Proceedings of The Institute of Acoustics

FUZZY SETS AND THEIR APPLICATIONS TO SPEECH RECOGNITION

J DWYER AND P A LEE

POLYTECHNIC OF THE SOUTH BANK

Introduction

The notion of a fuzzy set was first suggested by Zadeh (1965). Since then fuzzy set theory, together with directly related theories in fuzzy logic and fuzzy algorithms, have increasingly found applications in areas where ordinary set theory is inadequate.

One such area concerns the techniques required for the recognition of patterns. It is the intention in this paper to introduce the idea of a fuzzy set and to indicate some ways in which fuzzy sets can be applied to speech recognition. Particular reference is made to research into speech recognition in the department of electrical and electronic engineering at the Polytechnic of the South Bank.

Fuzzy sets

The theory of fuzzy sets concerns a subset A of a universal set $\mathcal{E}$ where the boundary between membership and non-membership in A is not precise. The value 1 is usually assigned to members that definitely belong to A, whereas value 0 is assigned to members that definitely do not belong to A.

Formally, a fuzzy set A in $\mathcal{E}$ can be defined as a set of ordered pairs

$$A = \{(x), \mu_A(x)\} \quad , \quad x \in \mathcal{E} \ ,$$

in which $\mu_A(x)$ is a membership function representing the grade of membership of x in A. The height of a fuzzy set A is the supremum of $\mu_A(x)$ over A and A is said to be normal if its height is 1. Otherwise A is called subnormal. Membership functions consistent with the specifications of a set are said to be admissible. Otherwise membership functions are called non-admissible.

As a simple illustration, Fig. 1 shows three different admissible membership functions $\mu_B^{I}(x)$, $\mu_B^{II}(x)$ and $\mu_B^{III}(x)$ corresponding to a fuzzy set B specified as the set of real numbers "close to" 0. Here the universal set in question is the set of real numbers, ie. $\mathcal{E} = \mathbb{R}$. The membership functions are given as

$$\mu_B^{I}(x) = \begin{cases} 1 & |x| \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_B^{II}(x) = \begin{cases} 1 - x/2 & |x| \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_B^{III}(x) = (1 + x^2)^{-1} \text{ for all } x$$

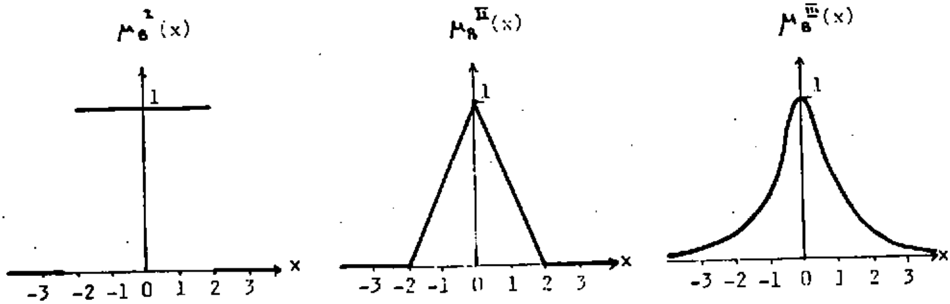


Fig. 1 Admissible membership functions

Clearly $\mu_B^{I}(x)$ corresponds to the ordinary (non-fuzzy) set of points in the closed interval [-2, 2], since each and every point in this interval is assigned a value 1 whereas remaining points in $\mathbb{R}$ are assigned value 0. $\mu_B^{II}(x)$ and $\mu_B^{III}(x)$ ensure that transition between membership and non-membership is gradual rather than abrupt. However it can be argued that $\mu_B^{III}(x)$ gives a better representation for the set B as specified since it contains information that member x = 3, for example, is not as close to 0 as member x = 2. $\mu_B^{II}(x)$ on the other hand assigns x = 3 and x = 2 exactly the same value, namely 0. It should be noted that in general, a choice of membership function is essentially subjective.

Let us now consider how ideas regarding unary and binary operations on non-fuzzy sets are extended to fuzzy sets. For any two fuzzy sets C and D defined over $\mathcal{E}$ with membership functions $\mu_C(x)$ and $\mu_D(x)$ respectively we have the following definitions:-

(i) C and D are equal (written C = D) if and only if $\mu_C(x) = \mu_D(x)$ for all $x \in \mathcal{E}$.

(ii) C is contained in D (written $C \subseteq D$) if and only if $\mu_C(x) \leq \mu_D(x)$ for all $x \in \mathcal{E}$.

(iii) The union of C and D (written $C \cup D$) has a membership function given by $\mu_{C \cup D}(x) = \max(\mu_C(x), \mu_D(x))$ for all $x \in \mathcal{E}$.

(iv) The intersection of C and D (written $C \cap D$) has a membership function given by
$$\mu_{C \cap D}(x) = \max(\mu_C(x), \mu_D(x))$$
for all $x \in \mathcal{E}$.

(v) The complement of C (written $\bar{C}$) has a membership function given by

$$\mu_{\bar{C}}(x) = 1 - \mu_C(x)$$

for all $x \in \mathcal{E}$.

(vi) The product of C and D (written CD) has a membership function given by

$$\mu_{CD}(x) = \mu_C(x) . \mu_D(x)$$

for all $x \in \mathcal{E}$.

(vii) The concentration of C (written con(C)) has a membership function given by

$$\mu_{con(C)}(x) = \mu_C^2(x)$$

for all $x \in \mathcal{E}$.

(viii) The dilation of C (written dil(C)) has a membership function given by

$$\mu_{dil(C)}(x) = \mu_C^{1/2}(x)$$

for all $x \in \mathcal{E}$.

For further definitions concerning fuzzy set operations, an interested reader might wish to refer to Kandel (1982).

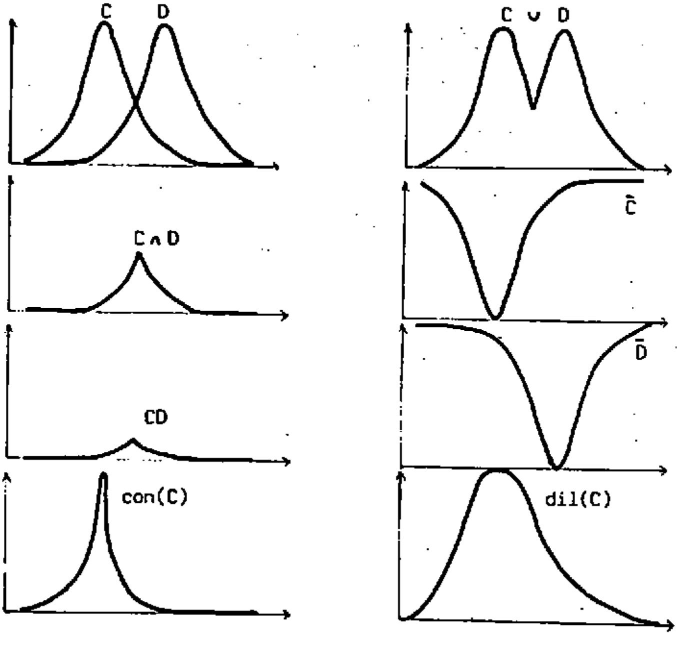


Fig. 2 Unary and binary operations on fuzzy sets

FUZZY SETS AND THEIR APPLICATIONS TO SPEECH RECOGNITION

Sketches of membership functions appropriate to the above definitions are shown in Fig. 2 for a particular choice of sets C and D. It can be easily shown that four of the five basic 'dual' laws of ordinary sets (ie. commutative, associative, distributive and identity) all apply to fuzzy sets. However the dual law of complementation viz.

$$C \wedge \bar{C} = \phi \quad ; \quad C \vee \bar{C} = \mathcal{E},$$

in which ϕ is the empty set and $\mathcal{E}$ is the universal set, does <u>not</u> hold true in general (cf. Fig. 3).

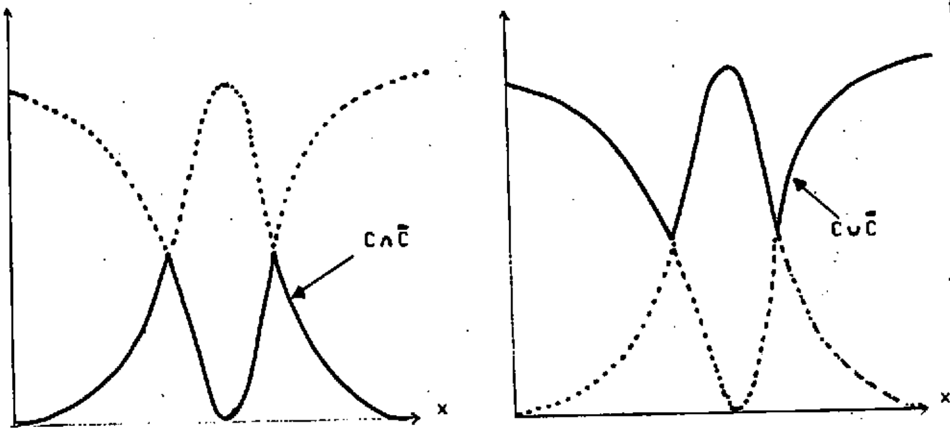


Fig. 3 The membership function corresponding to $C \wedge \bar{C}$ and $C \vee \bar{C}$

Extension of fuzzy set theory to fuzzy logic is straightforward. Fuzzy logic is a many-valued logic in which the truth of a proposition P(x) can assume any value in the closed interval [0,1]. Ordinary logic allows only two truth values, viz. 0 and 1.

Consider for example the proposition

$P_1(x)$: x "is close to" 0.

An admissible truth value function (t.v.f.) for $P_1(x)$ can be chosen to correspond directly with $\mu_B^{\bar{\bar{}}}(x)$ of Fig. 1. This truth value function is shown in Fig. 4.

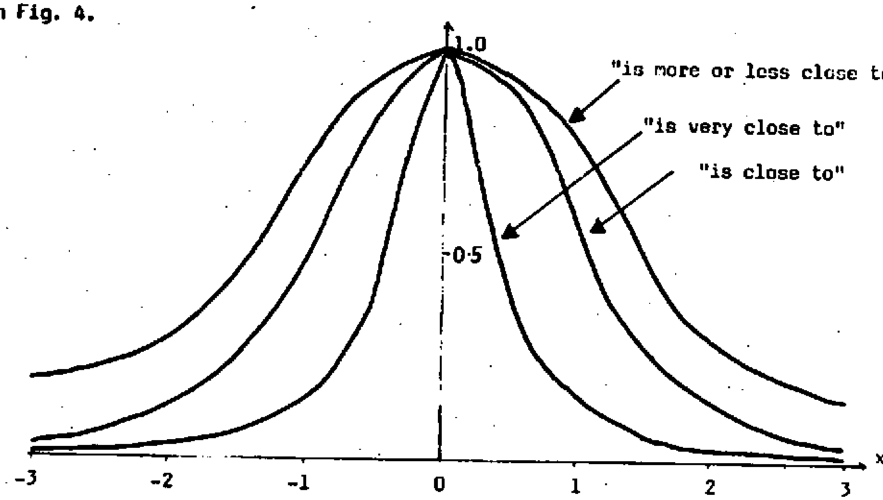


Fig. 4 Modification of a truth value function

Modification of a fuzzy proposition can be achieved by modification of its truth value function. Fig. 4 shows the modified predicates "is more or less close to" and "is very close to" obtained from $P_1(x)$ by dilation and concentration, respectively.

A binary fuzzy relation R in $\mathbb{E}_1 \times \mathbb{E}_2$ is a fuzzy subset of $\mathbb{E}_1 \times \mathbb{E}_2$, where X represents Cartesian product. For example, if $\mathbb{E}_1 = \mathbb{E}_2 = \mathbb{N}$, the set of natural numbers, a fuzzy relation matrix can be drawn up to describe R, the relation "is close to" for which $\mu(x_1, x_2) = (1+(x_1 - x_2)^2)^{-1}$ for all x_1 and $x_2 \in \mathbb{N}$.

R	1	2	3	4	5	→
1	1.0	0.5	0.2	0.1	0.06	→
2	0.5	1.0	0.5	0.2	0.1	→
3	0.2	0.5	1.0	0.5	0.2	→
4	0.1	0.2	0.5	1.0	0.5	→
5	0.06	0.1	0.2	0.5	1.0	→
↓	↓	↓	↓	↓	↓	↘

Fig. 5 A fuzzy relation matrix

This fuzzy matrix is shown in Fig. 5. As expected elements of the matrix away from the diagonal are valued less than 1, whereas elements situated on the diagonal have value 1.

FUZZY SETS AND THEIR APPLICATIONS TO SPEECH RECOGNITION

Applications

Given a sample pattern x and a set (or cluster) of patterns A, the recognition problem often is not whether x is or is not a member of A but the degree to which x is a member of A. Therefore, in a fuzzy set context, a recognition algorithm applied to x must yield the grade of membership $\mu_A(x)$. When an explicit description of this algorithm is known we shall say that the algorithm is transparent. Otherwise the algorithm is opaque.

Human perception involves the ability to adjust frames of reference so that it is possible to relate to a particular structure. Frequently, opaque algorithms are then used to recognise and classify. For instance a person cannot explicitly describe the algorithm used to assign a membership grade to a specific painting that classifies it in the fuzzy set of beautiful paintings.

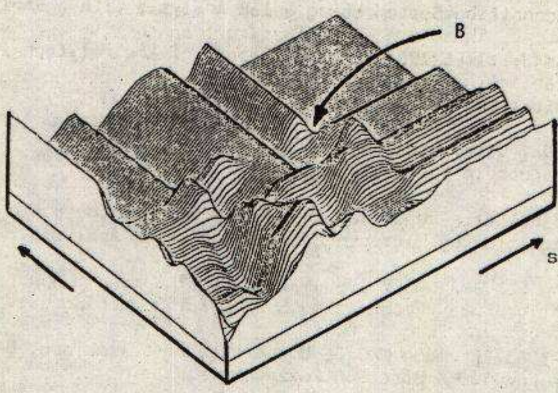
Clearly any simulation of the human recognition process requires the conversion of opaque algorithms to transparent algorithms. Accordingly there are at least two lines of attack provided by fuzzy set theory. These are respective approaches of template-similarity and representation of expert knowledge. While the latter approach is not discussed here in detail, it should be stressed that conventional 'knowledge' techniques based on predicate calculus are not applicable to many-valued propositions (see Zadeh 1983).

The template-similarity approach regards $\mu_A(x)$ as the degree of similarity between x and an 'ideal' pattern (template) representing A. In the case of speech recognition systems, feature extraction can yield a mathematical object, eg. a matrix, which is then matched with a template through a transparent algorithm and a decision made about the nature of the input speech.

Two systems based on this approach are described by Abu El-ata and Seymour (1983) and Lee (1983). Both systems concentrate on economy of algorithm to realise a low-cost implementation. The linear-programming approach of the former system matches the sample matrix and the template matrix using a simple algorithm in which differences between corresponding elements are computed. In the dynamic-programming approach of Lee (1983), the algorithm is more complicated. Template similarity is assessed via several stages, one of which involves the generation and comparison of similarity and dynamic- programming surfaces similar to those shown in Fig. 6. Further, Pal and Majumder (1980) use 'weighted distances' to measure the similarity of patterns associated with vowel and speaker recognition.

Remarks

The speech recognition group at the Polytechnic of the South Bank continues to investigate fuzzy sets and their applications to the entire recognition process. Fortunately with the advent of new initiatives in VLSI design (see, for example, Dwyer and Yates 1983), it is not necessary to concentrate on simplicity of algorithms. It is to be required only that any suitable transparent algorithm can be implemented on a parallel processing system with controlled (polynomial) complexity. Certainly there is currently much interest in the application of fuzzy sets to pattern recognition generally and a comprehensive list of references is given by Kandel (1982).



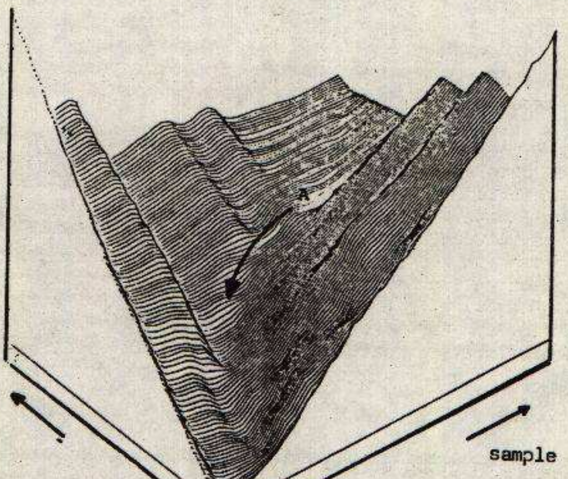


Fig. 6 Similarity and dynamic-programming surfaces (labelled B and A respectively)

References

1. M.A. ABU EL-ATA and J. SEYMOUR 1983 Acustica 54, 52-56
A Speech Recognition System Using Walsh Analysis with a Small Computer.

2. J. DWYER and J. YATES 1983 Research Report of the Polytechnic of the South Bank 5, 1-16
New Initiatives in VLSI Design.

3. A. KANDEL 1982 Wiley and Sons
Fuzzy Techniques in Pattern Recognition.

4. P. LEE 1983 Research Report of the Polytechnic of the South Bank 13, 1-56
Speech Recognition by Microprocessor Using Walsh Analysis.

5. S.K. PAL and D.D. MAJUMDER 1980 Fuzzy Sets Theory and App. (Durham, NC)
A Self-adaptive Fuzzy Recognition System for Speech Sounds.

6. L.A. ZADEH 1983 ERL memo M83/26 University of California, Berkeley
Commonsense Knowledge Based on Fuzzy Logic.

7. L.A. ZADEH 1965 Inf. Control 8, 338-353
Fuzzy Sets.