# Data Evaluation in Fuzzy Systems

MICHAEL GR. VOSKOGLOU
Mathematical Sciences, School of Technological Applications
Graduate T. E.I. of Western Greece
Meg. Alexandrou 1 – 263 34 Patras
GREECE
mvosk@hol.gr ; http://eclass.teipat.gr/eclass/courses/523102

*Abstract:* - The data obtained from the operation mechanisms of large and complex systems connected to science and real life applications are frequently characterized by uncertainty and fuzziness. The paper at hands proposes a model for evaluating such kind of data with the help of the corresponding membership degrees and possibilities. Examples are also presented illustrating the applicability of the proposed model in practice.

*Key-Words:* - Fuzzy set (FS), membership degrees, possibility, fuzzy data evaluation, fuzzy variables.

## 1 Introduction

In large and complex systems that frequently appear nowadays in science, technology and real life situations (like the socio – economic, the biological ones, etc.) many different and constantly changing factors are usually involved, the relationships among which are indeterminate. As a result the data obtained from their operation mechanisms cannot be easily determined precisely and in practice estimates of them are used.

While 50-60 years ago the unique tool in hands of the scientists for handling such kind of data, and situations of uncertainty in general, used to be the theory of *Probability*, nowadays the *Fuzzy Set (FS)* theory initiated by Zadeh in 1965 [1] and its extensions and generalizations that followed in the recent years [], have given a new dynamic to this field.

In the article at hands a model is developed for evaluating a system's fuzzy data in terms of the corresponding fuzzy possibilities. The rest of the article is organized as follows: In Section 2 the general model is developed and an application to learning mathematics is presented illustrating its applicability in real situations. In Section 3 the general model is extended for studying the combined results of the evaluation of fuzzy data obtained from two (or more) different sources and an example is provided on a market's research to emphasize the usefulness of this extension for tackling real life problems. The article closes with the final conclusions stated in Section 4.

## 2 The General Model for Evaluating Fuzzy Data

The reader is considered to be familiar to the basics of the FS theory and the book [3] is proposed as a general reference on the subject.

Assume that one wants to study a system's behavior consisting of *n* components (objects), $n \geq 2$, during a process involving vagueness and/or uncertainty. Denote by $S_i$ , i=1,2,3 the main steps of that process * and by *a, b, c, d, e* the linguistic labels of very low, low, intermediate, high and very high success respectively of the system components in each step.

Set $U = \{a, b, c, d, e\}$. A FS $A_i$ in $U$ will be associated to each step $S_i$, i = 1, 2, 3. For this, if $n_{ia}$, $n_{ib}$, $n_{ic}$, $n_{id}$, $n_{ie}$ denote the numbers of the system components that faced very low, low, intermediate, high and very high success respectively at stage $S_i$, we define the *membership degree* $m_{A_i}(x)$ of each $x$ in $U$ by

$$m_{A_i}(x) = \frac{n_{ix}}{n} \quad (1) \quad .$$

Then the FS $A_i$ in $U$ associated to $S_i$ is of the form:

$$A_i = \{(x, m_{A_i}(x)): x \in U\}, i=1, 2, 3 \quad (2).$$

In order to represent all possible *profiles (overall states)* of the system components during the corresponding process a *fuzzy relation*, say *R*, in $U^3$ (i.e. a FS in $U^3$) is considered of the form:

* For reasons of simplicity we consider here three steps, but the model holds in general for a finite number of steps

$$R= \{(s, m_R(s)): s=(x, y, z) \in U^3\} \quad (3).$$

Usually in practical applications the degree of success of each system's component in a certain step of the process depends upon the degree of its success in the previous step. Under this assumption and in order to define properly the membership function $m_R$, the following definition is given:

*Definition*: A profile $s=(x, y, z)$, with $x, y, z$ in $U$, is said to be *well ordered* if $x$ corresponds to a degree of success equal or greater than $y$ and $y$ corresponds to a degree of success equal or greater than $z$.

For example, $(c, c, a)$ is a well ordered profile, while $(b, a, c)$ is not. The membership degree of a well ordered profile $s$ is defined now to be equal to the product

$$m_R(s) = m_{A_1}(x) \cdot m_{A_2}(y) \cdot m_{A_3}(z) \quad (4).$$

On the contrary, the degree of the profiles which are not well ordered is defined to be zero. In fact, if for example the profile $(b, a, c)$ possessed a nonzero membership degree, then at least one of the system components demonstrating a very low performance at step $S_2$ would perform satisfactorily at the next step $S_3$, which is impossible to happen.

However, they are also real situations in which the performance of each component at each step does not depend on its performance in the previous steps (e.g. see Example 2 of Section 3). In such cases the membership degrees of all profiles are defined by equation (4).

Next, for reasons of brevity, we shall write $m_s$ instead of $m_R(s)$. Then the *fuzzy probability* $p_s$ of the profile $s$ is defined by

$$P_s = \frac{m_s}{\sum_{s \in U^3} m_s} \quad (5)$$

However, according to the British economist Shackle [4] and many other researchers after him, the human behaviour can be better studied by using the possibilities rather of the several profiles, than their probabilities. The *possibility* $r_s$ of the profile $s$ is defined by

$$r_s = \frac{m_s}{\max\{m_s\}} \quad (6).$$

In equation (6) $\max\{m_s\}$ denotes the greatest value of $m_s$ for all $s$ in $U^3$. In other words the possibility of $s$ expresses the "relative membership degree" of $s$ with respect to $\max\{m_s\}$.

The following application to the process of *learning* a subject matter in the classroom illustrates the applicability of the present model to real life situations:

**Example 1:** There is no doubt that learning is one of the fundamental components of the human cognitive action. There are very many different theories and models developed by psychologists, educators and other cognitive scientists for the description of the mechanisms of learning, Nevertheless, although the process of learning differs in details from person to person, it is in general accepted that it involves *representation* and *interpretation* of the input data in order to produce the new knowledge (step $S_1$), *generalization* of this knowledge to a variety of situations (step $S_2$) and *categorization* of the generalized knowledge by embodying it to the individual's appropriate cognitive structures, widely termed as *schemas of knowledge* (step $S_3$). In this way the individual becomes able to derive from memory the suitable in each case piece of knowledge for facilitating the solution of related composite and complex problems (e.g. see [5]).

On the other hand, the process of learning is usually connected with uncertainty and vagueness. In fact, the learner is in many cases not sure about the good understanding of a new concept or topic and also the teacher is in doubt about the degree of acquisition of a new subject matter by students. Consequently, the use of principles of the FS theory could be a valuable tool in the effort of a more effective description of the mechanisms of learning.

The following experiment took place some time ago at the Graduate Technological Educational Institute of Western Greece, in the city of Patras, during the teaching (in three teaching hours) of the definite integral to a group of 35 students of the School of Management and Economics.

In the instructor's short introduction, during the first teaching hour, the concept of the definite integral was introduced through the need of calculating the area between a curve and the x-axis, but the fundamental theorem of the integral calculus, connecting the indefinite with the definite integral of a continuous in a closed interval function, was stated without proof. Then the students were left to work alone on their papers and the instructor was inspecting their efforts and reactions giving from time to time the proper hints and instructions. His intension was to help students to understand the basic methods of calculating a definite integral in terms to the already known methods for the indefinite integral (step $S_1$ of the model).

It was observed that 17, 8 and 10 students respectively achieved intermediate, high and very high understanding of the new subject. In other words, in terms of the model one obtains that $n_{ia}=n_{ib}=0$, $n_{ic}=17$, $n_{id}=8$ and $n_{ie}=10$. Therefore the step of representation-interpretation of the process of learning can be represented as a FS in U in the form

$$A_1 = \{(a, 0), (b, 0), (c, \tfrac{17}{35}), (d, \tfrac{8}{35}), (e, \tfrac{10}{35})\}.$$

At the second teaching hour a series of exercises involving the calculation of improper integrals as limits of definite integrals and of the area under a curve (or among curves) was given to students for solution. The target in that case was to help students to generalize the new knowledge to a variety of situations (step $S_2$ of the model). Working in the same way as above it was found that the step of generalization can be represented as a FS in U in the form

$$A_2 = \{(a, \tfrac{6}{35}), (b, \tfrac{6}{35}), (c, \tfrac{16}{35}), (d, \tfrac{7}{35}), (e, 0)\}.$$

At the third teaching hour a number of composite problems was forwarded to students for solution, involving applications to economics, such as the calculation of the present value in cash flows, of the consumer's and producer's surplus resulting from the change of prices of a given good, of probability density functions, etc ([6], Chapter 17). The target this time was to help students to relate the new information to their existing schemas of knowledge (step $S_3$ of the model). In that case it was found that the step of categorization can be represented as a FS in U in the form

$$A_3 = \{(a, \tfrac{12}{35}), (b, \tfrac{10}{35}), (c, \tfrac{13}{35}), (d, 0), (e, 0)\}.$$

Then the membership degrees of all student profiles involved in the fuzzy relation (3) were calculated. For example, for $s = (c, b, a)$ one finds that $m_s = m_{A_1}(c). m_{A_2}(b). m_{A_3}(a) = \frac{17}{35} x \frac{6}{35} x \frac{12}{35} \approx 0.029$.

It turns out that the profile $(c, c, c)$ possesses the greatest membership degree, which is equal to 0.082. Therefore the possibility of each profile $s$ is calculated by $r_s = \frac{m_s}{0.082}$. For example the possibility of $(c, b, a)$ is equal to $\frac{0.029}{0.082} \approx 0.353$, while the possibility of $(c, c, c)$ is equal to 1, etc.

The total number of the student profiles is obviously equal to the total number of the ordered samples with replacement of three objects taken from five, i.e. equal to $5^3$. Among all those profiles the profiles possessing non zero membership degrees and their possibilities are presented in Table 1.

**Table 1:** Student profiles with non zero membership degrees

| $A_1$ | $A_2$ | $A_3$ | $m_s$ | $r_s$ |
|---|---|---|---|---|
| *c* | *c* | *c* | 0.082 | 1 |
| *c* | *c* | *a* | 0.076 | 0.927 |
| *c* | *c* | *b* | 0.063 | 0.768 |
| *c* | *a* | *a* | 0.028 | 0.341 |
| *c* | *b* | *a* | 0.028 | 0.341 |
| *c* | *b* | *b* | 0.024 | 0.293 |
| *d* | *d* | *a* | 0.016 | 0.195 |
| *d* | *d* | *b* | 0.013 | 0.159 |
| *d* | *d* | *c* | 0.021 | 0.256 |
| *d* | *a* | *a* | 0.013 | 0.159 |
| *d* | *b* | *a* | 0.013 | 0.159 |
| *d* | *b* | *b* | 0.011 | 0.134 |
| *d* | *c* | *a* | 0.031 | 0.378 |
| *d* | *c* | *b* | 0.026 | 0.317 |
| *d* | *c* | *c* | 0.034 | 0.415 |
| *e* | *a* | *a* | 0.017 | 0.207 |
| *e* | *b* | *b* | 0.014 | 0.171 |
| *e* | *c* | *a* | 0.039 | 0.476 |
| *e* | *c* | *b* | 0.033 | 0.402 |
| *e* | *c* | *c* | 0.042 | 0.512 |
| *e* | *d* | *a* | 0.025 | 0.305 |
| *e* | *d* | *b* | 0.021 | 0.256 |
| *e* | *d* | *c* | 0.027 | 0.329 |

All the above calculations have been made with accuracy up to the third decimal point. The fuzzy data of Table 1 give not only quantitative information, but also a qualitative view of the student behaviour in the classroom during the learning process. This is obviously very useful to the instructor for organizing his/her future teaching plans.

## 3 Combined Results of Fuzzy Data

Frequently in practical applications it becomes necessary to study the *combined results* of the behaviour of *k* different groups of a system's components, $k \geq 2$, during the same process (e.g. the combined performance of two or more student classes in solving the same problems).

For measuring the degree of evidence of the combined results of the *k* groups, it is necessary to define the *combined probability p(s)* and the *combined possibility r(s)* of each profile *s* with respect to the membership degrees of *s* in all the groups involved. The values of *p(s)* and *r(s)* can be defined with respect to the *pseudo-frequency*

$$f(s) = \sum_{t=1}^{k} m_s(t) \quad (7)$$

and they are equal to

$$p(s) = \frac{f(s)}{\sum_{s \in U^3} f(s)} \quad (8)$$

and

$$r(s) = \frac{f(s)}{\max\{f(s)\}} \quad (9)$$

respectively, where *max{f(s)}* denotes the maximal pseudo-frequency.

Obviously the same procedure could be applied if one wanted to study the combined results of the behaviour of a single group during *k* different activities (e.g. the combined performance of a student class during the solution of two or more different problems).

The following example on a *market's research* illustrates the importance of the above procedure:

**Example 2:** A company performed a market's research about the degree of the consumer preference for its negotiable products, which was characterized by the fuzzy linguistic labels *a, b, c, d, e* defined in Section 2. The research was performed separately for men and women and for three different categories of age, namely $C_1$: 18-30 years, $C_2$: 31-50 years and $C_3$: over 50 years old.

Denote by $A_1$ (t), $A_2$ (t) and $A_3$ (t) respectively the FSs representing the consumers' degree of preference for each of the above three categories of age, where the variable t takes the values t = 1 for men and t = 2 for women. Such kind of FSs, whose entries depend on the values of a variable, are usually referred as *fuzzy variables.*

According to the collected data the FSs $A_i$ (t), for i = 1, 2, 3 and t = 1, 2 were found to be as follows:

$A_1$ (1) = {(*a*, 0), (*b*, 0), (*c*, 0.486), (*d*, 0.228), (*e*, 0.286)}

$A_2$ (1) = {(*a*, 0.171), (*b*, 0.171), (*c*, 0.4), (*d*, 0.257), (*e*, 0)}

$A_3$ (1) = {(*a*, 0.343), (*b*, 0.0286), (*c*, 0.371), (*d*, 0), (*e*, 0)}

$A_1$ (2) = {(*a*, 0), (*b*, 0.2), (*c*, 0.5), (*d*, 0.3), (*e*, 0)}

$A_2$ (2) = {(*a*, 0.2), (*b*, 0.267), (*c*, 0.533), (*d*, 0), (*e*, 0)}

$A_3$ (2) = {(*a*, 0.4), (*b*, 0.3), (*c*, 0.3), (*d*, 0), (*e*, 0)}.

In this example the degree of the customer preferences in each age category does not depend on the previous categories. Therefore the calculation of the membership degrees of all the customer profiles is done by the product law of equation (4). For example, for the profile *s* = (*c, c, a*) one finds that

$m_s(1) = 0.486 \text{ x } 0.4 \text{ x } 0.343 \approx 0.67$ and

$m_s(2) = 0.5 \text{ x } 0.5 \text{ x } 0.33 \approx 0.107$.

It turns out that the above profile has the greater pseudo-frequency *f(s)* = 0.67 + 0.107 = 0.174 and therefore its combined possibility is equal to 1, while the combined possibilities of all the other profiles are calculated by $r(s) = \frac{f(s)}{0.174}$.

The membership degrees, the pseudo-frequencies and the combined possibilities of all the customer profiles with nonzero pseudo-frequencies are presented in Table 2.

**Table 2:** Customers' profiles with non zero pseudo-frequencies

| $A_1$ | $A_2$ | $A_3$ | $m_s(1)$ | $m_s(2)$ | $f(s)$ | $r(s)$ |
|---|---|---|---|---|---|---|
| *b* | *b* | *b* | 0 | 0.016 | 0.016 | 0.092 |
| *b* | *a* | *b* | 0 | 0.012 | 0.012 | 0.069 |
| *b* | *c* | *b* | 0 | 0.032 | 0.032 | 0.184 |
| *b* | *b* | *a* | 0 | 0.021 | 0.021 | 0.121 |
| *b* | *b* | *c* | 0 | 0.016 | 0.016 | 0.092 |
| *b* | *a* | *a* | 0 | 0.016 | 0.016 | 0.092 |
| *b* | *a* | *c* | 0 | 0.012 | 0.012 | 0.069 |
| *b* | *c* | *a* | 0 | 0.042 | 0.042 | 0.241 |
| *b* | *c* | *c* | 0 | 0.032 | 0.032 | 0.184 |
| *c* | *c* | *c* | 0.072 | 0.080 | 0.152 | 0.874 |
| *c* | *a* | *c* | 0.082 | 0.030 | 0.112 | 0.644 |
| *c* | *b* | *c* | 0.031 | 0.040 | 0.071 | 0.408 |
| *c* | *d* | *c* | 0.046 | 0 | 0.046 | 0.264 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *c* | *c* | *a* | 0.067 | 0.107 | 0.174 | 1 |
| *c* | *c* | *b* | 0.056 | 0.008 | 0.064 | 0.368 |
| *c* | *a* | *a* | 0.028 | 0.040 | 0.068 | 0.391 |
| *c* | *a* | *b* | 0.024 | 0.030 | 0.054 | 0.310 |
| *c* | *b* | *a* | 0.028 | 0.053 | 0.081 | 0.466 |
| *c* | *b* | *b* | 0.024 | 0.040 | 0.064 | 0.368 |
| *c* | *d* | *a* | 0.043 | 0 | 0.043 | 0.247 |
| *c* | *d* | *b* | 0.036 | 0 | 0.036 | 0.207 |
| *d* | *d* | *a* | 0.020 | 0 | 0.020 | 0.115 |
| *d* | *d* | *b* | 0.017 | 0 | 0.017 | 0.098 |
| *d* | *d* | *c* | 0.022 | 0 | 0.022 | 0.126 |
| *d* | *a* | *a* | 0.013 | 0.024 | 0.037 | 0.213 |
| *d* | *a* | *b* | 0.011 | 0.018 | 0.029 | 0.167 |
| *d* | *a* | *c* | 0.015 | 0.018 | 0.033 | 0.190 |
| *d* | *b* | *a* | 0.013 | 0.032 | 0.045 | 0.259 |
| *d* | *b* | *b* | 0.011 | 0.024 | 0.035 | 0.201 |
| *d* | *b* | *c* | 0.014 | 0.024 | 0.038 | 0.218 |
| *d* | *c* | *a* | 0.031 | 0.064 | 0.095 | 0.546 |
| *d* | *c* | *b* | 0.026 | 0.048 | 0.074 | 0.425 |
| *d* | *c* | *c* | 0.034 | 0.048 | 0.082 | 0.471 |
| *e* | *a* | *a* | 0.017 | 0 | 0.017 | 0.098 |
| *e* | *a* | *b* | 0.014 | 0 | 0.014 | 0.080 |
| *e* | *a* | *c* | 0.018 | 0 | 0.018 | 0.103 |
| *e* | *b* | *a* | 0.017 | 0 | 0.017 | 0.098 |
| *e* | *b* | *b* | 0.014 | 0 | 0.014 | 0.080 |
| *e* | *b* | *c* | 0.018 | 0 | 0.018 | 0.103 |
| *e* | *c* | *a* | 0.039 | 0 | 0.039 | 0.224 |
| *e* | *c* | *b* | 0.033 | 0 | 0.033 | 0.190 |
| *e* | *c* | *c* | 0.042 | 0 | 0.042 | 0.241 |
| *e* | *d* | *a* | 0.025 | 0 | 0.025 | 0.144 |
| *e* | *d* | *b* | 0.021 | 0 | 0.021 | 0.121 |
| *e* | *d* | *c* | 0.027 | 0 | 0.027 | 0.155 |

The above calculations have been made again with accuracy up to the third decimal point. The fuzzy data of Table 2 give to the company a detailed idea about the consumers' preferences for its products.

# 4 Conclusion

The management and evaluation of the fuzzy data obtained by the operation mechanisms of large and complex systems is very important for real life and science applications. A model has been developed in the present work for evaluating such kind of data in terms of the corresponding membership degrees and possibilities. Examples were also presented, for the process of learning a subject-matter in the classroom and for a market's research, illustrating the applicability and usefulness of the model to practical problems.

The general character of the proposed model enables its application to a variety of other human and machine activities (e.g. see the book [7] for a description of such kind of activities) and this is one of our main targets for future research.

*References:*

[1] Zadeh, L.A., Fuzzy Sets, *Information and Control*, 8, 338-353, 1965.

[2] Voskoglou, M. Gr., Multi-Valued Logics: A Review, *International Journal of Applications of Fuzzy Sets and Artificial Intelligence*, 9, 5-12, 2019.

[3] Klir, G. J. & Folger, T.A. , *Fuzzy Sets, Uncertainty and Information*, London, Prentice Hall Int., 1988.

[4] Shackle G.L.S., *Decision, Order and Time in Human Affairs*, Cambridge, Cambridge Univ. Press, 1961.

[5] Voss, J.F., Learning and transfer in subject matter learning: A problem solving model, *Int. J. Educ. Research*,11, 607-622, 1987.

[6] Dowling E.T., *Mathematics for Economists,* Schaum's Outline Series, New York, Mc Graw – Hill, 1980.

[7] Voskoglou, M. Gr., *Finite Markov Chain and Fuzzy Logic Assessment Models: Emerging Research and Opportunities*, Createspace.com. – Amazon, Columbia, SC, USA, 2017.