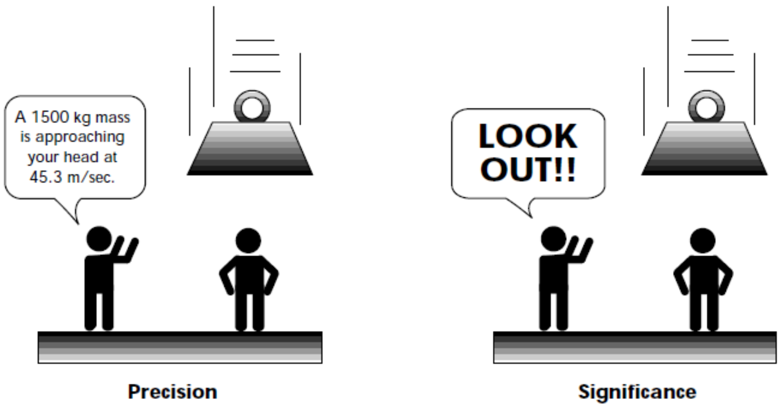
Massachusetts Institute of Technology

Introduction to fuzzy logic



Franck Dernoncourt
franck.dernoncourt@gmail.com

MIT, January 2013

Contents

List of Figures

Chapter 1

Introduction

Fuzzy logic is an extension of Boolean logic by Lotfi Zadeh in 1965 based on the mathematical theory of fuzzy sets, which is a generalization of the classical set theory. By introducing the notion of degree in the verification of a condition, thus enabling a condition to be in a state other than true or false, fuzzy logic provides a very valuable flexibility for reasoning, which makes it possible to take into account inaccuracies and uncertainties.

One advantage of fuzzy logic in order to formalize human reasoning is that the rules are set in natural language. For example, here are some rules of conduct that a driver follows, assuming that he does not want to lose his driver's licence:

If the light is red...	if my speed is high...	and if the light is close...	then I brake hard.
If the light is red...	if my speed is low...	and if the light is far...	then I maintain my speed.
If the light is orange...	if my speed is average...	and if the light is far...	then I brake gently.
If the light is green...	if my speed is low...	and if the light is close...	then I accelerate.

Intuitively, it thus seems that the input variables like in this example are approximately appreciated by the brain, such as the degree of verification of a condition in fuzzy logic.

To exemplify each definition of fuzzy logic, we develop throughout this introductory course a fuzzy inference system whose specific objective is to decide the amount of a tip at the end of a meal in a restaurant, depending on the quality of service and the quality of the food.

1.1 Set theory refresher

> *A set is a Many that allows itself to be thought of as a One.* Georg Cantor.

To begin with, a quick refresher on the classical sets can be useful if you haven't dealt with them for long time.

The classical set theory simply designates the branch of mathematics that studies sets. For example, 5, 10, 7, 6, 9 is a set of integers. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 is the set of integers between 0 and 10. 's' 'd'; 'z', 'a' is a set of characters. "Site", "of", "zero" is a set of words. We can also create sets of functions, assumptions, definitions, sets of individuals (that is to say, a population), etc. and even sets of sets!

Note that in a set, the order does not matter: 7, 6, 9 denotes the same set as 9, 7, 6. However, to improve readability, it is convenient to classify the elements in ascending order, ie 6, 7, 9. Usually, a set is denoted by a capital letter: thus, we write A = 6, 7, 9. The empty set is denoted $\emptyset$: it is a remarkable since it contains no element. This seems unnecessary at first glance, but in fact, we will often use it.

Sets are often represented in graphic form, typically by circles, as figure 1.1 illustrates.

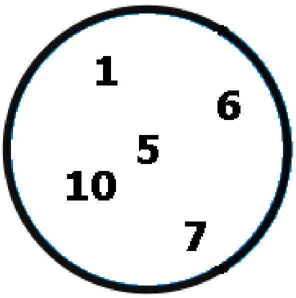


Figure 1.1: Graphical representation of the set $\{1, 5, 6, 7, 10\}$

The concept of belonging is important in set theory: it refers to the fact that an element is part of a set or not. For example, the integer 7 belongs to the set 6, 7, 9. In contrast, the integer 5 does not belong to the set 6, 7, 9. Membership is symbolized by the character in the non-membership and by the same symbol, but barred ossible. Thus, we have $7 \in \{6, 7, 9\}$ and $5 \notin \{6, 7, 9\}$.

A membership function (also called indicator function or characteristic function) is a function that explicit membership or not a set E. Let f be the characteristic function of the set $E = \{6, 7, 9\}$, and x is any integer:
TODO Math formula

This concept of membership is very important for this course because fuzzy logic is based on the concept of fuzzy membership. This simply means that we can belong to a set to 0.8, in contrast to classical set theory where as we have just seen membership is either 0 (not owned) or 1 (part).

In order to manipulate classical ensembles and make something interesting, we define a set of operations, which are very intuitive. Figures 1.2, 1.2 and

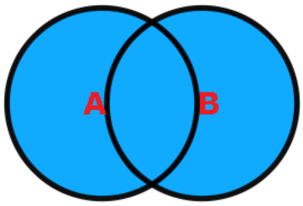


Figure 1.2: Union of two sets, denoted $A \cup B = \{x \in A ext{or} x \ in B\}$. $A \cup B$ corresponds to the blue area. For example if $A = \{6; 7; 9\}$ et $B = \{1; 5; 6; 7; 10\}$, then $A \cup B = \{1; 5; 6; 7; 9; 10\}$.

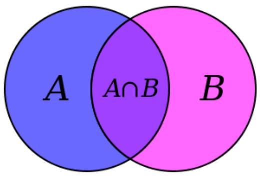


Figure 1.3: Intersection of two sets, denoted $A \cap B = \{x \in A et x \in B\}$. For example, if $A = \{6; 7; 9\}$ and $B = \{1; 5; 6; 7; 10\}$, then $A \cap B = \{6; 7\}$.

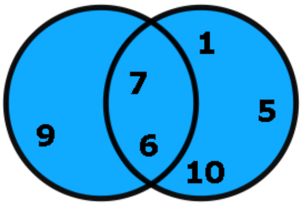


Figure 1.4: Here is the graphical representation of the sets $A = \{6, 7, 9\}$ and $B = \{1, 5, 6, 7, 10\}$. We see immediately that $A \cup B = \{1, 5, 6, 7, 9, 10\}$ and $A\ capB = \{6, 7\}$.

But classical set theory is not the subject of this course, so we stop here. However, as fuzzy logic is based on the concept of fuzzy, we see now the kind of problems that we face and that we will solve in the next sections: how to define such a union if memberships are not either 0 or 1?

Chapter 2

Fuzzy logic

As complexity rises, precise statements lose meaning and meaningful statements lose precision. Albert Einstein.

2.1 Fuzzy sets

Fuzzy logic is based on the theory of fuzzy sets, which is a generalization of the classical set theory. Saying that the theory of fuzzy sets is a generalization of the classical set theory means that the latter is a special case of fuzzy sets theory. To make a metaphor in set theory speaking, the classical set theory is a subset of the theory of fuzzy sets, as figure 2.1 illustrates.

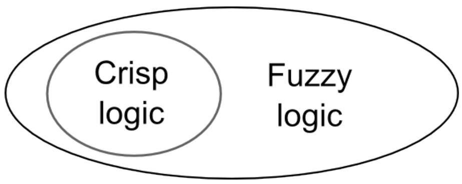


Figure 2.1: "The classical set theory is a subset of the theory of fuzzy sets"

Fuzzy logic is based on fuzzy set theory, which is a generalization of the classical set theory [Zadeh, 1965]. By abuse of language, following the habits of the literature, we will use the terms fuzzy sets instead of fuzzy subsets. The classical sets are also called clear sets, as opposed to vague, and by the same token classical logic is also known as Boolean logic or binary.

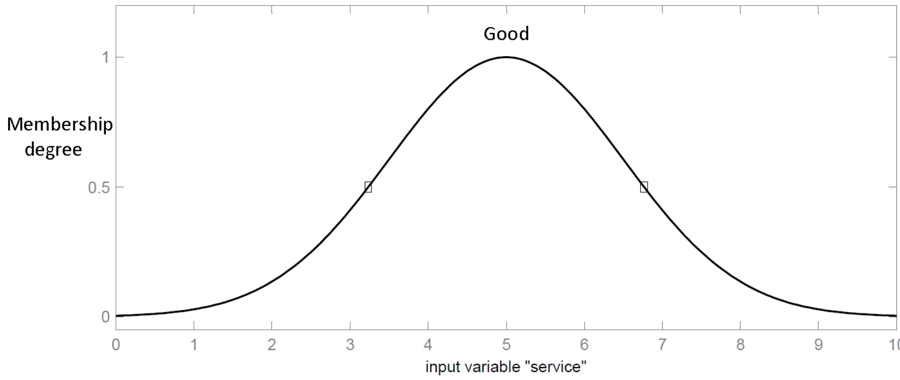


Figure 2.2: Membership function characterizing the subset of 'good' quality of service

The figure 2.2 shows the membership function chosen to characterize the subset of 'good' quality of service.

Definition 1.
Let X be a set. A fuzzy subset A of X is characterized by a ***membership function****. $f^a : X \to [0, 1]$. (In theory, it is possible that the output is greater than 1, but in practice it is almost never used.)*
Note: This membership function is equivalent to the identity function of a classical set.

In our tip example, we will redefine membership functions for each fuzzy set of each of our three variables:

- Input 1: quality of service. Subsets: poor, good and excellent.
- Input 2: quality of food. Subsets: awful and delicious.
- Output: tip amount. Subsets: low, medium and high.

The shape of the membership function is chosen arbitrarily by following the advice of the expert or by statistical studies: sigmoid, hyperbolic, tangent, exponential, Gaussian or any other form can be used.

The figure 2.3 shows the difference between a conventional set and a fuzzy set corresponding to a delicious food.

The figure 2.4 compare the two membership functions corresponding to the previous set.

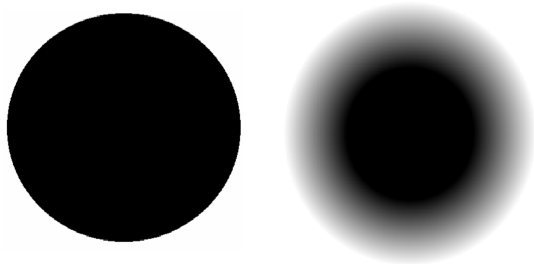


Figure 2.3: Graphical representation of a conventional set and a fuzzy set

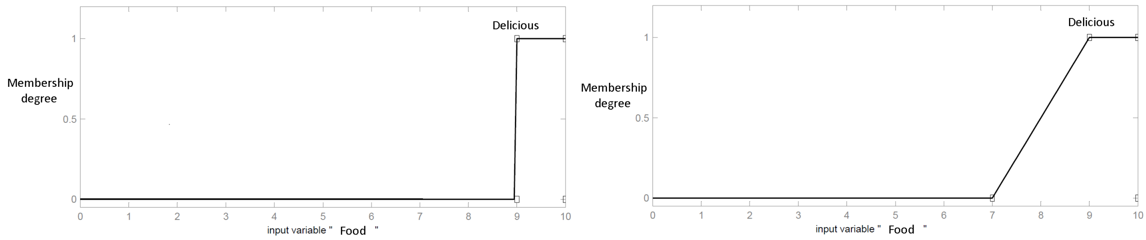


Figure 2.4: Comparison between a identity function of a conventional set and a membership function of fuzzy set

In order to define the characteristics of fuzzy sets, we are redefining and expanding the usual characteristics of classical sets.

Fuzzy set have a number of properties. Here are definitions of the most important properties, but they are notnecessary for understanding of the course. If you want, you can go now directly to the next section.

Let X be a set and A a fuzzy subset of X and μ_A the membership function characterizing it. $\mu_A(x)$ is called the membership degree of x in A.

Definition 2.
*The **height** of A, denoted h(A), corresponds to the upper bound of the codomain of its membership function:* $h(A) = \sup\{\mu_A(x) \mid x \in X\}$.

Definition 3.
*A is said to be **normalised** if and only if* $h(A) = 1$. *In practice, it is extremely rare to work on non-normalised fuzzy sets.*

Definition 4.
*The **support** of A is the set of elements of X belonging to at least some A (i.e.*

the membership degree of x is strictly positive). In other words, the support is the set $\text{supp}(A) = \{x \in X \mid \mu_A(x) > 0\}$.

Definition 5.
The ***kernel*** *of A is the set of elements of X belonging entirely to A. In other words, the kernel* $noy(A) = \{x \in X \mid \mu_A(x) = 1\}$. *By construction,* $\text{noy}(A) \subseteq \text{supp}(A)$.

Definition 6.
An α***-cut*** *of A is the classical subset of elements with a membership degree greater than or equal to*α *:* $\alpha\text{-cut}(A) = \{x \in X \mid \mu_A(x) \geqslant \alpha\}$.

Another membership function for an average tip through which we have included the above properties is presented in Figure 2.5.

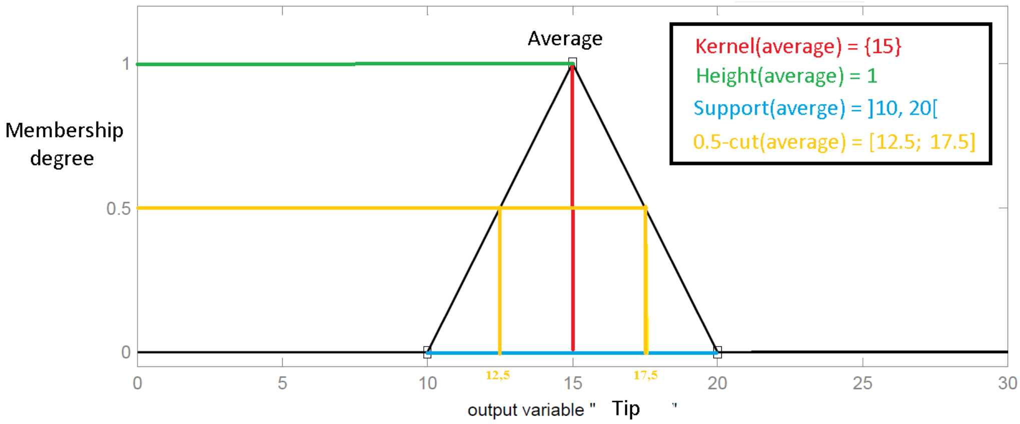


Figure 2.5: A membership function with properties displayed

We can see that if A was a conventional set, we would simply have $\text{supp}(A) = noy(A)$ and $h(A) = 1$ (ou $h(A) = 0$ si $A = \emptyset$). Our definitions can therefore recover the usual properties of classical sets. We will not talk about the cardinality property because we will not use this concept later in this course.

2.2 The linguistic variables

The concept of membership function discussed above allows us to define fuzzy systems in natural language, as the membership function couple fuzzy logic with linguistic variables that we will define now.

Definition 7.
Let V be a variable (quality of service, tip amount, etc.), X the range of values of

the variable and T_V a finite or infinite set of fuzzy sets. A ***linguistic variable*** *corresponds to the triplet (V, X, T_V).*

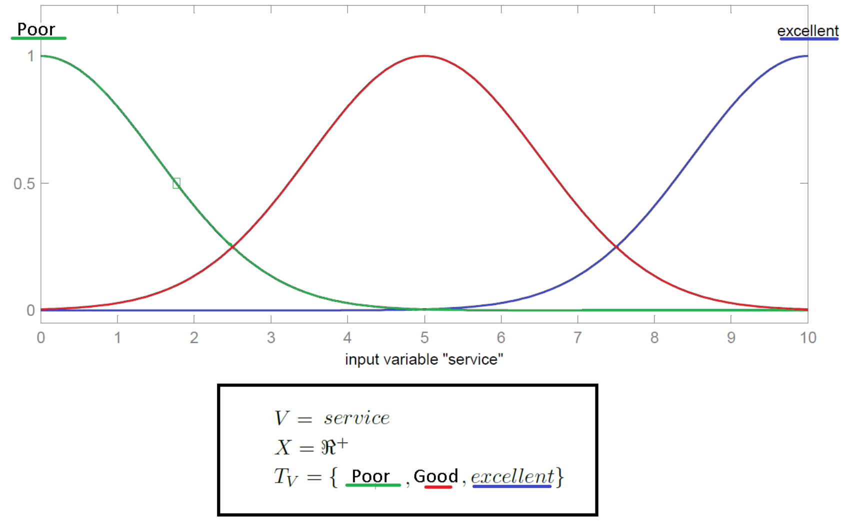


Figure 2.6: Linguistic variable 'quality of service'

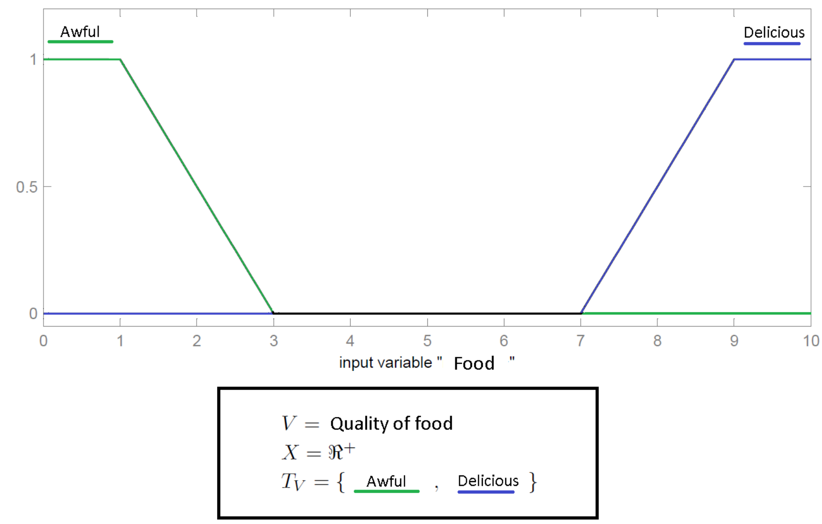


Figure 2.7: Linguistic variable 'quality of food'

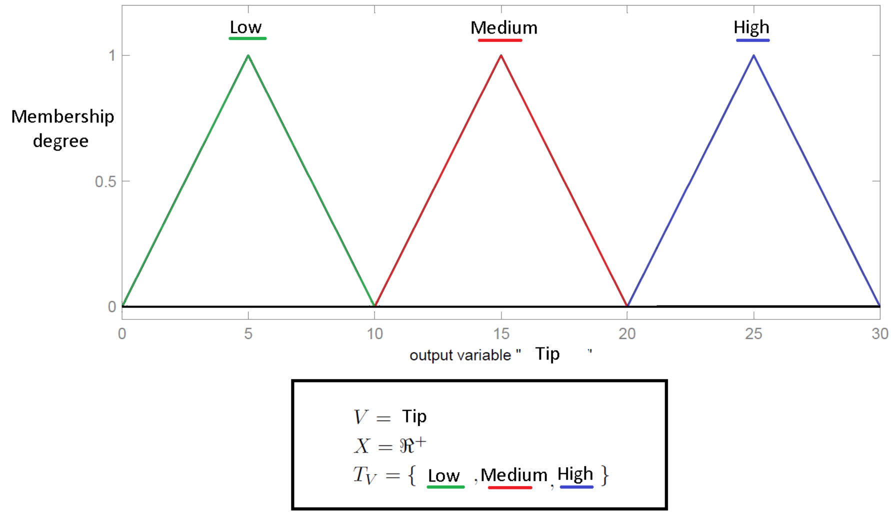


Figure 2.8: Linguistic variable 'tip amount'

When we define the fuzzy sets of linguistic variables, the goal is not to exhaustively define the linguistic variables. Instead, we only define a few fuzzy subsets that will be useful later in definition of the rules that we apply it. This is for example the reason why we have not defined subset "average" for the quality of the food. Indeed, this subset will not be useful in our rules. Similarly, it is also the reason why (for example) 30 is a higher tip than 25, while 25 however belongs more to the fuzzy set "high" as 30: this is due to the fact that 30 is seen not as high but very high (or exorbitant if you want to change adjective). However, we have not created of fuzzy set "very high" because we do not need it in our rules.

2.3 The fuzzy operators

In order to easily manipulate fuzzy sets, we are redefining the operators of the classical set theory to fit the specific membership functions of fuzzy logic for values strictly between 0 and 1.

Unlike the definitions of the properties of fuzzy sets that are always the same, the definition of operators on fuzzy sets is chosen, like membership functions. Here are the two sets of operators for the complement (NOT), the intersection (AND) and union (OR) most commonly used:

Name	Intersection AND: $\mu_{A \cap B}(x)$	Union OU: $\mu_{A \cup B}(x)$	Complement NOT: $\mu_{\bar{A}(x)}$
Zadeh Operators MIN/MAX	$min\,(\mu_A(x), \mu_B(x))$	$max\,(\mu_A(x), \mu_B(x))$	$1 - \mu_A(x)$
Probabilistic PROD/PROBOR	$\mu_A(x) \times \mu_B(x)$	$\mu_A(x) + \mu_B(x) - \mu_A(x) \times \mu_B(x)$	$1 - \mu_A(x)$

With the usual definitions of fuzzy operators, we always find the properties of commutativity, distributivity and associativity classics. However, there are two notable exceptions:

- In fuzzy logic, the law of excluded middle is contradicted: $A \cup \bar{A} \neq X$, i.e. $\mu_{A \cup \bar{A}}(x) \neq 1$.
- In fuzzy logic, an element can belong to A and not A at the same time: $A \cap \bar{A} \neq \emptyset$, i.e. $\mu_{A \cap \bar{A}}(x) \neq 0$. Note that these elements correspond to the set $\mathrm{supp}(A) - \mathrm{noy}(A)$.

2.4 Reasoning in fuzzy logic

In classical logic, the arguments are of the form:

$$\begin{cases} If\ p\ then\ q \\ p\ true\ then\ q\ true \end{cases}$$

In fuzzy logic, fuzzy reasoning, also known as approximate reasoning, is based on **fuzzy rules** that are expressed in natural language using linguistic variables which we have given the definition above. A fuzzy rule has the form:
If $x \in A$ *and* $y \in B$ *then* $z \in C$, with A, B and C fuzzy sets.

For example:
'If (the quality of the food is delicious), then (tip is high)'.

The variable 'tip' belongs to the fuzzy set 'high' to a degree that depends on the degree of validity of the premise, i.e. the membership degree of the variable 'food quality' to the fuzzy set 'delicious '. The underlying idea is that the more propositions in premise are checked, the more the suggested output actions must be applied. To determine the degree of truth of the proposition fuzzy 'tip will be high', we must define the fuzzy implication.

Like other fuzzy operators, there is no single definition of the fuzzy implication: the fuzzy system designer must choose among the wide choice of fuzzy implications already defined, or set it by hand. Here are two definitions of fuzzy implication most commonly used:

Name	Truth value
Mamdani	$min\,(f_a(x), f_b(x))$
Larsen	$f_a(x) \times f_b(x)$

Notably, these two implications do not generalize the classical implication. There are other definitions of fuzzy implication generalizing the classical implication, but are less commonly used.

If we choose the Mamdani implication, here is what we get for the fuzzy rule *'If (the food quality is delicious), then (tip is high)'* where the food quality is rated 8.31 out of 10:

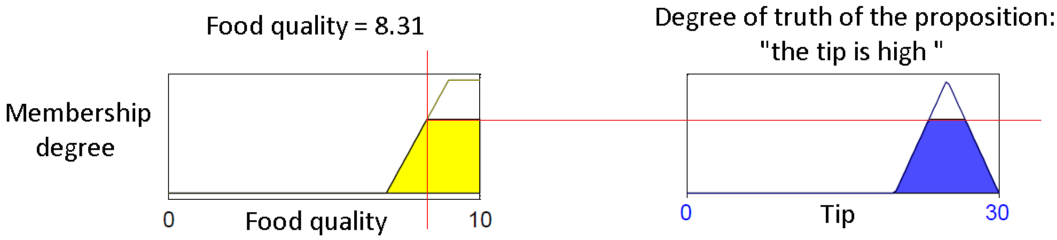


Figure 2.9: Example of fuzzy implication

The result of the application of a fuzzy rule thus depends on three factors:

1. the definition of fuzzy implication chosen,
2. the definition of the membership function of the fuzzy set of the proposition located at the conclusion of the fuzzy rule,
3. the degree of validity of propositions located premise.

As we have defined the fuzzy operators AND, OR and NOT, the premise of a fuzzy rule may well be formed from a combination of fuzzy propositions. All the rules of a fuzzy system is called the **decision matrix**. Here is the decision matrix for our tip example:

If the service is bad or the food is awful	then the tip is low
If the service is good	then the tip is average
If the service is excellent or the food is delicious	then the tip is high

The figure 2.10 shows what we get for fuzzy rule *'If (the service is excellent and the food is delicious), then (tip is high)'* where the quality of service is rated 7.83 out of 10 and the quality of food 7.32 out of 10 if we choose the Mamdani implication and the translation of OR by MAX.

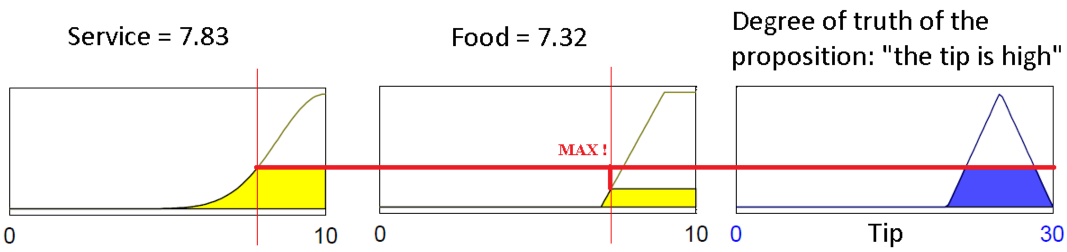


Figure 2.10: Example of fuzzy implication with conjunction OR translated into a MAX

We will now apply all the 3 rules of our decision matrix. However, we will obtain three fuzzy sets for the tip: we will aggregate them by the operator MAX which is almost always used for aggregation. The figure ?? shows this aggregation.

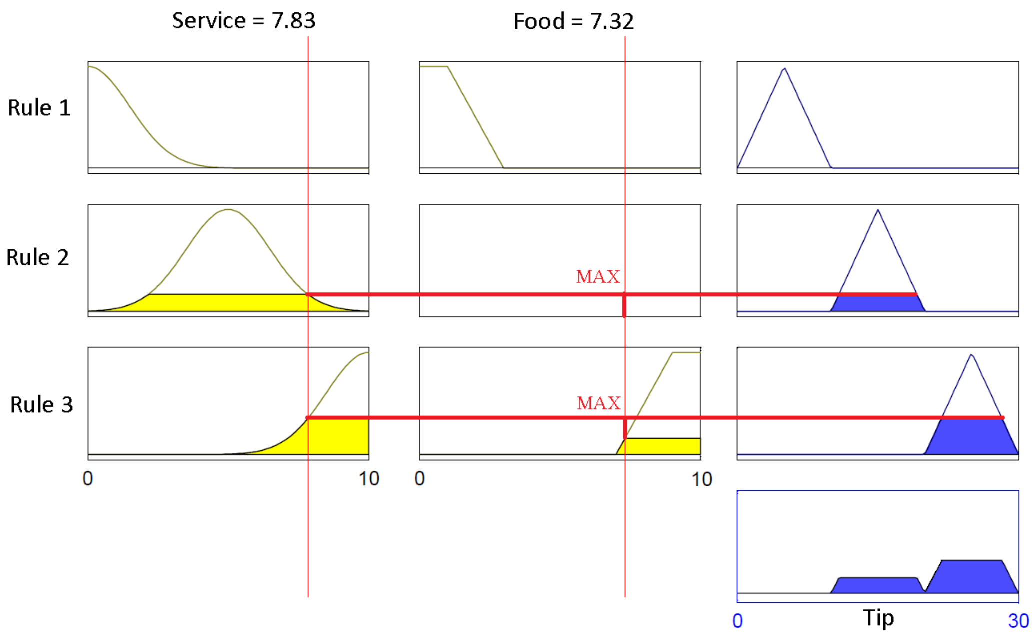


Figure 2.11: Example of fuzzy implication using the decision matrix

As we see, we now has to make the final decision, namely decide how much the tip will be knowing that the quality of service is rated 7.83 out of 10 and quality of food 7.32 out of 10. This final step, which allows to switch from the fuzzy set resulting from the aggregation of results to a single decision, is called the **defuzzification**.

2.5 The defuzzification

As with all fuzzy operators, the fuzzy system designer must choose among several possible definitions of defuzzification. A detailed list can be found in the research article [Leekwijck and Kerre, 1999]. We will briefly present the two main methods of defuzzification: the method of the mean of maxima (MeOM) and the method of center of gravity (COG).

The MeOM defuzzification sets the output (decision of the tip amount) as the average of the abscissas of the maxima of the fuzzy set resulting from the aggregation of the implication results.

$Décision = \frac{\int_S y \cdot dy}{\int_S dy}$
where $S = \{y_m \in R, \mu(y_m) = SUP_{y \in R}(\mu(y))\}$
and R is the fuzzy set resulting from the aggregation of the implication results.

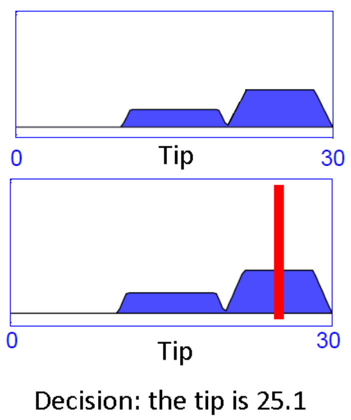


Figure 2.12: Defuzzification with the method of the mean of maxima (MeOM)

The COG defuzzification is more commonly used. It defines the output as corresponding to the abscissa of the center of gravity of the surface of the membership function characterizing the fuzzy set resulting from the aggregation of the implication results.

$$Décision = \frac{\int_S y \cdot \mu(u) \cdot dy}{\int_S \mu(u) \cdot dy}$$

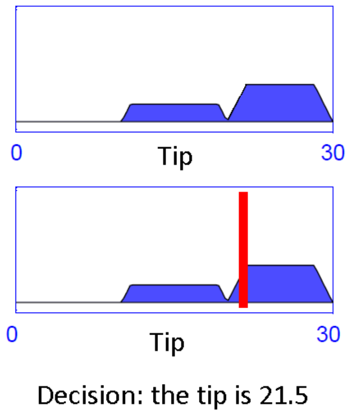


Figure 2.13: Defuzzification with the method of center of gravity (COG)

This definition avoids the discontinuities could appear in the MeOM defuzzification, but is more complex and has a greater computational cost. Some work as [Madau D., 1996] seek to improve performance by searching other methods as effective but with a lower computational complexity. As we see in the two figures showing the MeOM and COG defuzzifications applied to our example, the choice of this method can have a significant effect on the final decision.

2.6 Conclusions

In the definitions, we have seen that the designer of a fuzzy system must make a number of important choices. These choices are based mainly on the advice of the expert or statistical analysis of past data, in particular to define the membership functions and the decision matrix.

Here is an overview diagram of a fuzzy system:

In our example,

- the **input** is 'the quality of service is rated 7.83 out of 10 and quality of food 7.32 10',

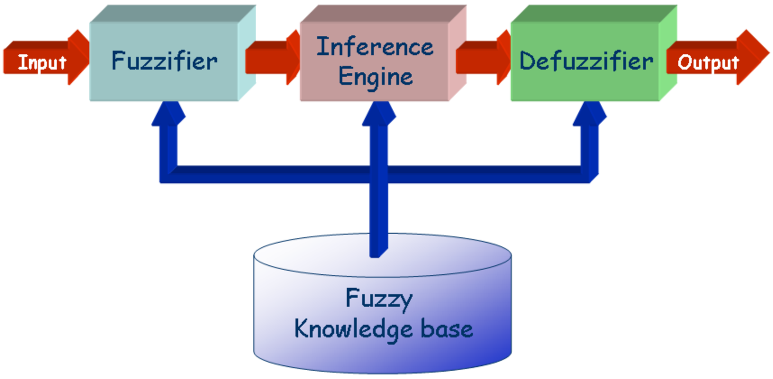


Figure 2.14: Overview diagram of a fuzzy system:

- the **fuzzifier** corresponds to the 3 linguistic variables 'service quality', 'food quality' and 'tip amount',
- the **inference engine** is made of the choice of fuzzy operators,
- the **fuzzy knowledge base** is the set of fuzzy rules,
- the **defuzzifier** is the part where has to be chosen the method of defuzzification,
- the **output** is the final decision: 'the tip amount is 25.1'.

It is interesting to see all the decisions based on each variable with our fuzzy inference system compared to the decisions that we would get using classical logic:

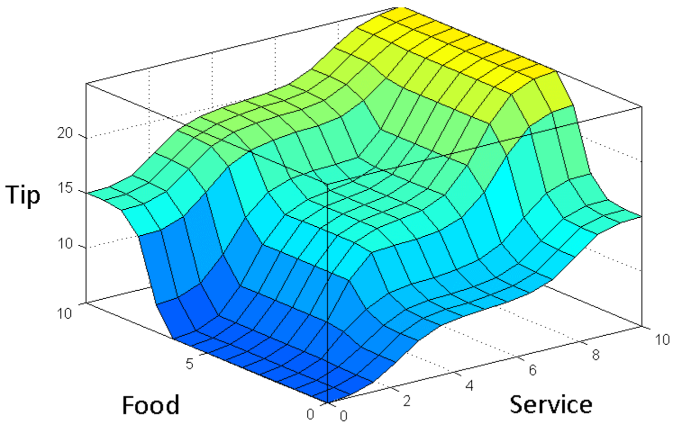


Figure 2.15: Decisions of a system based on fuzzy system

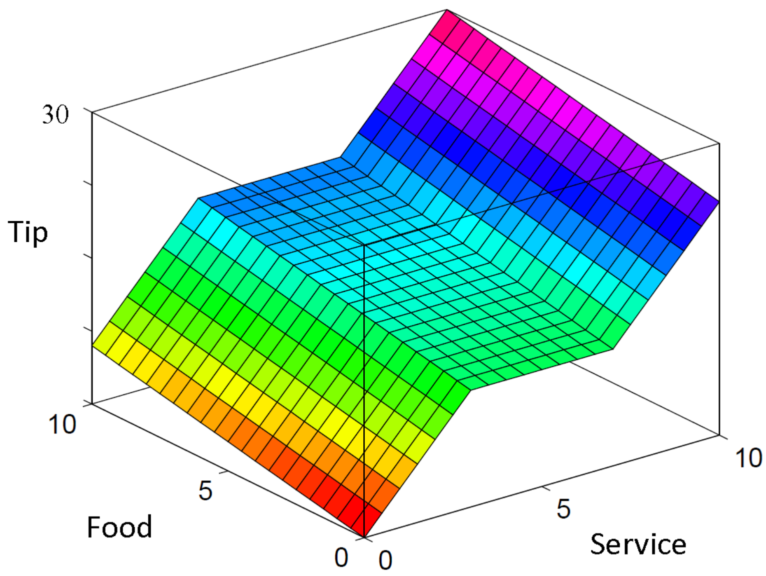


Figure 2.16: Decisions of a system based on classical logic

Thus, fuzzy logic allows to build inference systems in which decisions are without discontinuities, flexible and nonlinear, i.e. closer to human behavior than classical logic is. In addition, the rules of the decision matrix are expressed in natural language. This has many advantages, such as include knowledge of a non-expert computer system at the heart of decision-making model or finer aspects of natural language.

We will see in the third chapter how we can train a fuzzy inference system for the specific problem.

Chapter 3

Training fuzzy inference systems

3.1 Neuro-fuzzy systems

Neuro-fuzzy systems were introduced in the thesis of Jyh-Shing Roger Jang in 1992 under the name “Adaptative-Networks-based Fuzzy Inference Systems” (ANFIS) [Jangi, 1992]. They use the formalism of neural networks by expressing the structure of a fuzzy system in the form of a multilayer perceptron.

A multilayer perceptron (MLP) is a neural network without cycle. The input layer is given a vector network and the network returns a result vector in the output layer. Between these two layers, the elements of the input vector are weighted by the weights of the connections and mixed in the hidden neurons located in the hidden layer. Figure 3.1 illustrates an example of a feedforward neural network.

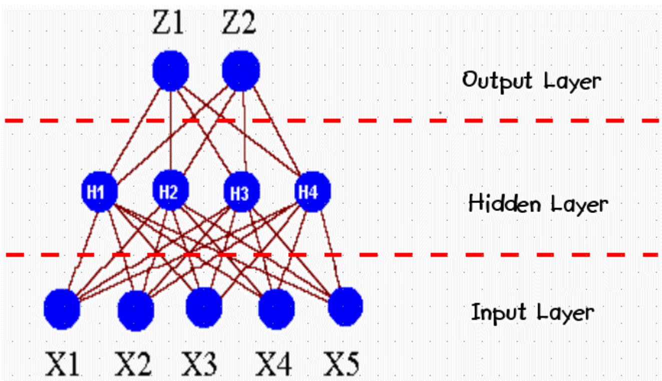


Figure 3.1: Example of a feedforward neural network

Several activation functions for the output layer are commonly used, such as linear, logistic or softmax. Similarly, there are several error backpropagation algorithms that optimize the learning of weights from the mistakes made between the values computed by the network and the actual values: Conjugate gradient optimization, Scaled Conjugate Gradient, Quasi-Newton optimization, and so on.

Figure 3.2 presents an example of the organization of a multilayer perceptron representing the neuro-fuzzy system:

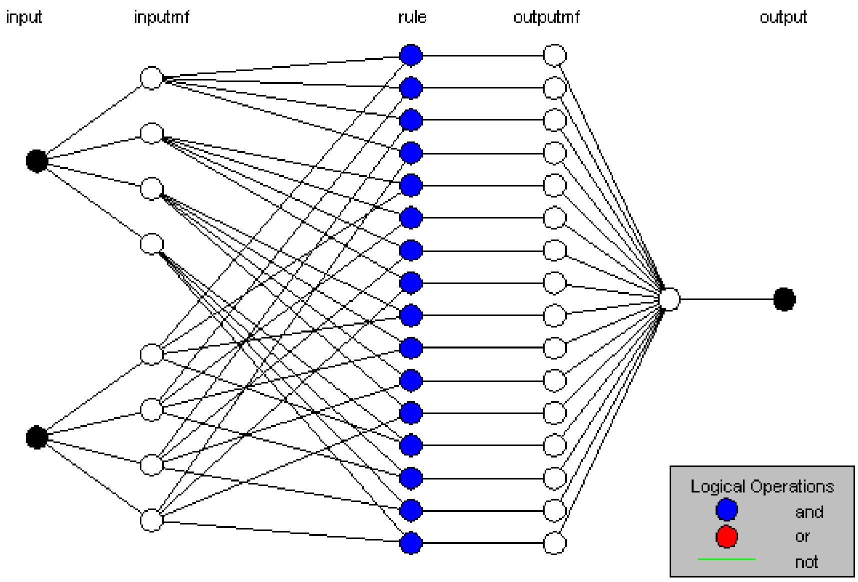


Figure 3.2: Structure of a neuro-fuzzy system

TODO: expand

3.2 Evolutionary computation

TODO

Chapter 4

Acknowledgments

I want to thank (in chronological order) :

2011 Jean Baratgin and Emmanuel Sander for letting me the opportunity to write this course on fuzzy logic,

2011 Alp Mestan and Romuald Perrot for their help in making this course online on Developpez.com,

2011 3DArchi, Julien Plu and Antoine Dinimant for their technical advice,

2011 Patriarch24 for proof reading the French version,

2012 The SiteDuZero community for inciting me to publish a second version of the course with much-needed improvements,

2013 Una-May O'Reilly for motivating me to translate this course into English.

Bibliography

[Jangi, 1992] Jangi, R. (1992). *Neuro-Fuzzy modeling: Architecture, Analysis and Application*. PhD thesis, University of California, Berkeley. [cited at p. 18]

[Leekwijck and Kerre, 1999] Leekwijck, W. V. and Kerre, E. E. (1999). Defuzzification: criteria and classification. *Fuzzy Sets and Systems*, 108(2):159 – 178. [cited at p. 14]

[Madau D., 1996] Madau D., D. F. (1996). Influence value defuzzification method. *Fuzzy Systems, Proceedings of the Fifth IEEE International Conference*, 3:1819 – 1824. [cited at p. 15]

[Zadeh, 1965] Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338 – 353. [cited at p. 5]