Fuzzy Models in Evaluation of Information Uncertainty in Engineering and Technology Applications

Leon REZNIK
School of Communications & Informatics
Victoria University of Technology
P.O. Box 14428, Melbourne City MC VIC 8001 Australia
Email: Leon.Reznik@vu.edu.au

Binh PHAM
Faculty of Information Technology
Queensland University of Technology
GPO Box 2434 Brisbane QLD 4001 Australia
Email: b.pham@qut.edu.au

Abstract

The paper studies the problem of information uncertainty evaluation in modern engineering and technology applications and especially system design. It analyses virtual environment design and engineering measurement. Information typical for those applications is classified according to its uncertainty types. Uncertainty sources are identified. Fuzzy theory models are proposed. Examples of their applications in characteristic problems are given.

I. INTRODUCTION

Vigorous development of Internet, computer graphics and animation and other information technologies has significantly expanded the range of volumes and sources of information available for decision making. On the other hand, in order to reach a high-quality decision in business, engineering and social applications nowadays one has to fuse information of different kinds (e.g. numerical, statistical, textual, visual, audio) from a variety of sources (e.g. engineering measurement systems, expert's opinions, images, sound tracks). Due to their nature, these sources differentiate in reliability and uncertainty of the information produced. Uncertainty also can be influenced by the characteristics of procedures and tools applied in information acquisition and processing. As the information sources incorporate both engineering and human as well as heterogeneous systems, the characteristics mentioned above might include psychological personal profiles and emotional behaviour, which IT in the past had trouble handling.

As nowadays information delivery includes the engagement of complex (including hardware, software, and a component now often called brainware) IT systems, the open availability of credible measurement and test methods is an important step toward assuring the quality of such information and promoting competitiveness in the information technology market. Many national and international organizations are now working toward reaching this goal, including the US National Institute of Standards and Technology (NIST), national metrology institutes in the European Community and Japan, professional organizations such as IEEE and ACM,and industry groups such as Open Group (X/Open) and Underwriters Laboratories [1].

Uncertainty is the main challenge for the fusion of a variety of information, both in how to reduce the degree of uncertainty and in how to describe the uncertainty that inevitably remains. The problem of estimating uncertainty (or reliability, imperfectness, impreciseness) of the information source and uncertainty of the information after its propagation and fusion with other information streams has become very important for decision making in different IT and engineering applications, especially in system design.

II. INFORMATION CLASSIFICATION AND UNCERTAINTY

According to L.Zadeh all the information commonly available in system design can be classified into three groups:

1) factual information which is numerical and measurement-based;
2) pseudo-measurement based and pseudo-numerical information (e.g. "checkout time is 11.00");
3) perception based information which is mainly linguistic or could be image and sound-based (e.g. "Robert is an honest and nice-looking person").

Those three groups are commonly supposed to differ in the degree of uncertainty, which the corresponding information has. However, does it mean that those groups have a fundamental, ontological differences as one traditionally presumes?

III. UNCERTAINTY EVALUATION IN MODERN APPLICATIONS

Uncertainty in the last two information groups occur in many important applications, among them are:

- Virtual environment
- System identification, modelling, and control
- Internet, networks and software testing and certification

The ways of dealing with uncertainty differentiate depending on the application area and tools traditionally applied within that field. To date, in most cases, information used in virtual environment design is often assumed to be of a precise (or crisp) nature. This assumption simplifies the representation of the models and their management. However, there are many situations where imprecise data exist and need to be catered for. The impreciseness may concern the *spatial* and *temporal* aspects of objects, their *attributes* or their *relationships* with each other, or with the environment itself. The impreciseness may come from different sources. It may be due to vagueness in human interpretation or intent, or to be given in natural language or visual and audio terms (e.g. locating an object that is *rather* oblong; placing an object *near* another object; activating an event *just after* another event). These situations occur when exactness is not required or is not possible to specify; or when exactness would unnecessarily limit options. Impreciseness may also be due to *missing* data (e.g. when an object is occluded from another), or uncertainty in *taxonomic* definition (e.g. is this a group of trees or a forest?). Some information may be provided in visual or audio forms. In addition, in group decision making process, there are fuzziness in both human preferences and the concept of majority (e.g. one option is *slightly* preferred than another option, *almost all* participants agree with this alternative).

As a virtual environment or system changes, such impreciseness also *propagates* and can exert serious effects on the operation and management of the environment or the system. Exact or crisp models when used in these cases would not be able to faithfully simulate the richness and subtlety of the real work space. Furthermore, in the worst case, misleading or incorrect outcomes may result. Thus, there is a need to articulate the problems arisen from such imprecise data with the view to construct appropriate models for their representation, as well as suitable methods for processing and manipulating them within the environment.

On the contrary, in engineering design and modelling since the early 90's and in particular since the publication of the International Standards Organisation (ISO) "Guide to the Expression of Uncertainty in Measurement" [3], there has been a widening recognition that uncertainty of measurement is no less critical than the value of the measurement result itself. So far, the procedures of uncertainty evaluation have been developed and formalised up to the level of national and international standards for the first information group. The methods of reducing uncertainty, such as calibration, are known as well. Those standards mainly apply methods based on probabilistic models and mathematical statistics, that prevents their application to the second and third information groups. However, there is no doubt that fusion of information from all available sources would significantly improve the quality of decisions made in system design process. Furthermore, these standards have been widely criticised during last years for their inability to cover all sources of uncertainty and it has been suggested that methods based on soft computing and fuzzy logic would be more suitable [4,5]. However, models and methodology for doing so have not been constructed yet.

IY. FEATURES OF THE PROPOSED MODELS

The methods of uncertainty evaluation presented in this paper are based on the same principles as applied in physical measurement science and stated in the ISO Guide [3] and apply fuzzy sets and logic models. Since its introduction in 1965, fuzzy logic has been applied for different purposes. However, the basic idea is its use for modelling "human like" logic and information. So it is prima-facie a most appropriate way to formalise and process an expert's information especially expressed in a natural language. Recent developments of fuzzy logic theory such as fuzzy constraints theory and calculus, possibility theory and especially the new theory of fuzzy information granulation [2] have prepared a necessary framework for a joint application of fuzzy and probabilistic models in measurement science. Fuzzy models calculus has been developed as well (see, for example, [7]).

Y. IDENTIFICATION OF UNCERTAINTY SOURCES AND THEIR CLASSIFICATION FOR APPLICATIONS

Table on fig.1 displays a list of uncertainty sources and feasible corresponding models [3,4].

In a virtual environment, uncertainty may occur in:

- Spatial location of objects
- Attributes of objects (e.g. shape, colour, visibility, personal characteristics)
- Attributes of environments (e.g. area, volume, boundary)
- Attribute relationship (e.g. more visible, better looking)
- Spatial relationships (e.g. direction, occlusion, topology)
- Typological attributes (e.g. group of objects, sub-objects)
- Temporal attributes and relationships (when an event occurs, and with respect to other events).

	Uncertainty source	Comments
A	Incomplete definition of the object or value	Incomplete definition may be caused by an impossibility or difficulty to compile an exact functional relationship or by the application of linguistic forms and rules. Fuzzy models may be particularly relevant here, also statistical depending on source.
B	Imperfect realisation of a definition	Could be due to a number of factors, statistical models relevant to physical limitations of experiment, fuzzy models relevant to conceptual limitations.
C	Non-representative sampling	The sample may not represent the defined object because of a limited sample size, inhomogeneity of the object under measurement, etc. Sampling implies an underlying distribution and refers primarily to statistical models, but to the degree that the sample may not be representative, may involve fuzzy models.
D	Inadequate knowledge of the effects of the environment	1)the model of environmental conditions may not cover all the influence factors or 2) the model may be made under slightly different conditions because of the environmental changes statistical modelling 3) the model may be based on expert's estimates or guesses, fuzzy and statistical models
E	Personal bias in description	May be difficult to model mathematically as it depends on a particular person, may vary with time, etc. Applicable to both fuzzy and statistical models.
F	Limited accuracy (high inaccuracy) of the information available	Applicable to both statistical and fuzzy models.
G	Inexact values of standards and reference materials	Can be reference sample or reported value from a real underlying distributions (statistical model)
H	Inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm	Usually uncertainty which is due to this reason has a relatively small value in comparison to other components. In some cases, it can be single value from literature references (statistical square distribution or fuzzy model).
I	Approximations and assumptions incorporated in the design method and procedure	Difficult to evaluate with statistical methods, here fuzzy modelling approaches may be particularly relevant.

Figure 1. Classification of uncertainty sources and feasible models

YI. REPRESENTATION OF UNCERTAIN INFORMATION AND UNCERTAINTY ESTIMATION IN VIRTUAL ENVIRONMENT DESIGN

To represent fuzziness for two dimensional spatial objects, we extend the concept of using intervals to that of bounding rectangles which completely enclose the object and whose sides are parallel to x and y axes. We use this definition instead of a bounding rectangle whose sides are parallel to the major axes of an object because although the latter represents more closely the shape of an object, the former is much faster for processing. *Minimum bounding rectangle* (MBR) which is the smallest of such rectangles has been often used to speed up the location of 2D objects and checking of object intersection and containment (see, for example, [6]). We therefore use MBBs for determining whether an object satisfies certain topological relationship. This concept can be extended to *minimum bounding boxes* (MBB) for similar purposes in 3D virtual worlds, e.g. to detect collision. If the MBBs of two objects are disjoint, then the objects are disjoint and no further work is necessary.

To represent *topological* and *directional* relationships, an *abstract spatial graph* will be used, which maps a volume into a point represented by a triple (rx,ry,rz), where rx,ry and rz are intervals along x,y and z direction. In this case, the MBB of an object or subobject is the volume of interest. Thus, an abstract spatial graph is constructed where each node is a MBB of an object in the virtual world. A weight will be added to each node to indicate the degree of participation in each relationship. This weight needs to be defined to suit the context of the relationship being considered. For example, for overlapping relationship, the weight can be computed as the ratio of the volume of a subobject to the volume of the MBB of the entire object. on the other hand, for a directional relationship, it would be more logical to define the weight to be the ratio of the extent of an object to the extent of the whole MBB along the direction of interest. Such weights can also be expressed in linguistic terms based on appropriate intervals (e.g. mostly overlapped for 70-98%, somewhat overlapped 30-70% and little overlapped 5-30%). A better option, however, is to express these weights in terms of fuzzy membership functions of a fuzzy set. These membership values (between 0 and 1) denote the semantic confidence of fuzzy linguistic terms. The main advantage of using fuzzy set is that impreciseness in data is retained as long as possible, hence the final decision when it is made, would give a more accurate result than in the case where a crisp approximation for imprecise data is used in every step leading to the final result. Due to page limitation, basic methods for construction of a fuzzy system is not covered here, they may be found in many text books.

Fuzzy sets will also be used to represent the fuzziness of other attributes of objects including typological attributes. For example, an object may be described as slightly twisted, moderately twisted or mostly twisted. Three functions can be formed to describe the fuzzy membership with respect to the parameter that controls the degree of twistness. Similarly, the degree in which an object may belong to a specific category (e.g. most likely, moderately likely or slightly likely) can also be expressed as fuzzy membership functions.

One major issue in virtual environments is how to manage the group decision making process. The impreciseness in this case occurs in two aspects: *fuzzy human preferences* (e.g. participant A slightly prefers option A to option B), and *fuzzy majority* (e.g. *most* instead of a rigid rule such as *at least* 2/3). Again, we will use fuzzy sets to represent both of these concepts and investigate appropriate *aggregation, implication* and *defuzzification* rules that can combine these preferences, deduce causes and effects, and finally derive a crisp value from fuzzy values in a meaningful way.

YII. REPRESENTATION OF UNCERTAIN INFORMATION AND UNCERTAINTY ESTIMATION IN MEASUREMENT SCIENCE AND SYSTEM DESIGN

According to a real-life measurement practice, the metrological characteristics can be given either as the limits of the allowed errors (deterministic form) or as the limits of the allowed values for some probabilistic and statistical characteristics (mean value, random component dispersion, confidence intervals). In both cases, a user does not really know how far from the given limits the actual values lay. Moreover, according to a number of standards the limits should be chosen from a given scale. Because of this, the limits are rounded down with the requested values being sometimes significantly lower than actually required. The probabilistic methods pretend to be objective. However, one can see that an application of the probabilistic and statistical methodologies in metrological analysis includes an assignment of a number of values such as a confidence level, an error probability, the significance level, etc. All those values are not calculated but assigned by an expert, who authored the corresponding document, or a person who performed metrological testing. All of these actually mean that real data about error characteristics are fuzzy in their nature. However, a user is recommended to consider those data as having probabilistic characteristics of some general aggregates, which are in turn assigned some probability distributions.

Considering all the information mentioned above the idea of measurement error formulation in terms of fuzzy systems theory looks rather reasonable. Some steps in this direction have been already made. In a number of the international [3,8,9] and national standards, the term "measurement error" has been replaced with the term "measurement uncertainty", which can be considered as more correspondent to fuzzy systems terminology. Publications, criticising the probabilistic models applied in measurement science, are now followed by a number of works, trying to formulate those models from the fuzzy theory point of view or to combine both theories [10-17]. For example, in [18,19] a priori fuzzy information about the object under measurement is applied to increase the measurement accuracy and/or reliability. In order to apply the fuzzy sets and systems methodology in metrology and measurement practice, one has to prove that this methodology is able to perform mathematical and logical operations with fuzzy values, intervals and functions, typical for measurement science and practice. Let us give a few examples illustrating a possibility of such model applications in measurement science.

Example 1 (see fig.2) demonstrates the result of nonlinear division operation applied to the fuzzy variable. The initial fuzzy variable is given by the membership function μ_A which has a flat top in the range between 1.8 and 2.2. The support set for this fuzzy variable is the interval (1.5, 2.5). The gravity centre for this membership function is 2.0. The measurement result which equals 2.0 can be modelled with this membership function. The result error has an unexcluded systematic component, which extends the interval of possible values to (1.8, 2.2) and a random component, which fuzzifies this interval boundaries. The division result is a fuzzy set with the membership function μ_B. Its gravity centre equals 0.512. Nonsymmetry and a bias are caused by the transformation nonlinearity. The boundaries of the corresponding support interval are (0.4, 0.64). One can see that with the application of fuzzy sets theory any mathematical operation produces some information not only about its result itself but some characteristics of its uncertainty as well. Such information may be given in a detailed form with the membership function or in a reduced form with a confidence interval.

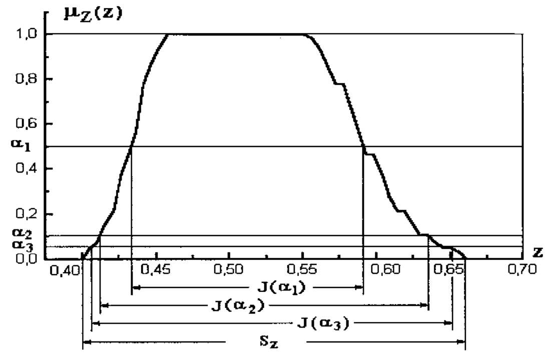

Figure 2. Example of a membership function with α-cuts marked

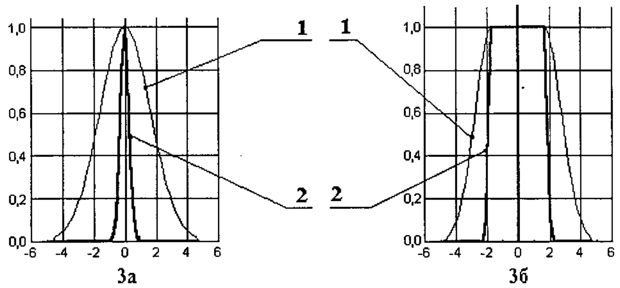

Figure 3. Examples of fuzzy variables averaging

Example 2 (see fig. 3a) illustrates an averaging operation of 36 measurement results. Each measurement is described with the unimodal membership function which has a shape of a gaussian distribution function. This sample of measurement results represents a case of excluded systematic errors. An application of the fuzzy sets methodology to an averaging operation results is the fuzzy set with a membership function 2. One can see that it has the same shape as initial functions but with the width 6 times narrower. Numerically, it coincides with the results produced by the statistical methods application.

Example 3 (see fig. 3b) is similar to the previous one. However, in this example a not excluded systematic error, which is given by the interval of possible values according to modern norms, is simulated. In the example this interval is limited between (-1.9, 1.9). These values serve as the boundaries for the membership function 1 flat top. The membership function 2 gives the averaging result of 36 initial membership functions. One can see that the averaging result has the same systematic error as its membership function has the same top. However, the result function slopes simulating a random error have become 6 times stepper.

These examples demonstrate the efficiency of the fuzzy sets methodology in averaging procedures, which are extremely popular in measurement results mathematical processing. One should note that this methodology allows not to separate systematic and random uncertainty components. One can also suppose that an interval theory may be applied to determine the measurement error characteristics. Further investigation of fuzzy models applications in measurement result processing should result in the development of a theoretically not contradicting and practically useful measurement uncertainty theory, which can cover all stages of measurement information processing.

References

1. Wakid S.A., D.R. Kuhn, and D.R. Wallace Toward Credible IT Testing and Certification, IEEE Software, vol. 16, No.4, 1999, pp. 39-47.
2. Zadeh L.A. Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic, Fuzzy Sets and Systems, vol. 90, No.2, pp.111-127, 1997.
3. Guide to the Expression of Uncertainty in Measurement. International Organisation for Standardisation, Geneva, 1995.
4. Reznik L., K.P. Dabke Evaluation of Uncertainty in Measurement: A Proposal for Application of Intelligent Methods in H. Imai/Ed. Measurement to Improve Quality of Life in the 21st Century, IMEKO – XV World Congress, June 13-18, 1999, Osaka, Japan, vol. II, p.93 – 100.
5. Xia X, Z.Wang and Y.Gao Estimation of Non-statistical Uncertainty Using Fuzzy-set Theory IEE Proceedings on Measurement, Science, and Technology, vol 11, 2000, pp.430-435.
6. Clementini E, Sharma J and Egenhofer M.J. Modelling topological and spatial relations: Strategies for query processing, *Computers and Graphics*, V18 (6), 1994, 815-822.
7. Reznik L. Fuzzy Controllers Butterworth-Heinemann, Oxford, 1997
8. International Vocabulary of Basic and General Standard Terms in Metrology.-ISO, Geneva, Switzerland, 1993
9. Quantifying Uncertainty in Analitical Measurement - Eurachem, 1995
10. Gonella L. Proposal for a Revision of the Measure Theory and Terminology // Alta Frequenza, vol. XLIV. № 10, 1975.
11. Destouches J.L., Fevrier P. New Trends in Expressing Results of Measurements // Proc. of IMEKO Coll., Budapest 1980.
12. Roskrut A. Zapis teorii pomiaru sformulowany w oparciu o teorie zbiorow rozmytych // Postepu technologii maszyn i urzadzen , No.2, 1983 (in Polish).
13. Otto M., Bandemer H. Calibration with Imprecize Signals and Concentrations Based on Fuzzy Theory // Chemometrics and Intellect. Laboratory Systems, № 1, 1986.
14. Mari L. Notes on Fuzzy Set Theory as a Tool for the Measurement Theory // Digest of the 12th Triennal World Congress IMEKO, vol. III, Beijing, China.
15. Mari L. The Role of Determination and Assignment in Measurement , Measurement, vol.21, No.3, pp.79-90, 1997
16. Reznik L. Towards Fuzzy Logic Application in Measurement Theory 1st International Conference "Soft Computing and Measurement" (SCM'98), St.Petersburg, Russia, June 22-26, 1998
17. Solopchenko G.N, L.Reznik, W.C.Johnson. Fuzzy Intervals as a Basis for Measurement Theory // Proc. of the First international Joint Conference of the North American Fuzzy Information Processing Society Biannual Conference, the Industrial Fuzzy Control and Intelligent Systems Conference, and the NASA Joint Technology Wofkshop on Neural Networks and Fuzzy Logic, pp. 405 - 406, San Antonio, USA, December 18-21, NASA, 1994
18. Reznik L. Fuzzy Models of a priori Information in the Measurements' Results Processing AMSE REVIEW, 1987, v.4, 2, p.1-12
19. Reznik L., Solopchenko G.N. Use of a priori Information on Functional Relations between Measured Quantities for Improving Accuracy of Measurement // Measurement, 1986, v.3, 2, p.61-69