Fuzzy Sets and Systems 21 (1987) 351–356
North-Holland

A NOTE ON RANDOM SETS AND THE THURSTONIAN SCALING METHODS*

Rami ZWICK
University of North Carolina, Chapel Hill, NC27514, USA

Received July 1985
Revised December 1985

We argue that if a vague concept is represented by a random set, then it is justifiable to use the law of comparative judgment to estimate the means of the membership function distributions. These means may be taken as the type 1 membership values.

Keywords: Random sets, Thurstonian scaling methods, Law of comparative judgment.

1. Introduction

The theory of fuzzy subsets as a purely mathematical discipline is independent of the psycholinguistic reality of fuzzy sets [2]. Any complete treatment of fuzzy subsets must assume well specified membership functions. However, membership value is not a primitive concept from a psychological point of view. Thus, when the theory of fuzzy subsets is used to model how experts or other people deal with vagueness in natural language, the specific scaling technique used to construct the membership function must be justified on psychological and logical grounds.

There is an apparent inconsistency in using a precise membership function to represent a vague concept. To account for this apparent paradox, Zadeh [17] defined type m fuzzy sets. Type m fuzzy sets are defined recursively as follows: a type 1 fuzzy set is an ordinary fuzzy set; a type m fuzzy set ($m > 1$) is an L-fuzzy set whose membership values are type $m-1$ fuzzy sets on $[0, 1]$. Type 2 fuzzy sets are easily interpreted, and are intuitively appealing because grades of membership in the first level cannot be assumed to be known precisely [9]. However, a conceptual difficulty in working with type m fuzzy sets is that it quickly leads to an infinite regress paradox. 'Generalized' fuzzy set theory [16] acknowledges the imprecision inherent in the first level of membership values due to the vague nature the 'concept' A is representing. This imprecision is modeled by putting a second level of membership values on the first. However, this second level is similarly seen to be imprecise and a third level of measurement is necessary. This process, based on the original logic, should continue ad infinitum. Furthermore, from a practical point of view, even two levels of measurement are

* This research was supported by contract MDA 903-83-K-0347 from the U.S. Army Research Institute. The views, opinions, and findings contained in this paper are those of the author and should not be construed as official Department of the Army position, policy, or decision.

highly undesirable. Estimating the membership function of higher order fuzzy sets would certainly be more difficult than estimating a single membership function.

An alternative, but closely related approach to representing vague concepts by fuzzy sets, is to use a probabilistic or random sets as the modeling tool [1, 3–7, 12] in which the value of the membership function is a random variable. Norwich and Turksen [12] pointed out that structuring a scaling procedure based on random sets allows us to view the inconsistency in a subject's response as an inherent fluctuation of the concept definition in the subject's mind.

If this approach is adopted, then the establishment of membership functions can be closely related to the well known psychometric scaling techniques developed by Thurstone [13, 14].

2. Thurstone's judgment scaling model

The Thurstonian approach postulates a psychological scale onto which stimuli are mapped. Each time a stimulus is presented it is presumed to be represented by a point along this scale, the location of which is determined by an unknown process. Because of momentary fluctuations in the organism, a given stimulus does not always excite the same value, and therefore the value is a random variable on the psychological continuum.

It is assumed that repeated presentations of the same stimulus produces a distribution (called a discriminal process) of such values along the psychological scale. On the assumption that the observer cannot directly report the value of the discriminal value, the scaling must be done indirectly. Equations based on the judgments of relations among stimuli can be deduced and used to estimate the modal discriminal value of each stimulus.

In this paper we will describe one set of equations known as the law of comparative judgment.

3. The law of comparative judgment

We assume that the stimuli are elements of the space of discourse (Ω) and that the underlying psychological continuum is 'F-ness', where F is a fuzzy concept (i.e., tall, long, etc.). Now suppose that we present a stimulus pair (x, y) to a subject a number of times. On each presentation a single discriminal value (a point on the psychological continuum) is elicited from the distribution of each stimulus. Each time the two stimuli are presented, the subject is required to judge which is higher on the psychological continuum (in our terminology the subject is required to judge the relative 'F-ness' of x and y). The basic assumption is that the discriminal processes can be represented by normal distributions on the underlying psychological continuum. Stimuli x and y will form two normal discriminal processes $\Psi(x)$ and $\Psi(y)$ with means μ_x and μ_y and standard deviations (called discriminal dispersions) of σ_x and σ_y.

If the two stimuli are presented together a large number of times, the discriminal differences themselves form a normal distribution with mean $\mu_x - \mu_y$ and standard deviation

$$\sigma_{x-y} = (\sigma_x^2 + \sigma_y^2 - 2r_{xy}\sigma_x\sigma_y)^{1/2},$$

where r_{xy} is the correlation between the two discriminal processes. It is assumed that the subject will respond that "stimulus x has more F-ness" or "is more F than stimulus y", if on that specific trial the discriminal value for stimulus x exceeds that for stimulus y on the psychological continuum ($\Psi(x) > \Psi(y)$).

Since $\Psi(x) - \Psi(y)$ is distributed normally with mean $\mu_x - \mu_y$ and variance σ_{x-y}^2 it follows that

$$[\Psi(x) - \Psi(y) - (\mu_x - \mu_y)]/\sigma_{x-y} \sim \mathrm{N}(0, 1).$$

Let $p(xy)$ be the theoretical proportion of times that stimulus x is judged higher than stimulus y on the psychological continuum, based on the theoretical discriminal distributions. Then

$$p(xy) = P[\Psi(x) > \Psi(y)] = P[\Psi(x) - \Psi(y) > 0].$$

Hence

$$P\{[(\Psi(x) - \Psi(y) - (\mu_x - \mu_y))/\sigma_{x-y}] > [-(\mu_x - \mu_y)/\sigma_{x-y}]\} = p(xy)$$

so that

$$-(\mu_x - \mu_y)/\sigma_{x-y} = z_{1-p(xy)}$$

We can thus write the equation

$$\mu_y - \mu_x = z_{1-p(xy)}(\sigma_x^2 + \sigma_y^2 - 2r_{xy}\sigma_x\sigma_y)^{1/2}. \tag{1}$$

Equation (1) is the complete form of the law of comparative judgment. Note that only a pair-wise ordinal judgment is required from the subject, which generally is preferable to a direct rating.

Our goal is to estimate μ_x and μ_y using the observed proportion ($p^0(xy)$) and to treat them as a first order approximation for each element's membership value regarding 'F-ness'. The law of comparative judgment is not solvable in the form of Eq. (1) since, regardless of the number of stimuli and observations, there are always more unknowns than observation equations. Thus, additional assumptions are required.

Thurstone [13] recognized five cases of the law of comparative judgment. We are interested here only in the case where replications are within a single individual, and we will adopt the Torgerson [15] classification of Thurstone's work.

Torgerson's condition A. In this condition is is assumed that the covariance term in all equations is a constant value ($2c$), hence Eq. (1) reduces

$$\mu_y - \mu_x = z_{1-p(xy)}(a_x + a_y)^{1/2} \tag{2}$$

where

$$a_x = \sigma_x^2 - c \quad \text{and} \quad a_y = \sigma_y^2 - c.$$

With n stimuli, there are n means (μ_i), and n a's (a_i) that are unknown. We can set one of the scale values to zero and one of the a's equal to one, leaving us with $2n-2$ unknowns. Since with n stimuli we have $\frac{1}{2}n(n-1)$ equations, a solution is possible when $[\frac{1}{2}n(n-1)] > [2n-2]$ or when $n \geq 4$. With $n > 4$ we can evaluate experimentally the goodness of fit of the solution.

Torgerson's condition C. In this more practical condition it is assumed that σ_{x-y} is constant for all pairs of stimuli (c). Hence Eq. (1) reduces to

$$\mu_y - \mu_x = c z_{1-p(xy)}. \tag{3}$$

This condition is attractive since a least-squares estimate of the scale values for complete and incomplete data exist. Given a matrix X, which contains the sample estimate $z_{1-p^0(xy)}$ of the theoretical values, Mosteller [10] has shown that a least-squares estimate of the scale values can be obtained by averaging the columns of matrix X (if matrix X contains no vacant cells). Solutions for condition C when matrix X is incomplete do exist. The interested reader should consult Torgerson [15], pp. 173–179.

To summarize, if the subject bases his or her judgment of stimulus dominance upon the difference between the two internal discriminal processes, then information on the distribution of such differences accumulates over a large number of trials. The probability with which one stimulus is judged to dominate another can then be converted into an estimate of the mean of this distribution of differences and consequently a measure of the distance separating the two stimuli along the psychological scale. The only evidence the experimenter has of these processes is the frequency with which one stimulus dominated another in a specific experiment. By adding some major assumptions, this is enough evidence to infer differences in scale values.

4. Major assumptions

All forms of the law of comparative judgment assume that each stimulus is independently compared a large number of times with each other stimulus. However if the subject can identify the stimulus pairs, there is the possibility that he or she will base later judgments on the memory of earlier judgments of the pair. Thus in order to use this procedure, the stimuli should be such that no extraneous differentiating cues are available to the subject. Furthermore, it is necessary to control the conditions that might introduce biasing effects by randomizing relative positions and orders and by use of counter balancing procedures.

It is important to realize that in order to use the Thurstonian approach, one for which a large enough set of stimuli will give more observational equations than unknowns, it is necessary to specify additional restrictions. In all cases a *normal distribution* is usually assumed. Thurstone cautioned that: "The only valid justification for bringing in the (normal) probability curve in this connection is that its presence can be experimentally tested". [14, pp. 368].

However as noted by Luce [8], testing the assumption of normality has proved rather less easy and less direct than Thurstone implied. Nevertheless based on an argument about neural schema, Luce [8] concluded that Thurstone's choice of the normal distribution was correct. If many different factors contribute to the inherent noise in the nervous system, and if the central nervous system averages the estimates from the individual fibers, and if these estimates are of comparable magnitude, then the Central Limit Theorem tells us that the resulting estimate is approximately normally distributed. For a more complete treatment of this idea see [8].

Another common assumption is that of *unidimensionality additivity*. In other words, it is assumed that all the stimulus comparisons take place along a unidimensional hypothetical psychological continuum. Furthermore in Torgeson's condition A it is assumed that the covariance term in the complete law of comparative judgment (Eq. 1) is constant for all pairs of stimuli. Finally in condition C it is assumed that σ_{x-y} is constant for all pairs of stimuli.

The validity of Thurstone's scaling techniques therefore depends on our willingness to accept the underlying assumptions and on the robustness of the techniques with regard to one or more violations of the assumptions. The advantage of using these techniques is that we can test the goodness of fit of the model to the data.

A single over-all test of goodness of fit for condition C has been given by Mosteller [11]. We define

$$\theta'(ij) = \arcsin(p^0(ij))^{1/2},$$

$$\theta''(ij) = \arcsin(p^s(ij))^{1/2},$$

$$n = \text{number of observations per pair of stimuli},$$

and

$$k = \text{number of stimuli},$$

where $p^0(ij)$ are the observed proportions and $p^s(ij)$ are the proportions derived from the estimated scale values. Then we can test goodness of fit by

$$\chi^2 = n \sum (\theta'(ij) - \theta''(ij))^2/821,$$

with $\frac{1}{2}(k-1)(k-2)$ degrees of freedom.

This test is sensitive to violations of additivity, and is affected unpredictably by unequal standard deviations of discriminal differences. If there is a reason to believe that the standard deviations are unequal, one might want to relax the conditions for case C and try some other case [15]. However, this would require a large number of stimuli and observations.

5. Conclusion

We have argued that from a psychological measurement point of view, it is preferable to represent a vague concept by a random set rather than a type m

fuzzy set. Doing so enables one to naturally justify and apply psychological scaling techniques developed by Thurstone and others.

References

[1] E. Czogala, Probabilistic Sets in Decision-Making and Control, Interdisciplinary System Research Vol. 81 (Verlag TUV, Rheinland, Koln, 1984).

[2] G.R. Dreyfuss, M. Kochen, J. Robinson and A.N. Badre, On the psycholinguistic reality of fuzzy sets, in: Papers from the Parasession on Functionalism, April 17, 1975 (Chicago Linguistic Society).

[3] I.R. Goodman, Fuzzy sets as equivalence classes of random sets, in: R.R. Yager, Ed., Recent Development in Fuzzy Sets and Probability Theory (Pergamon, New York, 1982) 327–342.

[4] I.R. Goodman, Some new results concerning random sets and fuzzy sets, Inform. Sci. 34 (1984) 93–113.

[5] I.R. Goodman and H.T. Nguyen, Uncertainty Models for Knowledge-Based Systems (North-Holland, Amsterdam, 1985).

[6] K. Hirota, Concept of probability sets, IEEE Conference on Decision and Control, New Orleans, LA (1975) 1361–1366.

[7] K. Hirota, Concepts of probability sets, Fuzzy Sets and Systems 5 (1981) 31–46.

[8] R.L. Luce, Thurstone's discriminal process fifty years later, Psychometrika 42 (1977) 461–489.

[9] M. Mizumoto and K. Tanaka, Some properties of fuzzy sets of type 2, Inform. and Control 31 (1976) 312–340.

[10] F. Mosteller, Remarks on the method of pair comparisons: I. The least squares solution assuming equal standard deviations and equal correlations, Psychometrika 16 (1951) 3–11.

[11] F. Mosteller, Remarks on the method of pair comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed, Psychometrika 16 (1951) 207–218.

[12] A.M. Norwich and I.B. Turksen, Stochastic fuzziness, in: M.M. Gupta and E. Sanchez, Eds., Approximate Reasoning in Decision Analysis (North-Holland, Amsterdam, 1982) 13–22.

[13] L.L. Thurstone, A law of comparative judgment, Psychological Rev. 34 (1927) 273–286.

[14] L.L. Thurstone, Psychophysical analysis, Amer. J. Psychology 38 (1927) 368–389.

[15] W.S. Torgerson, Theory and Methods of Scaling (Wiley, New York, 1958).

[16] S. Watanabe, A generalized fuzzy set theory, IEEE Trans. Systems Man Cybernet. 8 (1978) 756–760.

[17] L.A. Zadeh, Quantitative fuzzy semantics, Inform. Sci. 3 (1971) 159–176.