, A. K. Datta, and S. K. Pal, "Computer recognition of vowel speech sounds using three-dimensional weighted discriminant function," IEEE Computer Group Repository, 1976.
- [22] D. D. Majumder and S. K. Pal, "The concept of fuzzy sets and its application in pattern recognition problems," *Proc. CSI, 76 Convention*, Hyderabad, India, no. SD 02, Jan. 1976.
- [23] —, "On some applications of fuzzy algorithm in man-machine communication research," *J. Inst. Telecom. Electron. Eng.*, vol. 23, pp. 117-120, Mar. 1977.
- [24] D. D. Majumder, S. K. Pal, and B. B. Chaudhuri, "Some experiments on computer recognition of speech pattern," *Proc. Int. Conf. Inform. Sc. Syst.*, Patras, Greece, August 1976, Hemisphere Publishing, USA (to appear).
- [25] A. K. Datta, "An experimental procedure for handwritten character recognition," *IEEE Trans. Comput.*, vol. C-23, pp. 536-545, May 1974.
- [26] K. K. Paliwal and P. V. S. Rao, "Computer recognition of isolated phonemes," *Proc. CSI, 76 Convention*, Hyderabad, India, no. SR 04, Jan. 1976.

On the Evolution Equations of Certain Social-Process Models

I. W. SANDBERG, FELLOW, IEEE

Abstract—The equations considered are a pair of general coupled first-order differential equations of the kind that arise in connection with two-country Richardson-type arms-race models as well as in connection with models of several other social processes. Weak sufficient conditions are given, which have a direct and often useful interpretation, under which for each starting point there is a unique solution of the pair of equations, and all solutions approach a common point as $t \rightarrow \infty$.

PRELIMINARIES

We use the following basically standard notation and definitions. The real interval $(-\infty, \infty)$ is denoted by R , and R^2 denotes the Euclidean space of real two-dimensional vectors, with norm $|\cdot|$ and zero vector θ . For each $v \in R^2$, v_1 and v_2 denote the components of v . For v and u in R^2 , the inequality $v \leq u$ means that $v_i \leq u_i$ for $i = 1, 2$. Similarly, for v and u in R^2 , $v < u$ means that $v_i < u_i$ for both values of i . The interval $[0, \infty)$ and the set $\{v \in R^2: v \geq \theta\}$ are denoted by R_+ and R_+^2 , respectively.

Let $f_1(\cdot)$ and $f_2(\cdot)$ denote continuous functions from $R_+ \times R_+$ into R , and let f be the function from R_+^2 into R^2 defined by $f_1(u) = f_1(u_1, u_2)$ and $f_2(u) = f_2(u_1, u_2)$ for all $u \in R_+^2$.

Assume the following.

1) f has the "off-diagonal monotonicity" property that for each i , $f_i(u) \geq f_i(v)$ for every u and v in R_+^2 such that $u \geq v$ and $u_i = v_i$.

Manuscript received January 3, 1977; revised March 14, 1977.

The author is with the Department of Mathematical Studies, Bell Laboratories, Murray Hill, NJ 07974.