Biomatemática **33** *(2023), 103–114* **ISSN 1679-365X**
A publication of Biomatematics group IMECC – Unicamp

Alternatives for solving an inaccurate data problem

Rodney C. Bassanezi[1],
IMECC, Unicamp - 13.083-859, Campinas/SP.

Abstract. *If there were no uncertainty the world would be very boring.* This paper presents a simple problem that cannot be solved with an exact solution, and also it does not have an approximate solution, coming from a stochastic process, because we have not collected a database. Mathematics based on fuzzy logic is the instrument that justifies, in a way, the intuitive resolution to some ill-posed problems.

Key words*: Uncertainty, fuzzy systems, approximate reasoning, rules-based systems.*

1. Introduction

Obtaining an exact answer to certain problems is not always possible, even when the situations presented are very simple. This often occurs when we have a poorly-posed problem, that is, the information is partial or inaccurate.

The objective of this work is to present a simple problem that cannot be solved with an exact solution and also, it does not have an approximate solution, coming from a stochastic process, because we have not collected a database. However, with a little knowledge of the situation, we can always come up with an "intuitive solution". Mathematics based on fuzzy logic is the instrument that justifies, in a way, the intuitive resolution to some ill-posed problems.

2. The problem

Let's consider a very simple problem, which may have a very reasonable intuitive answer, but which cannot be solved using resources provided by classical

[1]rodneyb@unicamp.br

mathematics because the informations or data are inaccurates.

Problem: Time spent on a trip (Barros and Bassanezi, 2015; Jafelice et al., 2012; Martins, 2023).

Determine the time spent on a bus trip from Campinas to S. Paulo which is subject to the following considerations:

- The distance is approximately $100\,km$;
- Speed should not exceed $120\,km/h$;
- Traffic is generally heavy and speed is also reduced at tolls;
- The bus is almost always late and you should buy the ticket well in advance. Waiting time at the bus station is never more than a half hour.

We cannot solve this problem using the process provided by classical mathematics. The formula for calculating the total time is

$$T = T^* + \frac{E}{V}$$

where, T^*: the waiting time, E: the distance to be traveled and V: the speed of the bus. If there were an experimental database, collected from other similar situations, then we could use the stochastic method to obtain an approximate solution to the problem.

An intuitive approach to solving this problem showed that every people who were asked for an answer to the problem gave resposes such "the time should be a little more than an hour" or "between 1 hour and 1 hour and a half" or even "no more than 2 hours". Such answers are based on the personal experiences of those who have already faced similar situations: "These buses run a lot but the traffic is heavy and/or if there are tolls, the speed decreases. In addition, the buses are usually delayed".

The intuitive calculations are then made considering a "reasonable" average speed of $90\,Km/h$ and a departure delay of $15\,min$. In this way, the total time would be $T = T_1 + T_2 = (1\,h$ and $11.1\,min) + 15\,min = 1\,h$ and $26.1\,min \simeq 1.43\,h$.

We will present some alternatives ways, using fuzzy logic in different contexts, to "solving" this problem. This procedure may justify the possible intuitive answers.

3. Fuzzy rules based

A "justification" of the operation, used in the "intuitive resolution" of the problem, can be given through an inference process from a database, considering the following "information" given in Table 1

Table 1: Fuzzy rule-base.

$\mathbf{V} \ddots \mathbf{A}$	$\mathbf{A}_b$	$\mathbf{A}_m$	$\mathbf{A}_a$
$\mathbf{V}_b$	T_a	T_a	T_{al}
$\mathbf{V}_m$	T_m	T_a	T_a
$\mathbf{V}_a$	T_b	T_b	T_m

where, the linguistic variables V, A and T are, respectively, "bus speed", "exit delay time" and "total trip time". The indices b, m, a and al indicate, in each linguistic variable, its qualities, that is: *low*, *medium*, *high* and *very high*.

Each rule R_i is of the form: "*If* V_i *and* A_j *then* T_{ij}", for example,

If the *speed is low* and the *delay is high* **then** the *travel time is very high*;

Linguistic fuzzy variables are given by fuzzy sets, in this case, by fuzzy numbers:

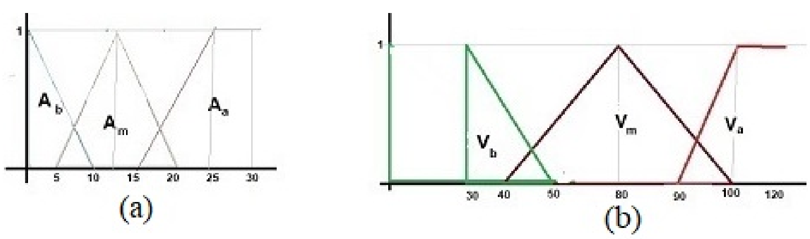

Figure 1: (a) A: exit delay time. (b) V: bus speed.

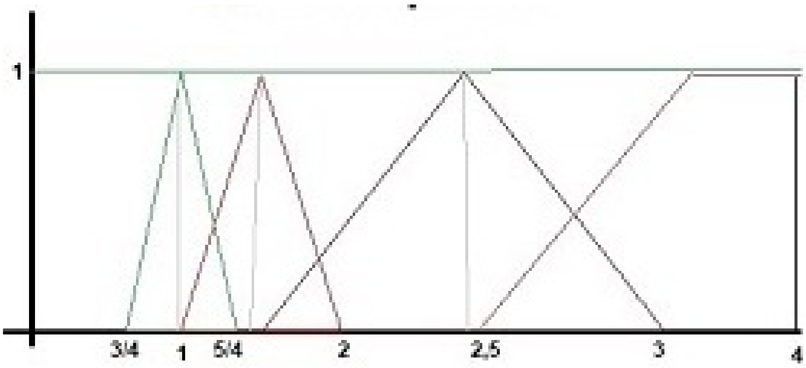

Figure 2: T: total trip time.

The most used inference method is the MAX-MIN, where the t-norm (minimum) is adopted to model the connective "and" and the t-conorm (maximum) to aggregate the fuzzy rules of the rule base. Usually Mamdani-type systems use MAX-MIN as t-norms and t-conorms.

Mamdani's method of inference operates as a weighted average between the input variables, A: *delay* and V: *speed*, to infer an output value for T: *Time spent traveling* (Figures 3 and 4):

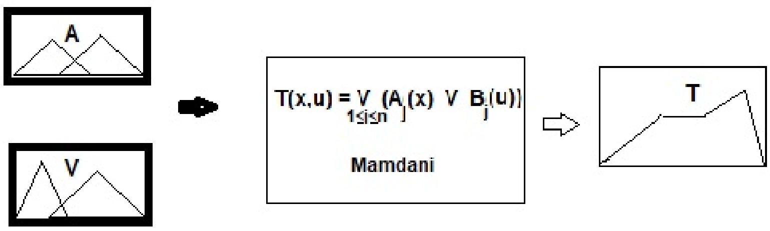

Figure 3: Mamdani inference method (Silva, 2005).

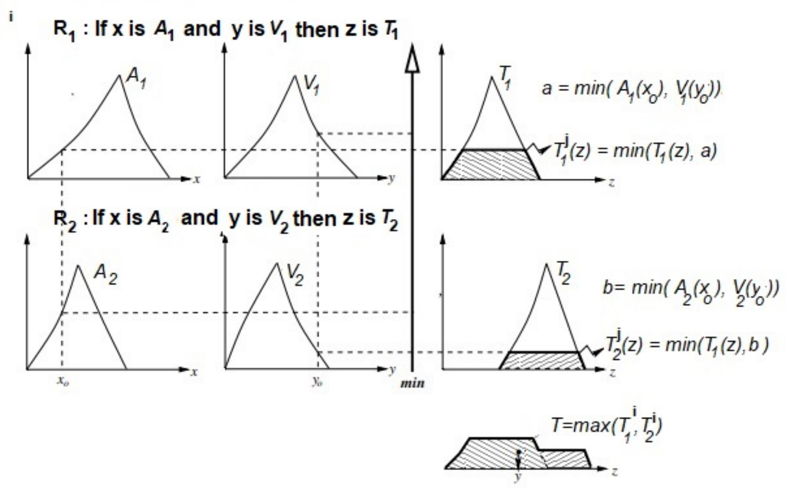

Figure 4: Mamdani inference method.

Figure 5 shows the different values of T for each input pair A_i and V_j:

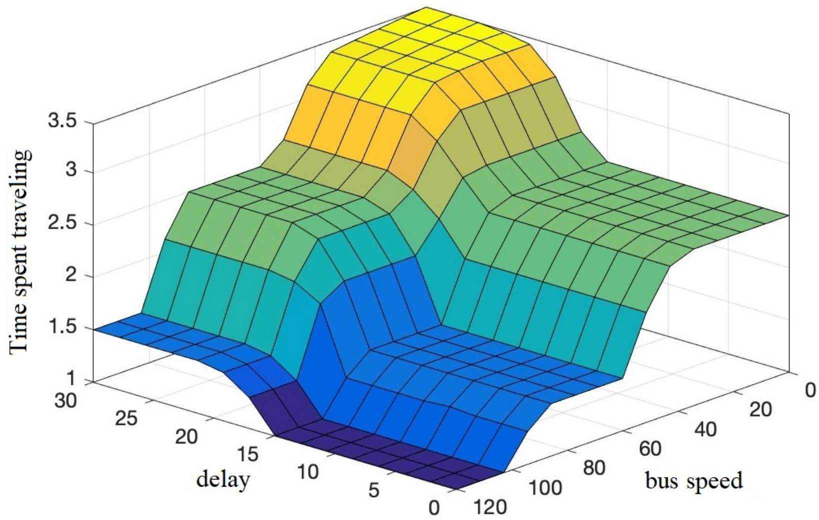

Figure 5: Values of the spent times T for each pair of subjective values of A and V.

Problems like this, which we “solve” based only on our intuition, can be formalized and developed with the help of fuzzy algebra of the number, thus justifying intuitive solutions.

4. Algebra of the fuzzy number

Definition 1. *A fuzzy subset A is called a fuzzy number when the* $\operatorname{supp} A \subset \mathbb{R}$ *and satisfies the following conditions:*

(i) all the α-levels of A are not empty for $0 \leqslant \alpha \leqslant 1$*;*

(ii) all α-levels of A are closed intervals of $\mathbb{R}$*;*

(iii) $\operatorname{supp} A = \{x \in \mathbb{R} : \varphi_A(x) > 0\}$ *is bounded.*

The arithmetic operations involving fuzzy numbers are closed linked to the interval arthmetic operations, that is,

Definition 2. *Let A and B be two fuzzy numbers and λ a real number, and Θ is one of the arithmetic operation between real numbers, then*

$$\varphi_{(A\Theta B)} = \sup_{\{(x,y):x\Theta y=z\}} \min\{\varphi_A(x), \varphi_B(y)\}. \tag{1}$$

Let's now consider the variables involved in the travel problem with being fuzzy numbers and determine the travel time $T = A + (D/V)$, where A is the waiting time at the bus station and D/V is the travel time.

Since the travel distance data is also approximate, we can consider it a fuzzy number centered on 100. It can be, for example, the triangular number $D = [90; 100; 110]$ whose degree of pertinence function is outlined in Figure 6:

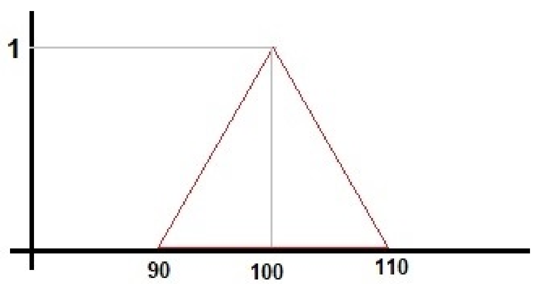

Figure 6: Fuzzy distance between the cities.

- The uncertainty of the speed of the bus can also be modeled by a fuzzy triangular number. Taking into account that it *never exceeds* $120\,km/h$ and that it has some *very low speeds*, we can assume that its function degree of pertinence is:

$$\varphi_D(x) = \begin{cases} 0 & \text{, if } x \leqslant 90 \\ \dfrac{x}{10} - 9 & \text{, if } 90 < x \leqslant 100 \\ 11 - \dfrac{x}{10} & \text{, if } 100 < x \leqslant 110 \\ 0 & \text{, if } x > 110 \end{cases}. \tag{2}$$

- The fact that the bus leaves almost always late indicates that we should also have an extra waiting time that does not exceed $1/2\,h$. This time can be modeled by a fuzzy number of the triangular type $[0; 0; 0.5]$, i.e.,

$$\varphi_{T_1}(x) = \begin{cases} 1 - 2x & \text{, if } 0 \leqslant x \leqslant 0.5 \\ 0 & \text{, if } x > 0.5 \end{cases}. \tag{3}$$

The time spent on the journey is obtained by the fuzzy set

$$T_2 = \frac{D}{V}$$

whose pertinence function is $\varphi_{T_2}(z) = \sup_{\left\{(x,y):\frac{x}{y}=z\right\}} \min\left[\varphi_D(x), \varphi_V(y)\right]$.

We have that α-levels of the T_2 are:

$$[T_2]^\alpha = [10\alpha + 90, 110 - 10\alpha]\left[\frac{1}{120 - 20\alpha}, \frac{1}{70\alpha + 30}\right] \tag{4}$$

and therefore,

$$\begin{aligned} \operatorname{supp} \varphi_{T_2} &= \left[\frac{11}{12}, \frac{9}{7}\right] \\ \varphi_{T_2}(x) &= 1 \Longleftrightarrow x = 100 \text{ e } y = 100. \end{aligned} \tag{5}$$

In addition,

$$[T_1]^\alpha = \frac{1-\alpha}{2};\ 0 \leqslant \alpha \leqslant 1.$$

The total time spent T is given by the delay time T plus the travel time T. The fuzzy number $T = T + T$ is such that its α-levels are given by:

$$\begin{aligned} [T]^\alpha &= \left[0, \frac{1-\alpha}{2}\right] + \left[\frac{10\alpha + 90}{120 - 20\alpha}, \frac{110 - 10\alpha}{70\alpha + 30}\right] \\ &= \left[\frac{10\alpha + 90}{120 - 20\alpha}, \frac{1-\alpha}{2} + \frac{110 - 10\alpha}{70\alpha + 30}\right]. \end{aligned} \tag{6}$$

The closed interval, given by *the support of* T, obtained from (6) when $\alpha = 0$, contemplates all possible solutions of the problem, i.e.,

$$\operatorname{supp} \varphi_T = \left[\frac{3}{4}, \frac{25}{6}\right]. \tag{7}$$

The membership function φ_T can be obtained by the inverses of the functions $f(\alpha) = \dfrac{10\alpha + 90}{120 - 20\alpha}$ e $g(\alpha) = \dfrac{1-\alpha}{2} + \dfrac{110 - 10\alpha}{70\alpha + 30}$ (Figure 6),

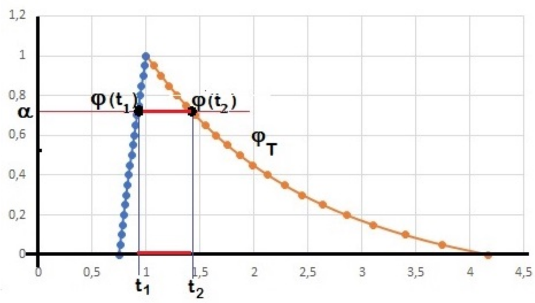

Figure 7: Membership function φ_T.

We observed that the answer $t = 1\,h$ has the higher degree of pertinence on T, that is, $\varphi_T(1) = 1$. The intuitive answer, given initially, $t = 1.43\,h$ would have a degree of pertinence $\varphi_T(1.43) = 0.71$, obtained from

$$1.43 = \frac{1-\alpha^*}{2} + \frac{110 - 10\alpha^*}{70\alpha^* + 30}.$$

In the fuzzy context we could also take a real answer, applying a deffuzzification to the fuzzy set T or considering the fuzzy expectancy of T. There are many methods of defuzzification that can be adopted, however the most used considers the "center of area" or centroid of the fuzzy set, that is

$$G(T) = \frac{\sum_{i=0}^{n} \alpha_i \varphi_T(x_i)}{\sum_{i=0}^{n} \varphi_T(x_i)} \quad \text{ou} \quad G(T) = \frac{\int_{\mathbb{R}} x\varphi_T(x)dx}{\int_{\mathbb{R}} \varphi_T(x)dx}. \tag{8}$$

Observation: If $[T]^\alpha = \{t \in \operatorname{supp} T : \varphi_T(t) \geqslant \alpha\}$ then, for each α with $0 \leqslant \alpha \leqslant 1$, we have on the problem :

$$[T]^\alpha = \left[\varphi_T^{-1}\left(\frac{10\alpha + 90}{120 - 20\alpha}\right), \varphi_T^{-1}\left(\frac{1-\alpha}{2} + \frac{110 - 10\alpha}{70\alpha + 30}\right)\right] = [t_1, t_2].$$

The center of area can be obtained by considering:

$$G(T) \cong \frac{\sum_{0 \leqslant i \leqslant n} \alpha_i(g(\alpha_i) + f(\alpha_i))}{\sum_{0 \leqslant i \leqslant n} \alpha_i(g(\alpha_i) - f(\alpha_i))} \tag{9}$$

that is:

$$G(T) \cong \frac{2.745}{0.993} = 2.764\, h.$$

5. Possibility distribution and fuzzy integral

If we consider the fuzzy set T (6), whose membership is shown in Figure 6, as a distribution of possibilities, we can obtain a defuzzification of T by means of the *fuzzy integral of possibility.* The concept of fuzzy measure was proposed by Sugeno (1974):

Definition 3. *Let $\wp$ be a σ-algebra on $\Omega \neq \varnothing$. A set-valued function $\mu : \wp \to [0,1]$ is a fuzzy measure if*

(i) $\mu(\varnothing) = 0$ *e* $\mu(\Omega) = 1$*;*

(ii) $\mu(A) \leqslant \mu(B)$ *whenever* $A \subseteq B$.

Definition 4. *A possibility distribution on the set Ω is a function $\varphi : \Omega \to [0,1]$ satisfaying* $\sup_{w \in \Omega} \varphi(w) = 1$.

We note that any normal fuzzy subset of Ω can be used to define a possibility distribution on Ω.

Definition 5. *Let $\wp$ be a σ-algebra on $\Omega \neq \varnothing$. A set-valued function $\Pi : \wp \to [0,1]$ is a possibility measure if it satisfies* *(i)* $\Pi(\varnothing) = 0$ *e* $\Pi(\Omega) = 1$*;*

(ii) For any family $\{A_{i \in J}\}$ *of subsets on* Ω *follows*

$$\Pi(\cup_{i \in J}) = \sup_{i \in J} \{\Pi(A_i)\}\,. \tag{10}$$

We note that Π is a fuzzy measure and $\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$ for all pair $A, B \in \wp$.

Definition 6. *Let* $f : \Omega \to [0, 1]$ *a function and* μ *a fuzzy measure on* Ω*. The integral fuzzy of* f *on* Ω *with respect to* μ *is the number*

$$\oint_{\Omega} f d\mu = \sup_{0 \leqslant \alpha \leqslant 1} [\alpha \wedge \mu\{w \in \Omega : f(w) \geqslant \alpha\}]. \tag{11}$$

We note that if $\mu = \Pi$ and φ is a possibility distribution on Ω then (cf. Martins, 2023, Theorem 3.15)):

$$\oint_{\Omega} f d\Pi_{\varphi} = \sup_{x \in \operatorname{supp} f} [f(x) \wedge \varphi(x)]. \tag{12}$$

In relation to the initial *problem of travel time* we can consider the fuzzy set (6) as a distribution of possibilities. Thus, we consider in the fuzzy integral, the pertinence degree function φ_T, whose α-levels are given by $[T]^{\alpha} = \left[\dfrac{10\alpha + 90}{120 - 20\alpha}, \dfrac{1 - \alpha}{2} + \dfrac{110 - 10\alpha}{70\alpha + 30}\right]$.

The fuzzy integral (12) works as a defuzzifier of sets that translate different possibilities. Examples:

(a) If A is the crisp fuzzy set 2, the possibility of the travel time being equal to A is 0.45.

(b) If A is the fuzzy set *approximately* 2, given by $A = (1.75; 2; 2.25)$, the possibility of the travel time A is 0.55 (Figure 7).

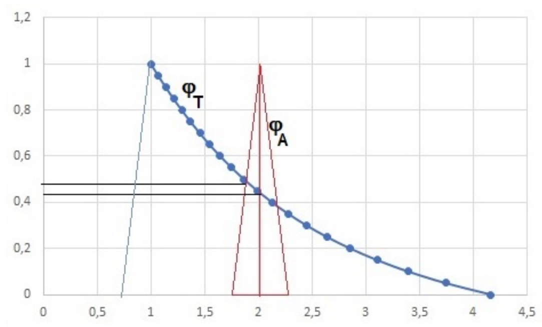

Figure 8: Possibilities of the fuzzy sets travel times.

6. Conclusion

In 1978 Zadeh published the first article on possibility measure (Zadeh, 1999). This article brings an important discussions dealing with the theoretical, semantic, and practical distinctions between possibility and probability. One of these distinction is with respect to an affirmation, seemingly naive, but quite common in day-to-day usage, "a fact is possible but unlikely". This suggests that, whatever concept we use for *possibility* (a measure that we denote by Π), we need have $\Pi(A) \geqslant P(A)$. Currently, this inequality has been much debated and is often called the *Principle of Consistency* (Klir and Yuan, 1995; Barros et al., 2017).

References

Barros, L. C. and Bassanezi, R. C. (2015). *Tópicos de lógica fuzzy e biomatemática*. UNICAMP/IMECC, Campinas, SP, 3rd edition.

Barros, L. C., Bassanezi, R. C., and Lodwick, W. A. (2017). *A first course in fuzzy logic, fuzzy dynamical systems, and biomathematics*. Studies in fuzziness and soft computing. Springer-Verlag Berlin Heidelberg, Berlin, 1st edition.

Jafelice, R. S. M., Barros, L. C., and Bassanezi, R. C. (2012). *Teoria dos conjuntos fuzzy com aplicações*, volume 17 of *Notas em Matemática Aplicada*. Sociedade Brasileira de Matemática Aplicada e Computacional, São Carlos, SP.

Klir, G. J. and Yuan, B. (1995). *Fuzzy sets and fuzzy logic.* Prentice Hall PTR, New Jersey.

Martins, A. (2023). *Integral de Sugeno e o Teorema Fundamental do Cálculo Fuzzy-Aplicações.* PhD thesis, IMECC–Unicamp, Campinas/SP.

Silva, J. D. M. (2005). *Análise de estabilidade de sistemas dinâmicos p-fuzzy com aplicações em biomatemática.* PhD thesis, IMECC–Unicamp, CCampinas/SP.

Zadeh, L. A. (1999). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 100:9–34.