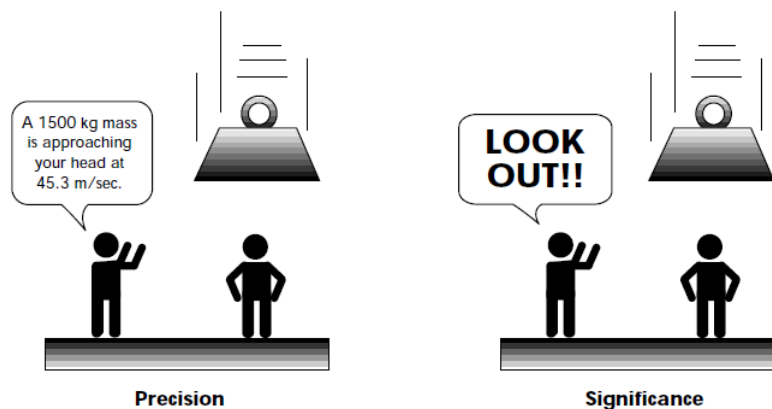Conservatoire National des Arts et Métiers

Fuzzy logic:
introducing human reasoning
within decision support systems?

Franck Dernoncourt
franck.dernoncourt@gmail.com

Professor: Elisabeth Métais

Paris, February 2011

Abstract

Fuzzy logic is based on solid mathematical foundations, including the mathematical theory of fuzzy sets, generalizing classical set theory. Firstly, we define fuzzy operators, which generalize operators of classical logic.

As a second step, we see how fuzzy logic can imitate human reasoning. We analyze the contribution of fuzzy logic for the modeling of human reasoning, and also experimentally investigate whether the decisions taken by humans correspond to decisions taken by fuzzy systems. To this end, given that the literature is deficient on this point, we design an experiment for that purpose and analyze the results.

We study the potential applications for databases and decision support systems in Chapter 5. How to integrate the advantages of fuzzy logic in the database? To which extent decision-making systems can use the flexibility of fuzzy logic?

We then analyze the potential applications for decision support systems and databases. We show that at the heart of the company, bringing together all the interesting information from the operational databases, decision systems could benefit greatly from fuzzy logic by giving the keys to human reasoning, allowing to refine the decision-making.

Database theorists know what fuzzy logic could bring them in terms of information modeling: queries more intuitive and more powerful on the one hand, the data more consistent with the reality on the other. Many papers have been written, but few significant achievements have followed. The lack of consensus on a standard is probably the main reason behind.

Contents

List of Figures

Chapter 1

Introduction

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

1.1 The issue

Knowledge available to humans are almost never perfect. These imperfections can be distinguished into two classes:

- **Inaccuracies**, which refer to knowledge whose validity is subject to question. For example, if we know someone bumped his head on a ceiling, we can guess that he is likely to be very tall.

- **Incertitudes**, which refer to knowledge that is not clearly perceived or defined. For example, instead of saying someone is 2 feet and 3 inches, we usually say that person is very tall.

In fact, these imperfections stem from the nature of man and the world: none of our senses and observation instruments allow us to reach an infinite precision and the world is based on the principle of continuity in the mathematical sense, as opposed to discrete values. This is one reason why it is so difficult to establish a system of measurement units and why any quantity is only approximate. Similarly, rare are the situations where we can be totally sure that this statement is true without imposing a prior set of assumptions.

As a result, the knowledge on which human reasoning relies on is almost always marred by a number of uncertainties and inaccuracies. We will not discuss here about scientific reasoning, the objective of which is just to get rid of any imperfections, but about all the usual reasoning that we do every day, constantly, on things, people and

thoughts surrounding us. This kind of reasoning ranges from driving a car to the treatment recommended by a doctor to his patient or to the decision to buy a loaf of bread for tomorrow morning.

Surprisingly, and fortunately, despite the vagueness that characterizes our world view, the quality of human reasoning is remarkable. Whether in simple or complex situations, decisions are generally very good compared to the vagueness and uncertainty of the problem's data.

A human being, as part of systems theory, is considered a system in itself [Mélès, 1971], which is able to reason on very imperfect data. His impressive performance, given the complexity and diversity of situations encountered, interested some researchers in systems theory in the 1960s, among whom Lotfi Zadeh, the founder of fuzzy logic.

1.2   Fuzzy logic

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 or-ange...	if my speed is aver-age...	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.

1.3 Decision-making systems

Fuzzy logic seems to reproduce the flexibility of human reasoning as to its consideration of the imperfections of available data. Therefore it would be interesting to use in the heart of **expert systems**, systems which aim to reproduce the cognitive mechanisms of an expert in a particular field. Fuzzy logic can also be used for **decision support system**, during the phase of data analysis for example.

A decision support system is a set of resources, tools and methods to collect, consolidate and return the data, whether tangible or intangible, of a firm in order to provide a decision support and enable those in charge of corporate strategy to have an overview of the activity.

The primary objective of such systems is as the name implies to assist in decision making. To achieve this goal, the company data are often grouped into a data warehouse, which, using the definition of Bill Inmon, one of the founders of decision systems [Inmon, 1992], is a collection of data that are thematic integrated, nonvolatile, historized and organized for decision making.

Fuzzy logic can be useful for decision making, either to discover rules or fuzzy inferences to better understand the data and thus inform decisions, or to execute fuzzy queries on expert knowledge.

1.4 Report's structure

In order to study the issues previously raised, we will first in Chapter 2 formally define what fuzzy logic is. As we shall see, fuzzy logic is based on solid mathematical foundations, including the mathematical theory of fuzzy sets, generalizing the classical set theory. It will therefore define operators for fuzzy generalized operators of classical logic. As the purpose of this report is not to enter into too theoretical considerations, we define the basics of fuzzy logic throughout an example of a complete fuzzy inference system.

Secondly, we will see how fuzzy logic can mimic human reasoning in chapters 3 and 4. This first introductive chapter was very general: it is necessary to further analyze the contribution of fuzzy logic with respect to modeling human reasoning as well as experimentally investigate whether the decisions taken by the correspond to human decisions of fuzzy systems. To this end, the literature is incomplete on this point an experiment was designed to this purpose and the results will be discussed.

Finally, we will see potential applications for databases and decision support systems in Chapter 5. How can we integrate the advantages of fuzzy logic in databases? To which extent decision-making systems can use the flexibility of fuzzy logic?

Chapter 2

Fuzzy logic

Vagueness is no more to be done away with in the world of logic than friction in mechanics. Charles Sanders Peirce.

To exemplify each definition, we will develop throughout this section a fuzzy inference system the real purpose of which is to determine the tip-up in the end of a meal in a restaurant, based on service quality and the quality of food.

2.1 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.
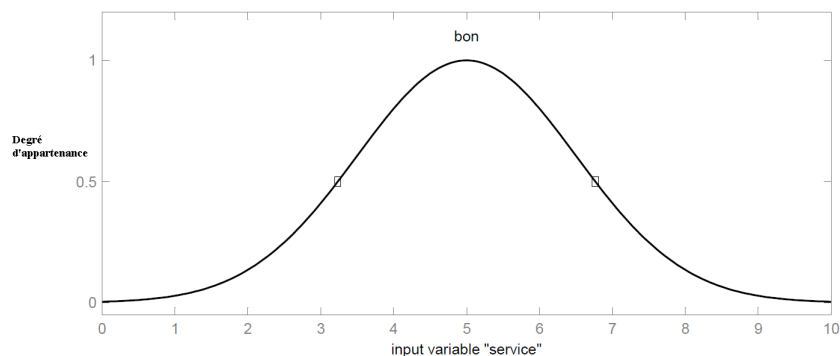
Figure 2.1: Membership function characterizing the subset of 'good' quality of service

The figure 2.1 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**. *$f^a : X \to [0,1]$.*
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 2.2 shows the difference between a conventional set and a fuzzy set corresponding to a delicious food.
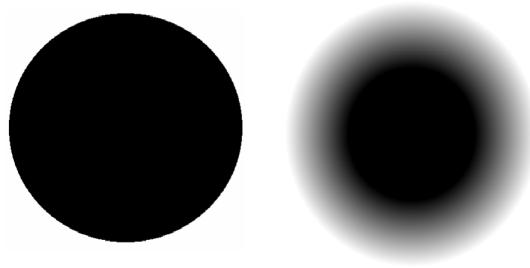
Figure 2.2: Graphical representation of a conventional set and a fuzzy set

The figure 2.3 compare the two membership functions corresponding to the previous set.

In order to define the characteristics of fuzzy sets, we are redefining and expanding the usual characteristics of classical sets.
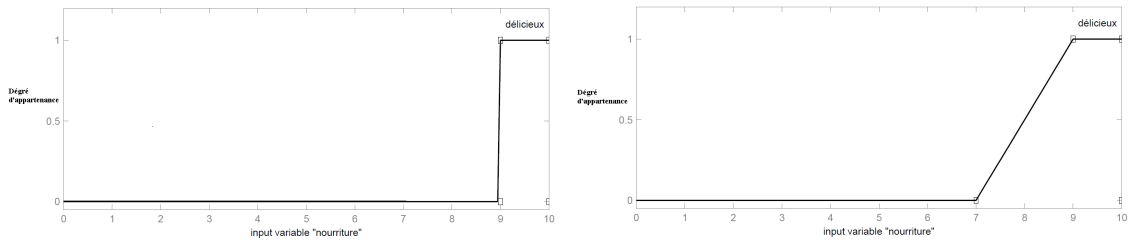
Figure 2.3: Comparison between a identity function of a conventional set and a membership function of fuzzy set

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 toα : $\alpha\text{-cut}(A) = \{x \in X \mid \mu_A(x) \geqslant \alpha\}$.*

Another membership function for an average tip through which we have included the above properties is presented in Figure 2.4.

We note 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 report.
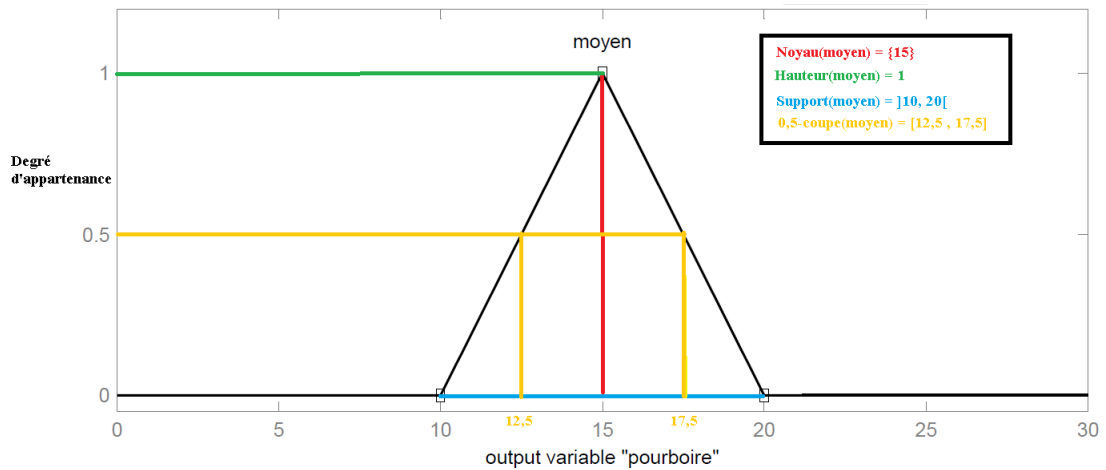
Figure 2.4: A membership function with properties displayed

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).*
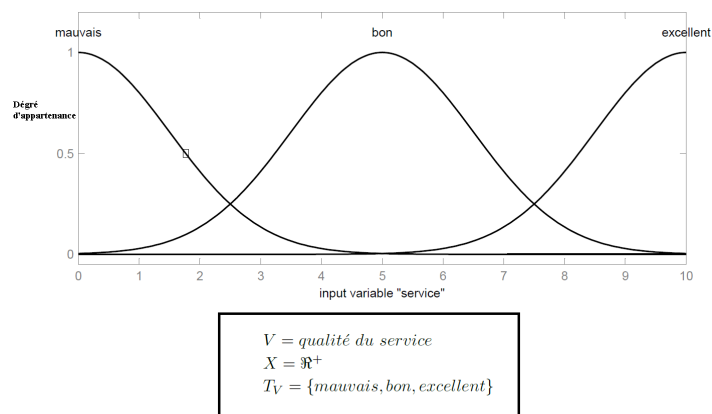
Figure 2.5: Linguistic variable 'quality of service'
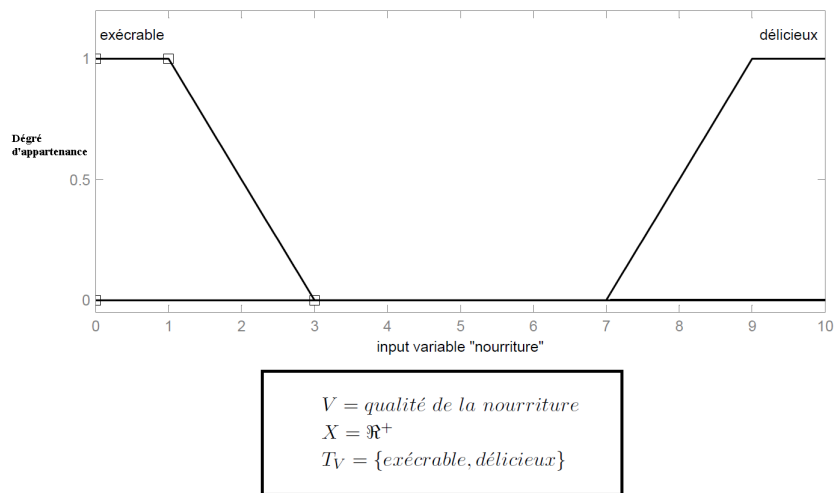
Figure 2.6: Linguistic variable 'quality of food'
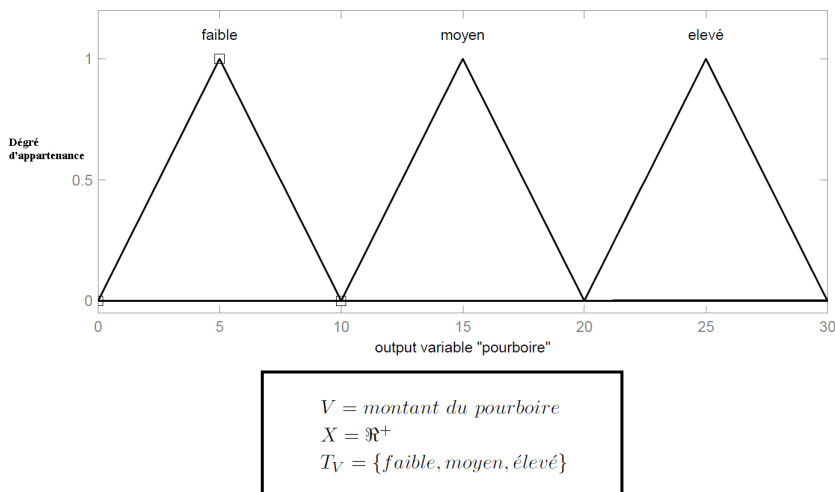
Figure 2.7: Linguistic variable 'tip amount'

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:

$$\begin{cases} Si\ p\ alors\ q \\ p\ vrai\ alors\ q\ vrai \end{cases}$$

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	$min\left(f_a(x), f_b(x)\right)$
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:
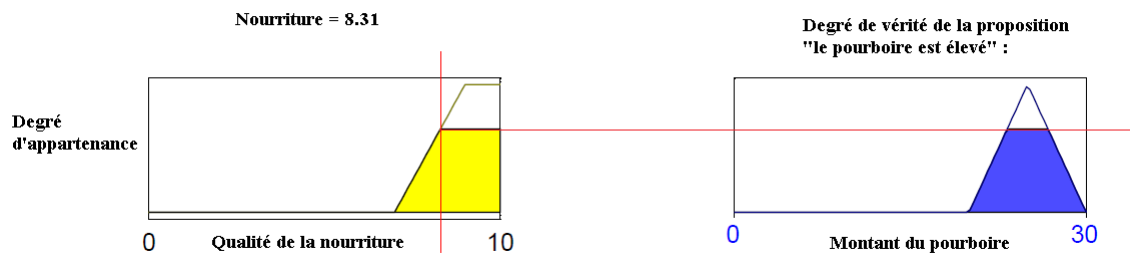
Figure 2.8: Example of fuzzy implication

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

If we choose the Mamdani implication and the translation of OR by MAX, here is what we get for the 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:
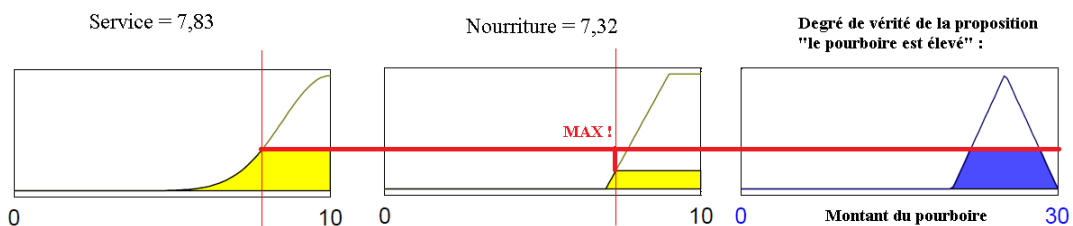
Figure 2.9: Example of fuzzy implication with OR translated by MAX

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.

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
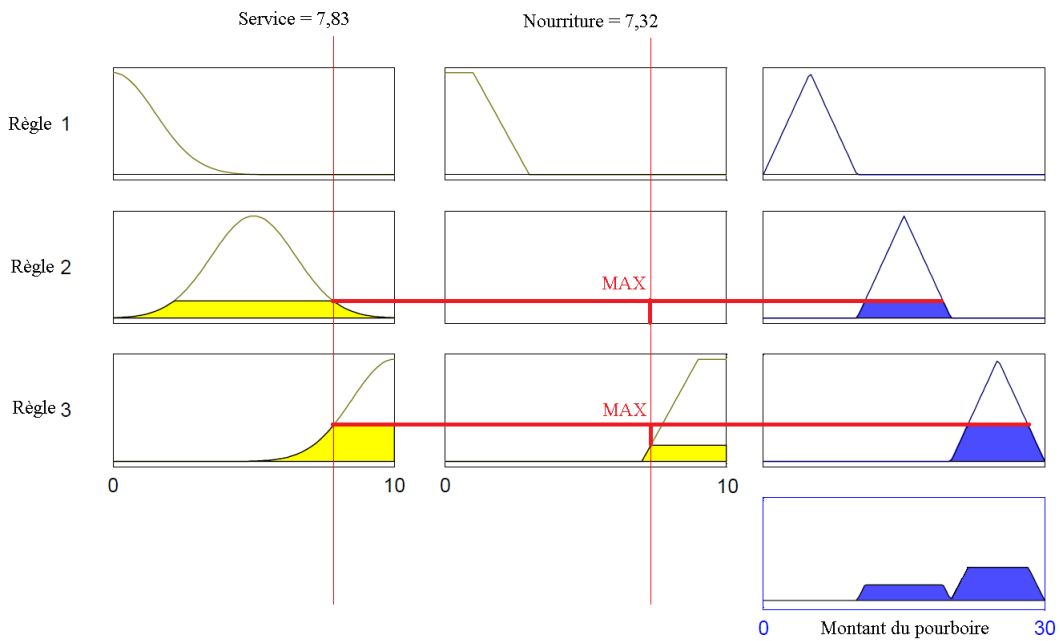
Figure 2.10: Example of fuzzy implication using the decision matrix

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.
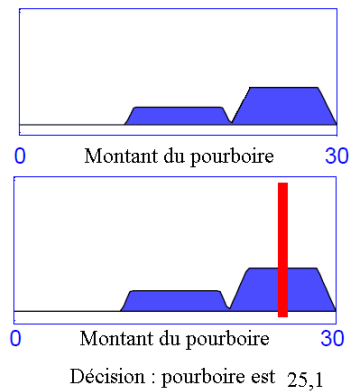
Figure 2.11: Defuzzification with the method of the mean of maxima (MeOM)

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.

$$Décision = \frac{\int_S y \cdot \mu(u) \cdot dy}{\int_S \mu(u) \cdot dy}$$
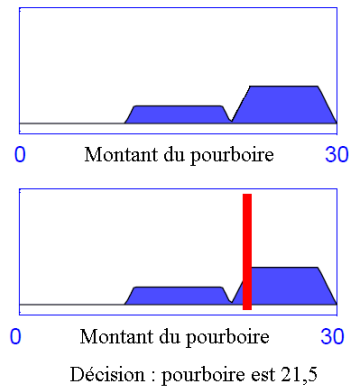
Figure 2.12: Defuzzification with the method of center of gravity (COG)

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,
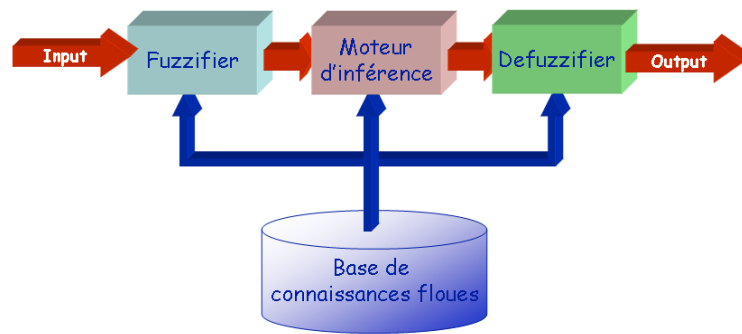
Figure 2.13: Overview diagram of a fuzzy system:

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

- 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:

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. We will see in the second chapter if human reasoning in decision-making processes as in the example of the tip has similar results to inferences from fuzzy systems.
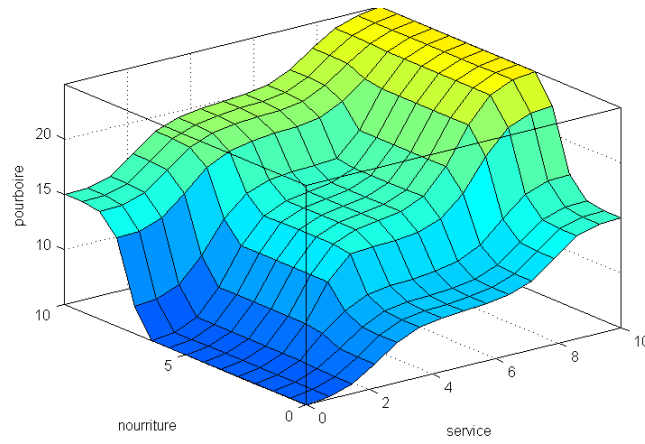
Figure 2.14: Decisions of a system based on fuzzy system
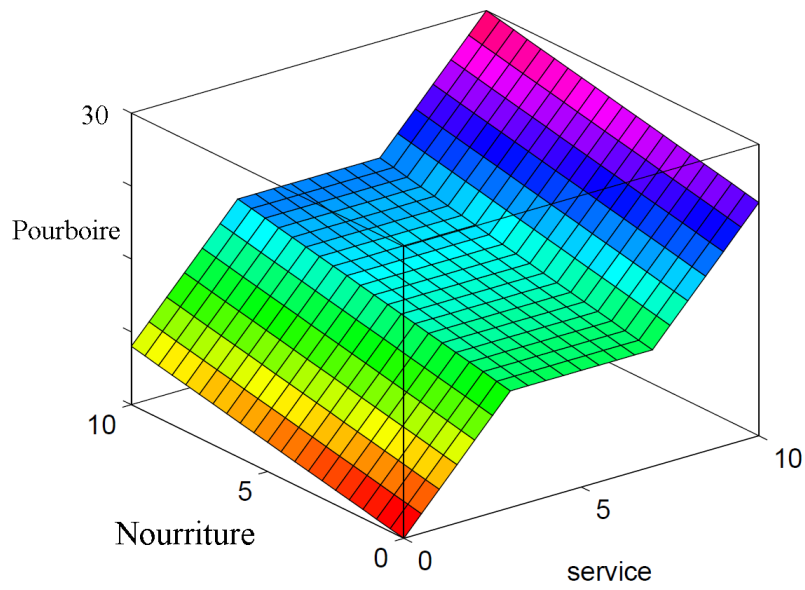
Figure 2.15: Decisions of a system based on classical logic

Chapter 3

Human reasoning

L'être humain va au-delà de l'information immédiatement disponible.
Jerome Bruner.

3.1   The problem

Reason is the ability to think, as it allows human beings to judge well and to apply this judgement to action. It is the ability to develop, from primitive knowledge of the real or hypothetical state of one's environment (**prémisses**), other knowledge or beliefs about the state of the environment (**conclusions**), through activities totally internalized. **Reasoning** means the activity of reason, i.e. the method of passing from premises to conclusion.

The 20th century was a major turning point for the understanding of human reasoning: assumptions that were previously firmly rooted in the common thought were totally challenged, and new theories have emerged to fill the gaps.

However, many experimental observations on human reasoning remain only partially explained: the purpose of this chapter is to see how fuzzy logic can represent a novel approach to the understanding of human reasoning.

To this end, we will address several issues that motivated the development of new theories of human reasoning from the 20th century and whether the fuzzy logic can provide an alternative response.

3.2   Non-additive probabilities

The beginning of probability theory dates back from the correspondence between

Pierre de Fermat and Blaise Pascal in 1654 and was axiomatized by Kolmogorov in 1933 [Shafer G., 2005]. One direct consequence of these axioms is that $P(\Omega \setminus E) = 1 - P(E)$, where Ω is the universe, and E some event.

Probabilities can be interpreted in two ways:

- Objectively: probabilities are values given for events in the context of gambling.

- Subjectively: probabilities correspond to the knowledge of a human being on an event or state of the world.

As the purpose of this report is to study human reasoning, we keep only the second interpretation. In this perspective, the probability value is unique to each human being.

However, it turns out that we can experimentally show that the probabilities used and generated by human reasoning does not follow the Kolmogorov axioms, as they do no respect one of Kolmogorov axioms' direct consequence $P(\Omega \setminus E) = 1 - P(E)$ previously seen.

3.2.1 Subadditivity

[Redelmeier DA, 1995] made the following experiment: the case of a patient admitted to a hospital was summarized and presented to doctors from Stanford. The latters were then asked to estimate the probability of each of the following events:

- The patient died while in hospital,

- The patient will leave alive the hospital but died within one year,

- The patient will die somewhere between 1 and 10 years,

- The patient will live at least 10 more years.

As these four events form a partition of the universe, whatever the life duration of the patient, 1 and only 1 of these events will be true. We would therefore expect that the sum of the probabilities is equal to 1. However, experimentally, we find that the average of the sum of these four probabilities is equal to 1.64 (confidence interval 95% : [1.34, 1.94]). This result shows that the sum of the probabilities of events of a partition of the universe is greater than 1 is called **subadditivity** of probability judgement.

The term subadditivity come from mathematics, which define a subadditive function as follows:

Definition 8.
*A function $f\colon A \to B$ is said to be **subadditive** iff A domain closed under addition, B is a codomain partially ordered closed under addition and $\forall x, y \in A, f(x + y) \leq f(x) + f(y)$.*

3.2.2 Superadditivity

Conversely, some articles such as [Cohen et al., 1956], [Macchi, 1999] and [Boven and Epley, 2003] show that under certain conditions, the probability judgement may reflect a **sur-additivité**, autrement dit que la somme des probabilités des événements d'une partition de l'univers est inférieure à 1.

Definition 9.
*A fontion F is said to be **superadditive** iff it is not subadditive.*

For example, [Macchi, 1999] asked students to evaluate the probability of two events:

- The freezing point of gasoline does not equal that of alcohol. What is the probability that the freezing point of gasoline is greater than that of alcohol?

- The freezing point of gasoline does not equal that of alcohol. What is the probability that the freezing point of alcohol is greater than that of gasoline?

Analysis of the results shows that on average the sum of these probabilities is about 0.9, instead of 1 as we might expect, the first event being the complement of the second. [Macchi, 1999] examines the various factors that may increase or decrease this amount. Broadly, we find that the more the subject is sure of his decision, the more likely he will have subadditive probabilities, and conversely the more uncertain he is about his decision, the more likely he will have superadditive probabilities. The **support theory** [Tversky and Koehler, 1994] predicts in more details the types of non-additivity and provides a unified explanation of the experimental results previously described.

Therefore, human reasoning ignores the axioms of Kolmogorov, who laid the foundations of probability theory. It does not allow to entirely model human reasoning.

3.2.3 Fuzzy logic and non-additivity

In fuzzy logic, as we saw in the first chapter, the law of excluded middle is contradicted ($A \cup \bar{A} \neq X$, i.e. $\mu_{A \cup \bar{A}}(x) \neq 1$) because the definition of the OR is classically $max\left(\mu_A(x), \mu_B(x)\right)$ (Zadeh min/max) or $\mu_A(x) + \mu_B(x) - \mu_A(x) \times \mu_B(x)$ (PROD/PROBOR). The problem of non-additivity does not arise. According to the chosen membership functions, we can have what we want: superadditivity, or subadditivity or sum equal to 1.

Let's model our last example on freezing points taken from [Macchi, 1999] in a fuzzy system:
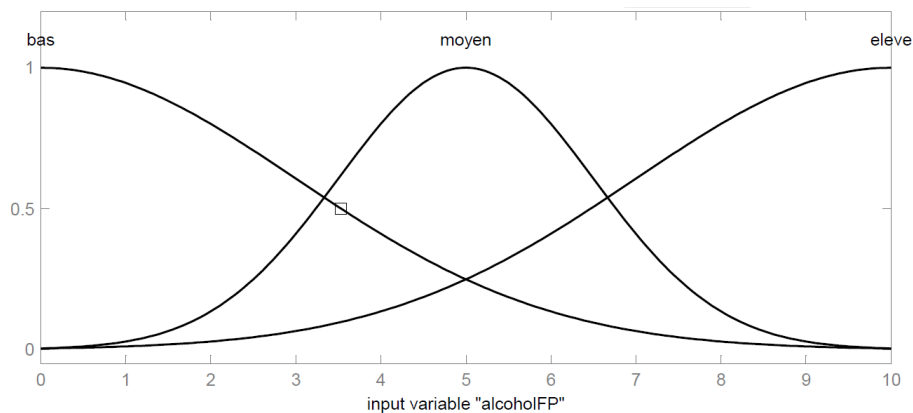
Figure 3.1: Variable freezing point (FP) of alcohol
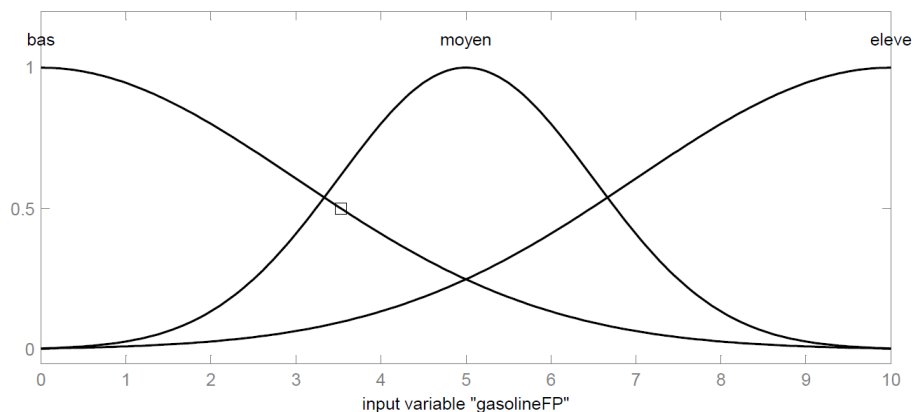
Figure 3.2: Variable freezing point (FP) of gasoline
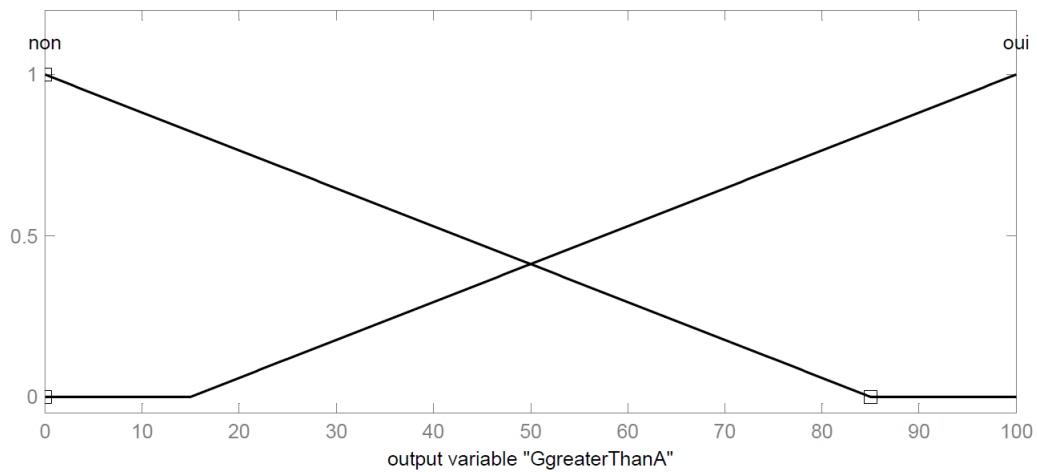
Figure 3.3: Variable freezing point (FP) of gasoline greater than alcohol (GgreaterThanA)

Here the choices of fuzzy operators we do (usual choice):

Operator	Choice
OR	max
AND	min
Implication	Mamdani
Aggregation	max
Defuzzication	COG

The decision matrix will be:

If FP alcohol is low	and if PC gasoline is average	then GgreaterThanA is yes.
If FP alcohol is low	and if PC gasoline is high	then GgreaterThanA is yes.
If FP alcohol is average	and if PC gasoline is low	then GgreaterThanA is no.
If FP alcohol is average	and if PC gasoline is high	then GgreaterThanA is yes.
If FP alcohol is high	and if PC gasoline is low	then GgreaterThanA is no.
If FP alcohol is high	and if PC gasoline is average	then GgreaterThanA is no.

where :

- PC alcohol stands for freezing point of alcohol,

- PC gasoline stands for freezing point of gasoline,

- GgreaterThanA is yes means that the freezing point of gasoline is higher than the freezing point of alcohol,

- GgreaterThanA is no means that the freezing point of gasoline is lower than the freezing point of alcohol.

Imagine that PC alcohol is equal to 5/10 and PC essence is equal to 7.95/10 (as were are not freezing experts, we chose to measure the freezing points on an arbitrary scale of 0 to 10). Our fuzzy system then gives the value GgreaterThanA = 66.7. With this value, we can find the membership degrees in fuzzy sets "yes" and "no":
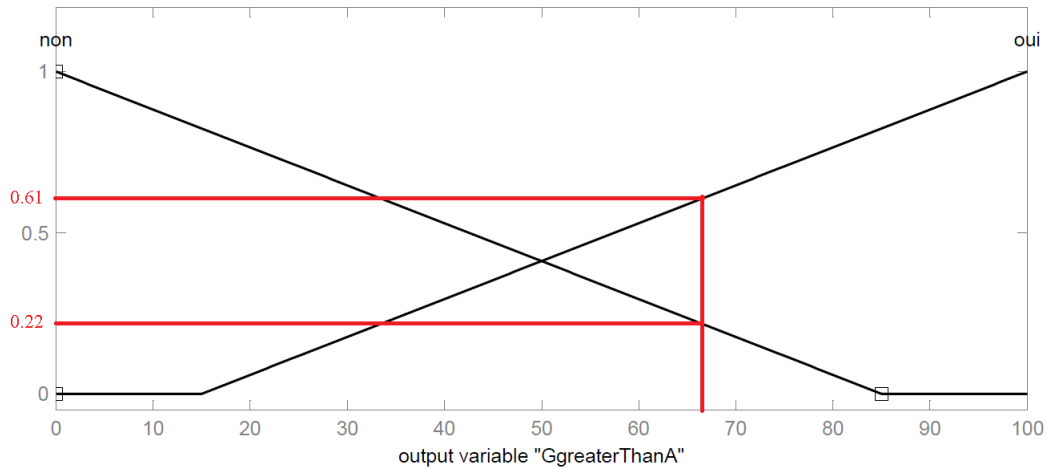
Figure 3.4: Superadditivity obtained by fuzzy logic

The degree of membership in fuzzy set "no" is $\frac{85-66,7}{85} \approx 0,22$ and the degree of membership in the fuzzy set "yes" is $\frac{66,7-15}{85} \approx 0,61$. We find that $0,22 + 0,61 < 1$, which reflects the on-additivity found experimentally in [Macchi, 1999].

More generally, the non-additivity that has been shown in many studies of experimental psychology can be modeled by choosing fuzzy variables and membership functions of their fuzzy sets.

3.3   The cognitive paradoxes of classical logic

Classical logic (binary) often result in conclusions that differ from what humans do in practice. We will see in this section whether fuzzy logic helps to explain some of these differences.

3.3.1 Sorites paradox

The sorites paradoxes arise when we ask questions like "How many grains does it take to do a heap?". The answer is difficult because in everyday language, no added grain can be identified as being the difference between a heap and not a heap. Since one single grain is not regarded as being a heap, it followed that two grains are not a heap, no more than 3, and so on. However, if we continue to add grains, after a number of grains that is not defined, we eventually get a heap.

In classical logic, the proposition "T is a heap" raises problems as it must be either true or false. However, we have just seen that there is no threshold expressed in number of grains above which a non-heap becomes a heap.

In fuzzy logic, this problem is less important because we can define "T is a job" as a fuzzy set: we do not have this "threshold" of classical logic. However, we must still define a membership function.

3.3.2 Paradoxes of material implication

In classical logic, implications are of the type "If A then B". By definition, the implication is always true when A is true. Thus, the implication "if $1 + 1 = 3$ then $1 + 1 = 2$" is considered true, while cognitively many people do not accept it as true. This is a paradox that stems from the definition of implication.

In fuzzy logic, the system designer chooses all fuzzy rules of the system. Having no interest in putting unnecessary rules like "if $1 + 1 = 3$ then $1 + 1 = 2$", there will be only useful rules for the purposes of the system.

3.3.3 Hempel's paradox

Consider the statement "All crows are black" (H). This sentence is logically equivalent to "all non-black objects are non-ravens" (H'). Indeed, by the law of contraposition, $P \rightarrow Q$ is equivalent to $\neg Q \rightarrow \neg P$.

As a result, the discovery of a black raven confirms (H) and also (H'), but also that the discovery of something non-black that is not a raven such as a pink flamingo or even a gray umbrella confirms (H') and therefore (H). This conclusion seems paradoxical.

In fuzzy logic, fuzzy sets and decision matrices are determined by the system designer. Nevertheless, it is possible to redefine these sets or fuzzy rules by learning, particularly

through neural networks. This combination between fuzzy logic and neural networks is called **neuro-fuzzy** [Jangi, 1992]. We will not go into details here but will do so in the chapter 4.3.4.

3.3.4   Wason selection task

The Wason selection task is an experiment that tests mastery of modus ponens and modus tollens in human subjects [P. C. Wason, 1966]. Here is what is proposed to individuals as described in the original experiment, "Four cards with numerals on one side and a letter on the other, are placed on a table. One side of each card is visible. The visible faces are: D, 7, 5, K. What card(s) do you need to turn over to determine the card(s) that violate(s) the following rule: 'If a card has a D on one side, then it has a 5 on the other side'. Do not return card unnecessarily, nor forget to return one."

Approximately 80% subjects are wrong in their answer. The most common error, namely returning the card 5 and forget the card 7, reveals two cognitive biases:

- verification bias, which is to seek more verification than refutation of the rule,

- matching bias, which is to focus on the items mentioned in the statement.

In logical terms, forgetting the card 7 shows poor control of the modus tollens. Choosing card 5 corresponds to the fallacy of affirming the consequent, that is to say, to confuse a simple relation of implication with a relation of logical equivalence.

Fuzzy logic does not predict the fallacy of affirming the consequent, but can still simulate it without any theoretical problem: the confusion between logical implication and equivalence can not be held unless it is deliberately introduced by the designer system. The same applies to the modus tollens.

Note that a few research groups are currently focusing on the fuzzy modus tollens, including the development of reasoning systems based on experience ([Zhaohao Sun and Sun., 2005]).

3.4   Natural language

Human reasoning mostly rely on statements made in natural language. Classical logic can not adequately maintain the links between the propositions such as:

- "This rabbit is small"

- "Some rabbits are small"

- "Some rabbits are very small"

- "Some rabbits are not very small"

We saw in the first chapter that fuzzy logic is based on the concept of linguistic variables, which allows to introduce considerable flexibility in the characterizations. To better take into account the proposals made in natural language, Zadeh has published three long articles in 1975 to extend the notion of linguistic variables [Zadeh, 1975a] [Zadeh, 1975b] [Zadeh, 1975c], which he resumed in 1978 [Zadeh, 1978] and he laid the foundations of the PRUF theory (*Possibilistic Relational Universal Fuzzy Language*).

In particular, beyond the simple linguistic variables, PRUF introduced four linguistic concepts can be modeled by fuzzy logic:

- modifiers. Example: " Xis very far",

- quantifiers. Example: " Most Chinese have black hair"

- qualifiers. Example: " It is likely that X is far",

- composition operators. Example: " X and Y is near is far." The composition operators correspond to the operators and/or, which we have already seen the first part).

For the sake of brevity, as our point is to show the flexibility of fuzzy logic for the formalization of natural language with respect to classical logic, we will briefly explain how to define modifiers.

Modifiers are adverbs that modify a fuzzy variable in order to increase or diminish its significance as "very", "moderate" or "more". For example, this allows us to obtain the fuzzy set "very dynamic" from the fuzzy set "dynamic" and " very".

Definition 10.
*A **linguistic modifier** is an operator m characterized by a function t_m such that if A is a fuzzy set characterized by membership function f_a, m(A) is a fuzzy set characterized by $f_{m(A)}$, with $f_{m(A)} = t_m(f(A))$.*

Some examples of modifiers introduced initially in the articles of Zadeh:

- "'very"' : $t_m(x) = x^2$;

- "'more or less"' : $t_m(x) = \sqrt{x}$;

- "'not"' : $t_m(x) = 1 - x$;
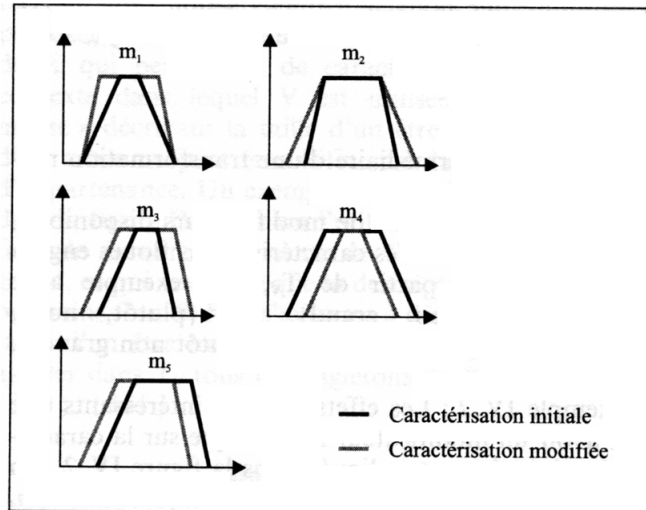
Figure 3.5: Effect of modifiers

Another way to define a modifier is to represent it directly as a fuzzy set [Ribeiro and Moreira, 2003], which allows to handle the same type of elements in the fuzzy system:
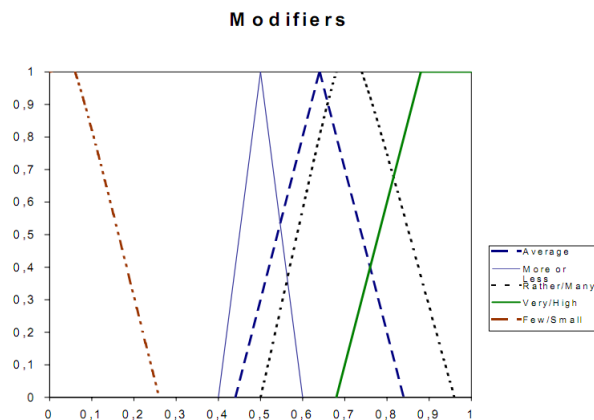
Figure 3.6: Membership functions of modifiers
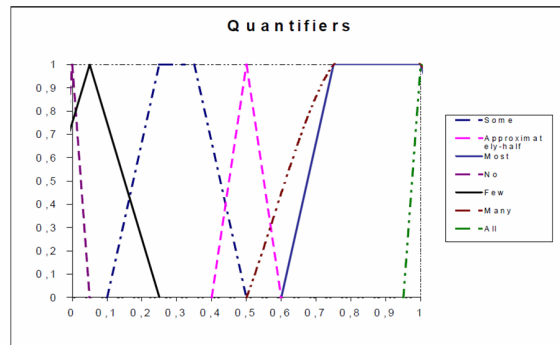
Figure 3.7: Membership functions of quantifiers
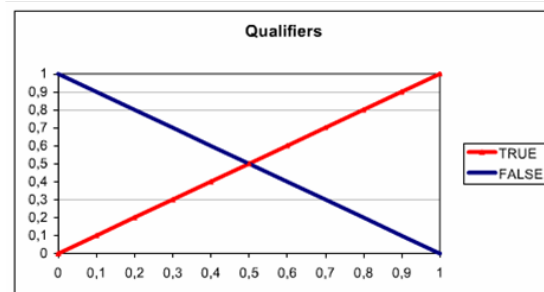
Figure 3.8: Membership functions of qualifiers

3.5 Conclusions

The purpose of this chapter was to assess the "proximity" between human reasoning and fuzzy logic. For this purpose, we studied several aspects of human reasoning that are difficult to model and experimentally undermined classical probabilies and logic.

Fuzzy logic seems to provide a satisfactory answer for the non-additivity of probabilities observed in human subjects. Fuzzy logic can also address or mitigate certain paradoxes that appear in classical logic. Nevertheless, it would be useful to widen the study to better analyze some aspects of these paradoxes in fuzzy logic, including the formalization of the Wason selection task.

An undeniable strength of fuzzy logic is its proximity to the natural language. In essence, the rules of the decision matrix are expressed in natural language, and the

basis of fuzzy logic based on fuzzy set theory that lends itself very well to express the linguistic variables.

Fuzzy logic seems to provide a good theoretical framework for human reasoning, at least its imitation. We will see in the next chapter if we can find experimentally traces of fuzzy inference in the decision made by human subjects as part of tasks that will be affected to them.

Chapter 4

Experiments

La vérité n'est pas l'exactitude. Henri Matisse.

4.1 Objectifs

There are few studies comparing the results of inference systems based on fuzzy logic with the results that come from human reasoning. In this chapter we will analyze two aspects of human reasoning through two experiments.

Firstly we will analyze a research paper in the field of medical informatics in which the authors attempted to design a system for decision support based on fuzzy logic to help anesthesiologists during surgery. The prospect will be then to study the type inference in humans when we give them very precise, certain objective variables measured by medical instruments.

In a second step, we will do our own experiment in the world of video games. We will study how gamers give scores to video games based on two criteria they evaluate themselves, contrary to previous experience when evaluating these two criteria is not made by humans but is given by an external source.

4.2 Decision system support for anaesthesiologists

In the study [Hamdi Melih Saraoglu, 2007], the authors have worked on the decision-making of anesthesiologists during surgery. An anesthesiologist decides the content of the gas as the patient breathes continuously: this gas must be more or less anesthetic according to the patient's condition at a given time of the operation.

To make this decision, the anesthesiologist is based on two main criteria (input):

- SAP: systolic arterial pressure (i.e. blood pressure),

- HRP: heart pulse rate.

Based on these two criteria, the anesthesiologist will take a decision (output):

- AO: rates of anesthetics in the gas (anesthesia output).
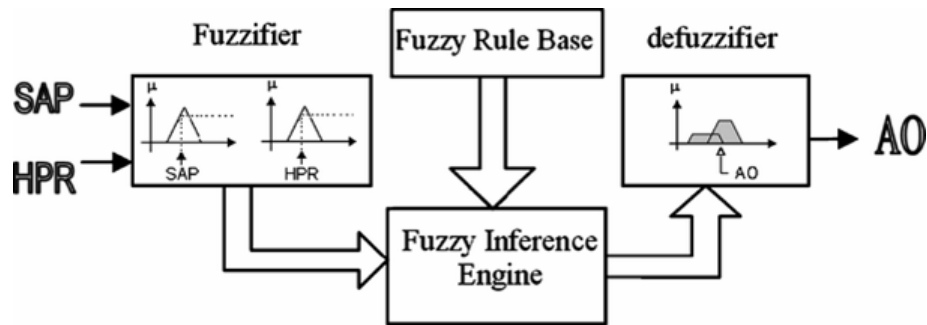
Figure 4.1: Overview of the inference system

Initially, the study authors define the artificial fuzzy system and in a second time they put it in a real situation in order to compare these results with the decisions made by human anesthesiologists. We will first briefly explain the fuzzy system chosen.

First, here are the fuzzy sets and the decision matrix:
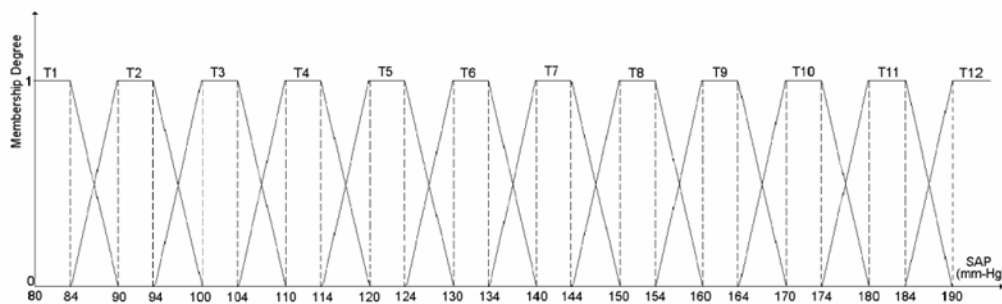
Figure 4.2: Fuzzy sets of the variable SAP
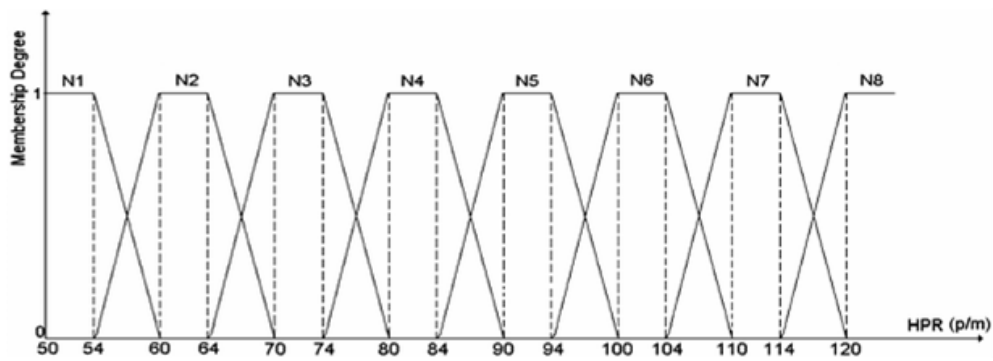
Figure 4.3: Fuzzy sets of the variable HPR
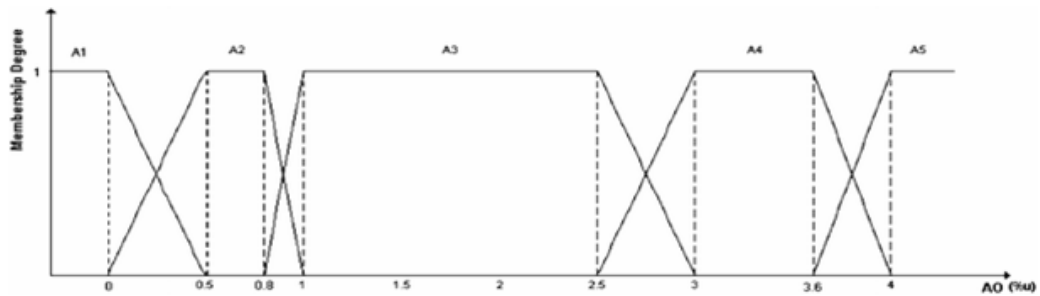
Figure 4.4: Fuzzy sets of the variable AO

	N1	N2	N3	N4	N5	N6	N7	N8
T1	A1	A1	A2	A2	A2	S	S	S
T2	A2	A2	A3	A3	A3	A4	A4	A4
T3	A2	A3	A3	A3	A3	A4	A4	A4
T4	A2	A3	A3	A3	A3	A4	A4	A4
T5	A2	A3	A3	A3	A3	A4	A4	A4
T6	A2	A3	A3	A3	A3	A4	A4	A4
T7	A2	A3	A3	A3	A3	A4	A4	A5
T8	S	A4	A4	A4	A4	A5	A5	A5
T9	S	A4	A4	A4	A4	A5	A5	A5
T10	S	A4	A4	A4	A4	A5	A5	A5
T11	S	A5	A5	A5	A5	A5	A5	A5
T12	S	A5	A5	A5	A5	A5	A5	A5

Figure 4.5: Decision matrix

Here the choices of operators and an example of fuzzy inference. This classic choice

using the Zadeh operators and the Mamdani implication is called the Mamdani's
fuzzy inference method.

Operator	Name
AND	min
OR	max
Implication	min
Aggregation	max
Defuzzification	COG
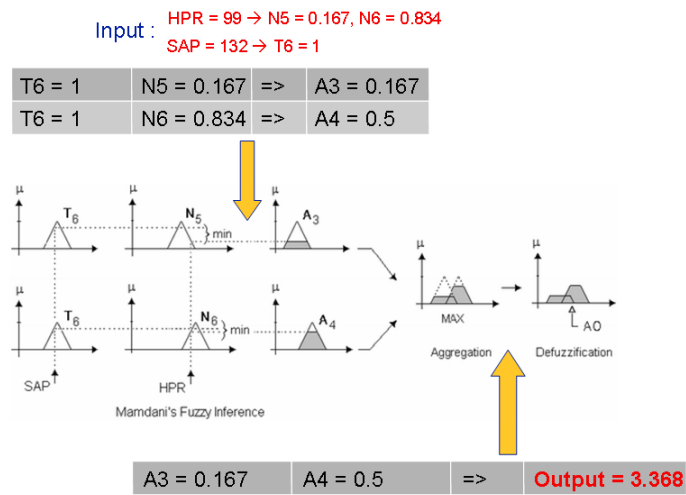
Figure 4.6: Example of inference

Now that the fuzzy system is established, let's compare it with the decisions taken
by anesthesiologists:

												Total
Patient 1	SAP (mmHg)		164	161	192	156	172	161	152	157	154	
	HPR (p/m)		65	96	99	80	72	81	73	78	73	
	AO	Anesthetist	2	2.5	4	3	3.5	2.5	2.5	2.5	2.5	25
		Fuzzy Logic	3.25	3.29	3.88	2.18	3.27	3.27	1.83	2.38	1.83	25.2
	Distance %											+1
Patient 2	SAP (mmHg)		149	184	139	182	163	158	158			
	HPR (p/m)		88	94	98	74	78	90	87			
	AO	Anesthetist	2	4	2	5	4	2	2			21
		Fuzzy Logic	1.85	3.27	2.61	3.27	3.27	2.61	2.52			19.4
	Distance %											−7.6
Patient 3	SAP (mmHg)		110	130	130	110	130	120	120			
	HPR (p/m)		90	70	70	68	72	75	76			
	AO	Anesthetist	2.5	2	2	2	2	1.5	1			12
		Fuzzy Logic	1.83	1.83	1.83	1.85	1.83	1.84	1.85			11.1
	Distance %											−7.5
Patient 4	SAP (mmHg)		150	110	115	110	110					
	HPR (p/m)		96	74	107	95	90					
	AO	Anesthetist	3	2	2	2	2					11
		Fuzzy Logic	2.18	1.83	3.26	2.00	1.83					11.1
	Distance %											+1

Figure 4.7: Comparison between the decisions of a fuzzy system and those of a human

This table provides a comparison between doses of anesthetic products given by the anesthesiologist and those given by the fuzzy system every five minutes based on the SAP and HPR variables of the patient. Four operations of different durations are compared: the operation of patient 1 lasted 45 minutes, the patients 2 and 3 35 minutes, and the patient 4 stayed on operating table only 25 minutes.

The article and the previous table, from which it is drawn, point out that the sum of the doses of anesthetics at the end of the operation is substantially similar between the anesthesiologist and the fuzzy system: for example, for the first operation, the anesthesiologist gives 25 AO while the fuzzy system gives 25.2 AO, which is virtually the same. The article concludes that the fuzzy system is efficient enough to mimic the decisions of anesthesiologists.

However, comparing only the sum of AO given is not at all satisfactory. Indeed, it could very well happen that the first decision of the dose of anesthetics to provide to the patient is lethal, and subsequent doses are almost nil, which would give an amount comparable to the sum of a prescription made by an anesthesiologist.

The Pearson correlation coefficient allows us to have a better idea as to the correspondence between experimental and theoretical results. Of all the operations, the Pearson correlation coefficient is 0.627, which corresponds to a moderate approval,

but often found in medical decisions when analyzing the difference between 2 physicians [Caroff, 2010].

A better indicator comparing experimental results with the results of the model is the square root of the mean square error (RMSE). In particular, the RMSE will allow us to compare several models.

Definition 11.

$$Let\ \mathbf{A}_1 = \begin{bmatrix} x_{1,1} \\ x_{1,2} \\ \vdots \\ x_{1,n} \end{bmatrix} \quad and \quad \mathbf{A}_2 = \begin{bmatrix} x_{2,1} \\ x_{2,2} \\ \vdots \\ x_{2,n} \end{bmatrix}.$$

The square root of the mean square error is calculated by the formula:

$$\mathrm{RMSE}(\mathbf{A}_1, \mathbf{A}_2) = \sqrt{\mathrm{MSE}(\mathbf{A}_1, \mathbf{A}_2)} = \sqrt{\mathrm{E}((\mathbf{A}_1 - \mathbf{A}_2)^2)} = \sqrt{\frac{\sum_{i=1}^{n}(x_{1,i}-x_{2,i})^2}{n}}.$$

Of all operations, we find a RMSE equal to 0.6877. This will allow us to compare these results with the following experiment.

4.3 Experiment of video games rating

4.3.1 Presentation

Now let's look at the case where the human makes a decision based on criteria which he assess the value himself. We designed an experiment for this report in order to study this type of reasoning.

In short, the experiment asks seven subjects passionate about video games (*gamers*) to give scores to video games. Each of them will give a score to thirty video games and three days later, we will ask them to give a score to two criteria comprising all aspects of a video game. We then study the relationship between the notes of the two criteria with the overall score, to see if such an inference system based on fuzzy logic may have a similar relationship.

4.3.2 Protocol

Each subject will have the following profile: male, between 20 and 30, gamer. The experiment will be conducted in two sessions which last approximately 15 minutes each, separated by three days. Full details of the protocol are in Annex B.

During the first session, they will be asked to give a score to thirty games of their choice on a scale of 0 to 100, a high score meaning that the player enjoys the game

During the second session, they will be asked to give a score to always between 0 and 100 two criteria comprising all aspects of a video game. The two criteria:

- Criterion 1: technical aspects, including graphics, fluidity, quality of the interface, quality of the controls, online mode and the number of addons/extensions.

- Criterion 2: non-technical aspects, including the duration of the game, story, fun to play, interest in strategies, business community and the possibility of learning.

4.3.3 Results

Here is the graph of a subject and the graph for all subjects. A blue dot corresponds to the rating of a video game. Our data are in three dimensions:

1. Dimension 1: technical score

2. Dimension 2: non-technical score

3. Dimension 3: Overall score
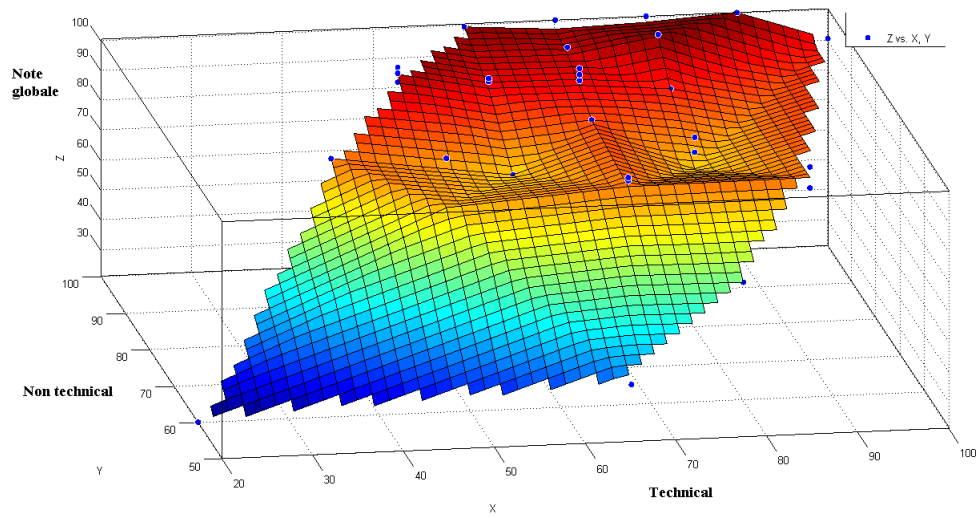
Figure 4.8: Surface generated from the scores of the subject 1 for the 2 criteria
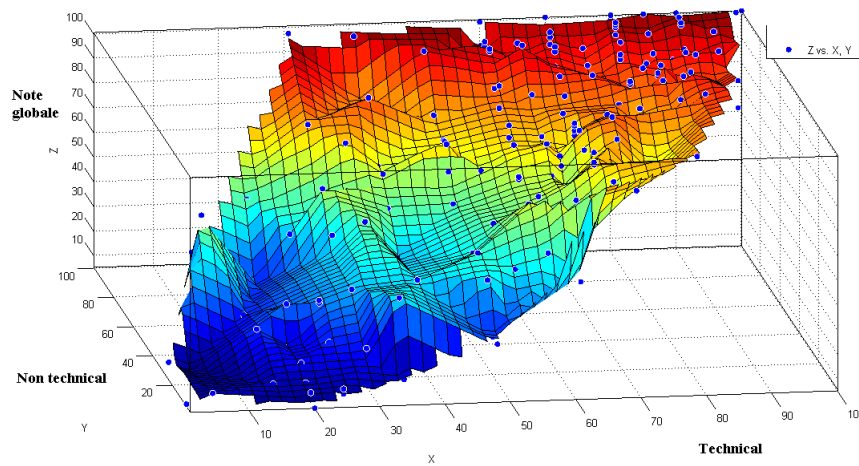
Figure 4.9: Surface generated from the scores of all subjects for the 2 criteria

We will compare the RMSE obtained with a polynomial prediction model with that obtained with a model based on a fuzzy system. The polynomial model with a RMSE equal to 7.91 when the degrees of X and Y is 1, and drops to 6.097 when X and Y have a degree equal to 5.

We must now develop a model based on a fuzzy system. Specifically, we need to

define the fuzzy sets as well as the fuzzy rules of the decision matrix. The problem is that if we manually created the model, we would have no assurance of optimality and therefore the comparison of RMSE would not be necessarily relevant. In order to optimize as much as possible our fuzzy system, we will refine our parameters using neural networks. Such systems are called neuro-fuzzy systems.

4.3.4 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.
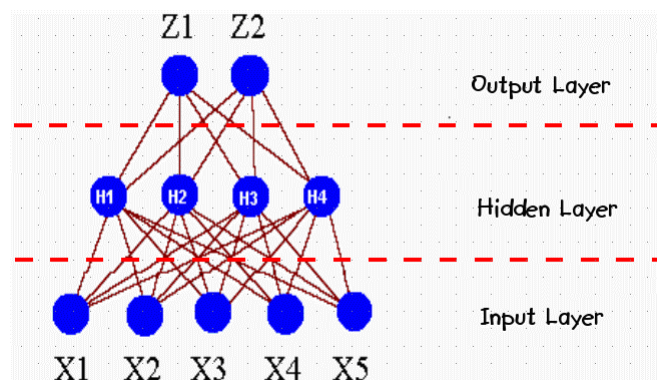
Figure 4.10: Example of a feedforward neural network

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.

Here is the organization of our multilayer perceptron representing the neuro-fuzzy system which we will use to model the scoring system:
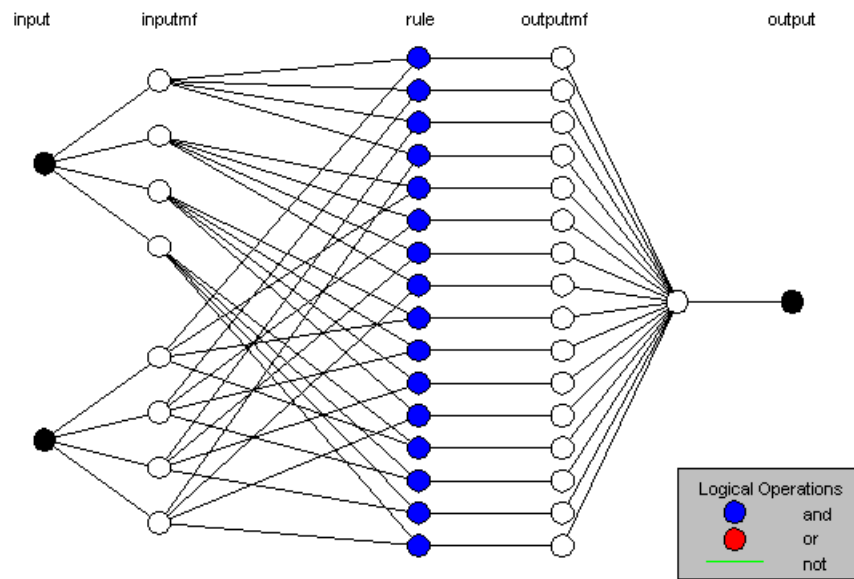
Figure 4.11: Structure of a neuro-fuzzy system

The field being new, we had to take into account several constraints in order to find a model for our video game scoring experiment. We had to fix before the learning phase:

- the number of input fuzzy sets: 4 per input,

- the shape of the membership function: Gaussian.

Here is the learning curve for our fuzzy system:
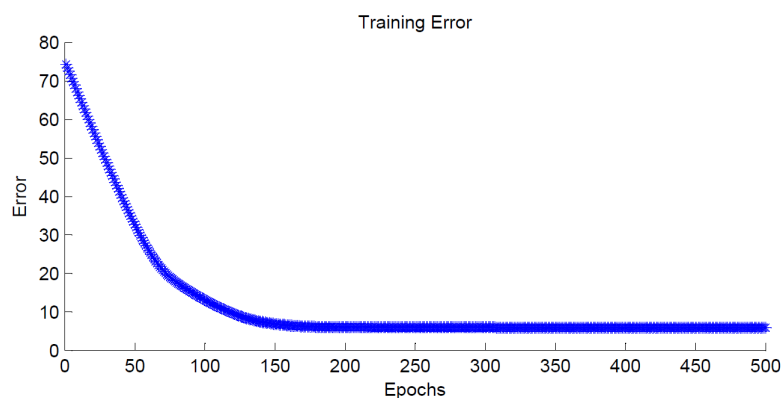
Figure 4.12: Learning curve of the neuro-fuzzy system

We obtain an RMSE of 5.9 after a few hundred steps of learning, far less than the RMSE of a polynomial of degree 1, and slightly lower than the RMSE of a polynomial of degree 5. This result is quite good, here is the decision surface obtained:
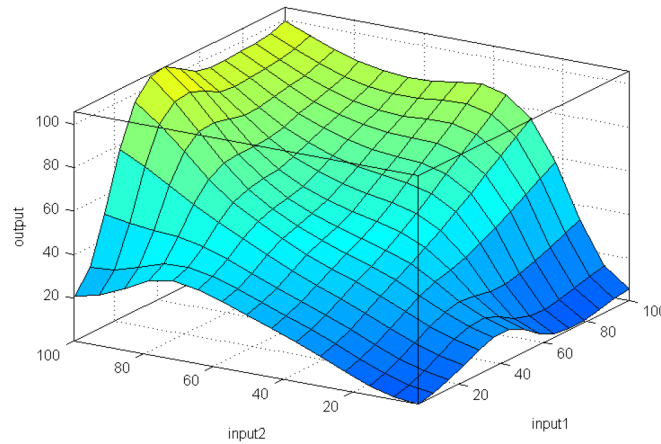
Figure 4.13: Decision surface of the neuro-fuzzy system

We notice that the decision surface of the fuzzy system we obtained has a form between a linear surface and a *classique* decision surface of a fuzzy system. The Pearson correlation coefficient between the two inputs is 0.786, meaning that subjects tended in the second stage of the experiment to be influenced by the technical non-technical notes that they had given before, or vice versa. However, even in these conditions which favor a polynomial model, the fuzzy model still has better results.

We see that the experimental data are close to the theoretical model, with the exception of two areas where extreme notes given to the technical side is very different from the note given to non-technical aspects, which is a case that never happens in practice because if the technical aspect of the game is really bad, it negatively impacts non-technical aspects, and vice versa if the non-technical aspects are too bad, the player stops playing right away before he can really assess the technical side.

Let's have a look at the decision matrix of our fuzzy system optimized by neural networks. As we have seen, we set to four the number of fuzzy sets for each of the two inputs:

- the set of fuzzy sets of input 1: $\{in_1mf_i, i \in [\![1,4]\!]\}$;

- the set of fuzzy sets of input 2: $\{in_2mf_i, i \in [\![1,4]\!]\}$;

- the set of fuzzy sets of output: $\{out_1mf_i, i \in [\![1, 4 \times 4]\!]\}$.

The set of fuzzy rules of our system is:
$\{If\ in_1mf_i\ and\ in_2mf_j\ then\ out_1mf_{4\times(i-1)+j}, (i, j) \in [\![1, 4]\!]^2\}$ which appears cognitively plausible.

4.3.5 Comparison with the previous experiment

For the experiment with anesthesiologists, we created a neuro-fuzzy model to see if we could improve the system created by hand in the article. Here are the decisions of our fuzzy system with respect to the decisions of the anesthesiologist during the four operations for which we have the data:
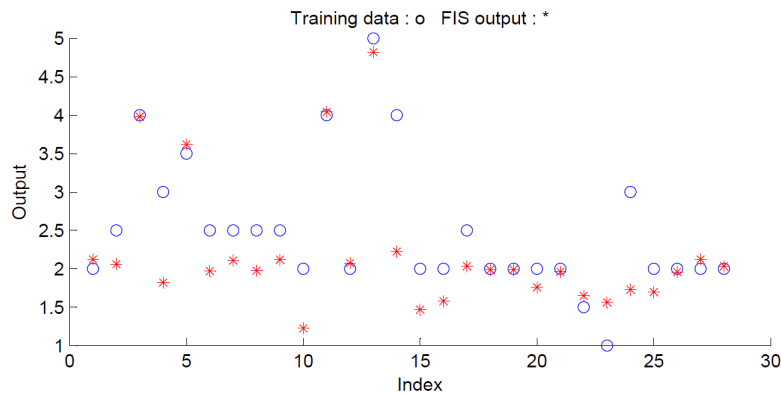
Figure 4.14: Decision surface of the neuro-fuzzy system

Here is the summary of the RMSEs:

Experiment	Model	RMSE*	Standardized RMSE **
Anesthesiologist	Neuro-fuzzy system (by hand)	0,6877	13,754
Anesthesiologist	Neuro-fuzzy system	0,569	11,38
Game scoring	Polynomial model with degree 1	7,91	7,91
Game scoring	Polynomial model with degree 5	6,097	6,097
Game scoring	Neuro-fuzzy system	5,9	5,9

* RMSE between model and experimental data
** RMSE standardized (i.e. normalised) because the RMSE is influenced by the

scale. The output of the experiment with anesthesia has values between 0 and 5, whereas the output of the experiment with the scores has values between 0 and 100. Therefore, we normalize the RMSE by multiplying by $\frac{100}{5} = 20$.

We see that for the experiment with the anesthesiologists, our fuzzy system has a slightly better RMSE than that proposed in the article. However, it should have more data because the article gives the data of only 4 operations, which gives us only 28 individuals in the statistical series, which is very inadequate to effectively optimize a fuzzy system with neural networks. Here are two decision surfaces. The first surface is the one that allowed us to obtain a RMSE equal to 0.569: we see that the form is quite correct except for high values of input 2, reflecting the lack of data there. The second is an example of over-learning increased by an excessive number of fuzzy sets, which has a RMSE of 0.212.
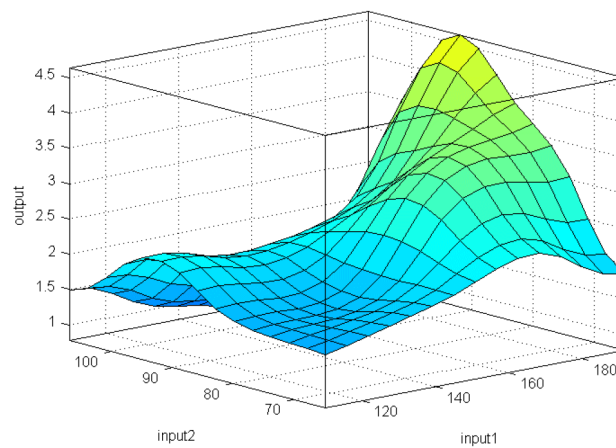
Figure 4.15: Decision surface of the neuro-fuzzy system
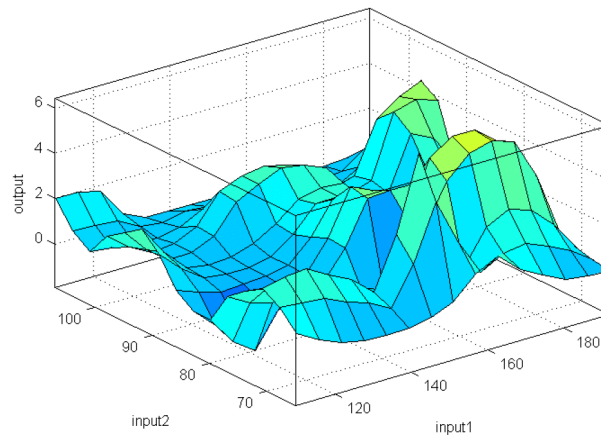
Figure 4.16: Decision surface of a over-learned neuro-fuzzy system

To compare the RMSE between the two experiments, after normalization we see that our fuzzy system for the game experiment has a RMSE (5.9) well below the fuzzy system defined by hand for anesthesiologists (13.754). This is not surprising as the number of data in the game experiment is about 10 times larger (267 against 28), which allowed the neural network of well optimized fuzzy system while making almost impossible to over-learning, because the number of fuzzy sets of inputs is limited.

It is yet difficult to compare the RMSE between the two experiments because in the first one the two input variables are almost uncorrelated, while in the second one they are more correlated.

4.4   Conclusions

In conclusion, fuzzy systems seem to be able to generate quite similar results from the decisions of human reasoning. Clearly, they are better than linear systems. In addition, fuzzy systems used in these experiments, particularly for the second, are very simple: two inputs and three or four fuzzy sets for each.

Fuzzy logic makes it possible to imitate in these examples some of the reasoning of the human subject, if we first define the parameters of fuzzy inference system (membership function, choice of the rule of implication and so on).

Chapter 5

Decision support systems

The most measurable sometimes hides the most important. René Dubos.

5.1 Introduction

Now that we have shown the contribution of fuzzy logic and found that decision-making issue based on human reasoning can be imitated to some extent by an expert system based on a fuzzy inference system, we will see in this chapter how databases and decision support systems can take advantage of that.

5.2 Fuzzy databases

5.2.1 Définitions

A **database** is a set of information stored in a computer device. A **database management system** (DBMS) is a software that can read or modify a database. A **data model** defines the organization of semantic entities stored in a database.

Here are the main data models, in chronological order [Gardarin, 2005]:

- **Hierarchical model**: introduced by IBM's IMS in 1964, data are represented by a tree structure where the only possible relationship is the association of father/son, access to data always being by his father. This model has significant limitations, particularly to describe complex structures.

- **Network model**: also appeared in 1964, driven by IDS2, this model is more flexible than the previous one, because the relationship between records can exist without restriction. However, this model requires navigating along the

base via a series of pointers to access an object, making the program dependent on the data structure as the hierarchical model.

- **Relational model**: founded by E. F. Codd in 1970 [Codd, 1970], this model is based on the notion of relation, which corresponds to a set of attributes defining a fact, e.g. a student has a name, and has attended a class since 09/21. Mathematically, this model is based on relational algebra.

- **Object model**: this model is the variation of the object paradigm for databases and appeared around 1985 [Atw, 1985]. Records are in the form of object, which allows among other things better communication with object oriented languages and benefit from the power of the concept of object to express complex, composite and polymorphic entities.
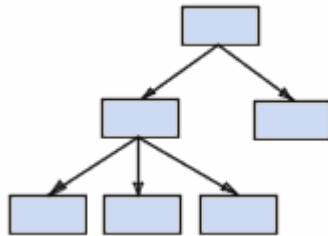
Figure 5.1: Hierarchical model
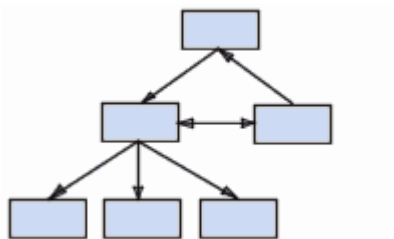
Figure 5.2: Network mode
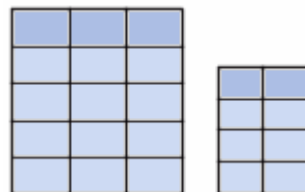
Figure 5.3: Relational model
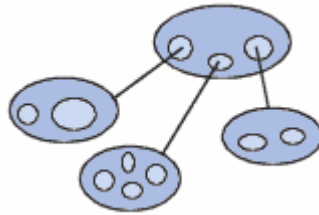
Figure 5.4: Object model

In this chapter we will focus on the relational databases (RDB), because the most common data model currently is the relational model. However, some of our ideas can be easily extended to object databases.

5.2.2 The relational model

The relational model is based on relational algebra defined by E. F. Codd as well as on the relational calculus. The relational algebra is very similar to classical set theory and defines the operations that can be performed on matrices containing a set of tuples. The relational calculus is an adaptation of the predicate calculus to RDB queries.

Relational algebra and relational calculus have the same expressive power: therefore, all queries that can be made using one can be expressed with the other. It was first proved by EF Codd in 1972 through an algorithm called *Codd's reduction algorithm* with which any expression of the relational calculus can be formulated equivalently in relational algebra.

The relational model thus finds its roots in first-order predicate calculus as well as in set theory. Therefore, it has the same drawbacks as the latter, including the difficulty to represent imperfect information. The idea behind fuzzy databases is to introduce at the heart of the RDB mechanisms of fuzzy logic to make handling of imperfect data more natural.

5.2.3 SQL

SQL (*Structured Query Language*) is as its name suggests a query language: it is used to perform operations on databases. It is recognized by nearly all relational DBMS.

SQL is a declarative language, i.e. it allows to describe the desired result, without describing how to get it, as opposed to imperative languages such as C or Java, where we describe the technical solution to obtain the desired result. In addition, its syntax aims to be close to natural language. As a result, SQL hides rather well the fact that, like relational databases, it is based on relational algebra and relational calculus.

Here is the form of a standard query expressed in SQL:

```
SELECT List of attribute names of the result
FROM Name of one or more relations
WHERE Logical conditions defining the result
```

To introduce fuzzy mechanisms at the heart of databases, we will extend SQL and incorporate concepts from fuzzy logic. But first, let's see what are the limitations of relational databases by providing concrete examples.

5.2.4   Limitations of the relational model

Consider the following query: "Give all the names of the men**very rich** and **living near** the Metro station Place d'Italie"'. Imagine that the data table consists of 6 entries:

Name	Annual salary (k€)	Distance to Place d'Italie (km)
Donald	15	0,5
Marpe	25	1,9
Cobin	35	0,5
Bill	80	0,1
Trov	95	0,8
Pixou	200	1,1

If we wanted to implement this query in SQL, we should set thresholds by hand. For example:

```
SELECT Name FROM Inhabitants
WHERE Salary > 75 AND Distance < 1;
```

This query returns Bill and Trove. We could refine the query:

```
SELECT Name FROM Inhabitants
WHERE (Salary > 60 AND Distance < 0,3)
```

```
OR (Salary > 80 AND Distance < 0,7)
OR (Salary > 90 AND Distance < 1);
```

This query returns Trove. We see that such a mechanism is still suffering from the adverse effects of thresholds: in our example, Pixou was never returned whereas he could certainly be of interest to the person making the query. This is easy to notice because we see the data at the same time we read requests. However, in a real case, these threshold effects are extremely dangerous because we can easily omit very interesting results. The figures 5.5 and 5.6 we saw in Chapter 2 illustrate this threshold effect and how fuzzy logic can address them.
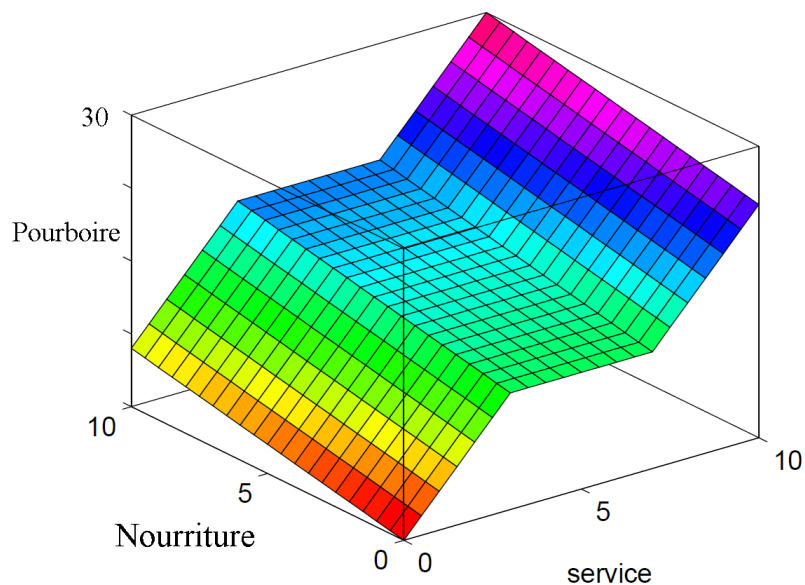
Figure 5.5: Threshold effect in conventional SQL queries

Furthermore, the notion of "rich" is partly subjective: if Pixou seems very rich for the average person, it is not certain that Bill Gates thinks so. Beyond the personal perception of a linguistic variable, we can well imagine that depending on the reasons for the query, memberships to categories differ. For example, if the request was present at the heart of a decision support system of a bank, it is clear that the bank wants to set itself the membership functions for particular categories. It would therefore be useful to be able to easily update the membership functions: in standard SQL language, this would mean to redefine the limits everywhere, with the risk of inconsistencies that it would bring. With a fuzzy mechanism, we only need to change the membership functions, which can be done graphically if the user requests it.
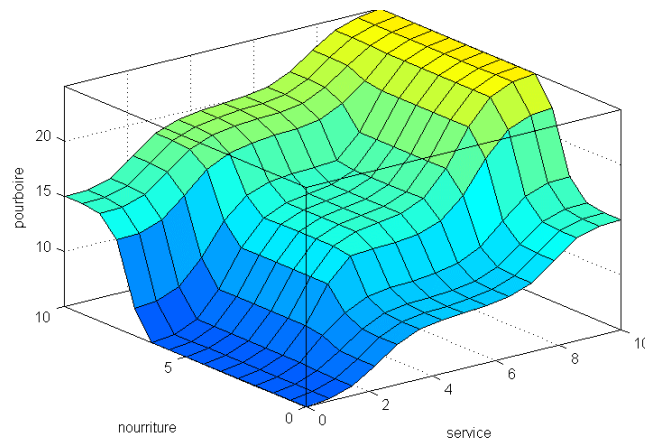
Figure 5.6: Decisions of a fuzzy system

In addition, it would be time-consuming to update this kind of queries and as the number of conditions in the *WHERE* condition, the size of queries explodes if we try to reduce some threshold effects for these new conditions. Suppose that in the previous query we wanted to add a condition, such as the size of the individual has be to very large. Our query looks like this:

```
SELECT Name FROM Inhabitants
WHERE (Salary > 60 AND Distance < 0,3 AND Height > 2,1 )
   OR (Salary > 65 AND Distance < 0,3 AND Height > 2,0 )
   OR (Salary > 70 AND Distance < 0,3 AND Height > 1,9 )
   OR (Salary > 70 AND Distance < 0,3 AND Height > 1,9 )
   OR (Salary > 80 AND Distance < 0,7 AND Height > 1,9 )
   OR (Salary > 90 AND Distance < 1,0 AND Height > 2,1 )
   OR (Salary > 100 AND Distance < 1,0 AND Height > 2,0 )
   OR (Salary > 110 AND Distance < 1,0 AND Height > 1,9 );
```

Finally, another problem arises: we have tried here individuals with the characteristic "very rich". Perhaps tomorrow we will want to find individuals with the characteristic "rich". As we formed the query, we have completely removed the compositionality of natural language. In other words, our previous SQL considers "very rich" as a single adjective, not an adjective (*riche*) qualified by a modifier (*very*). This simplification of natural language has the effect of not being able to define the notion of "rich" from "rich" and "very ", which leads to potential inconsistencies and increased complexity of the definition of thresholds.

As we see, to summarize our remarks above, this type of query is:

- easyly made inconsistent,

- difficult to update,

- subject to severe threshold effects,

- opaque with respect to the compositionality of natural language,

- painful to read.

Thus, the classical relational model has many limitations on the type of queries that we can do it. But there are also important limitations on the nature of the data that we store. The question is: how can we save an imperfect datum such as "client X's size is fairly tall"? Ignoring this information is potentially a loss: we must find a method to have databases contain this information.

At first glance, the idea would be to record the fact that X is 1.90 meters, by considering that on average a fairly tall person measures 1.90 meters. This solution quickly shows its weaknesses when the nature and extent of the imperfections of the data size changes depending on the individual. For example, if we know that Y actually measures 1.90 meters, then the size recorded for Y should be distinguishable from the size recorded for X.

Another finer solution would be to record both the measure and its degree of imperfection. For the previous example, we could save the measurement of X as the pair (1.90, 0.5) and measurement of Y as the pair (1.90, 0). With such couples, we could make the difference between the size of X and Y and quantify the range for the size of X. However, we should store it somewhere what means an imperfection of 0.5 and we could not introduce different types of defects for the same attribute.

Moreover, just like for queries we find it impossible to transcribe the compositionality of natural language in the records of the database. Thus, if we want to record the fact that the size of Z is very tall, we can not establish any link with the notion "fairly tall".

Therefore, it seems natural and useful to introduce fuzzy mechanisms to fill the shortcomings of the relational model based on classical Boolean logic.

5.2.5   Integration of fuzzy logic in a database

Many studies have examined the issue of integration of fuzzy logic in a database, such as [Buckles and Petry, 1982], [Zvieli and Chen, 1986] and [Galindo et al., 1999], however implementations are rare and often experimental [S. Skrbic, 2008]. The idea is to make possible the storage of information regardless of their imperfections. The database should provide a means to record the meaning of fuzzy information it contains.

For each attribute, we must define the universe of discourse, that is to say the set of all the fuzzy sets a variable can have. For example, the attribute "temperature" admits a set of fuzzy sets of type "hot", "cold", "warm". This kind of attributes corresponds to the linguistic variables seen in previous chapters. We can also make other types of attributes, such as fuzzy numbers. Theoretically, for each fuzzy set, we should have an algorithm to calculate the values of its membership function. In practice, the values commonly used follow standard forms (triangular, trapezoidal, Gaussian,...), as shown in Figures 5.7, 5.8 and 5.9 which illustrate membership functions for fuzzy number 55.
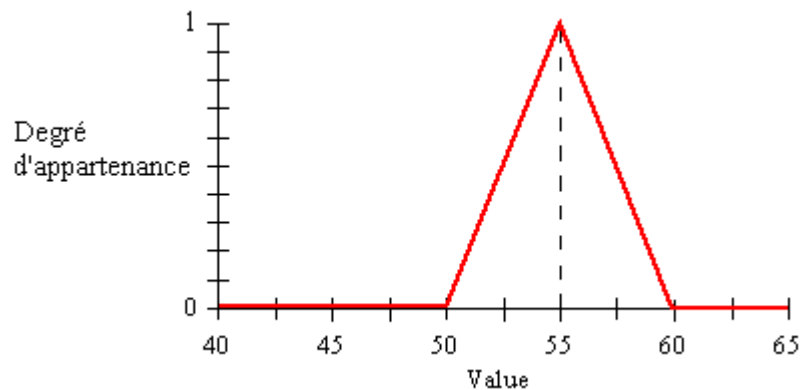
Figure 5.7: Fuzzy number - triangular form
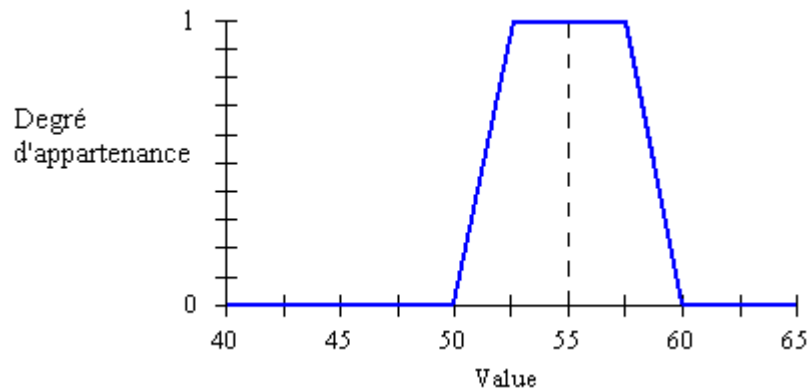
Figure 5.8: Fuzzy number - trapezoidal form
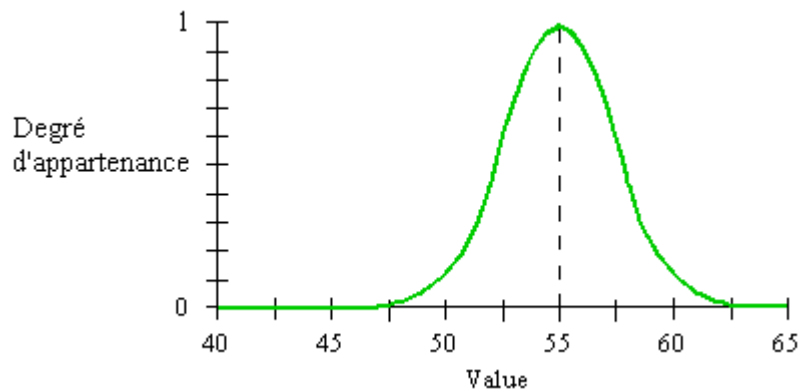
Figure 5.9: Fuzzy number - Gaussian form

To manipulate such data, an extended relational algebra was introduced by [Medina et al., 1994] and [P. Bosc, 1998]. We will not go into detail for the sake of brevity. The technical expression for this extension is found in the theoretical generalizations of SQL to fuzzy SQL as SQLf (f to fuzzy) introduced by [P. Bosc, 1995]. We will briefly present the characteristics of SQLf.

LThe structure of a SQLf query always follows the usual three clauses of SQL *SELECT*, *FROM* and *WHERE* (we shall speak here only of read requests). The clause *FROM* is identical with the standard SQL. The clause *WHERE* is the most interesting because this allows to contain Boolean conditions as well as fuzzy terms. Finally, the clause *SELECT* is close to the standard SQL and simply offers a new optional parameter to define a threshold below which the results are eliminated: indeed, the clause *WHERE* may contain conditions fuzzy which are processed via an extended

relational algebra that incorporates fuzzy mechanisms and the results returned have a membership degree with respect to all conditions and operators of *WHERE*.

Here is how a typical SQLf request looks like, to use the example of the previous section:

```
SELECT THRESHOLD 0.6 Name FROM Inhabitants
WHERE Salary = very high
  AND DistancePlaceItalie = small
```

In order to run this query, the variables wages and distance must first be defined, this time using fuzzy sets if information is imperfect. After that, we will indicate in the database what *rich* and *very* means. Then, we will choose the operator corresponding to the fuzzy *AND*. A fuzzy metaknowledge base will therefore be necessary to contain all information about the attributes.

The figure 5.10 shows the organization of a fuzzy database. We see that it contains two additional modules that fit into the context of a typical database:

- A layer containing a parser for analyzing the fuzzy query is added,

- A module containing the various characteristics pertaining to fuzzy (attributes, domain, language variables,...) is created in parallel with the classical database to contain the meanings of the fuzzy data present in the typical database.
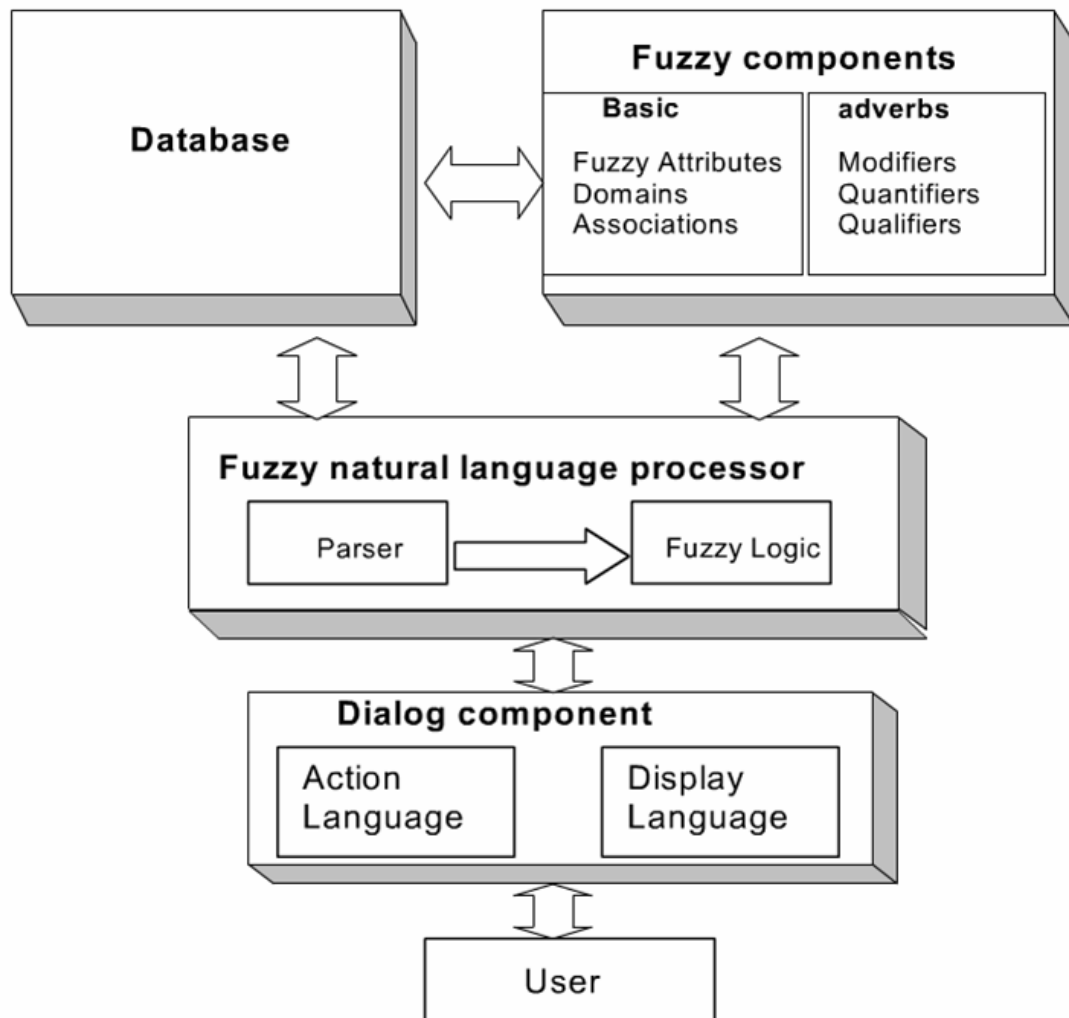
Figure 5.10: Organization of a fuzzy database

Unfortunately, there is very little implementation of fuzzy databases, and to our knowledge, all those that do exist are experimental. We briefly tried the software FQ(http://www.lcc.uma.es/ ppgg/FSQL) that implements FSQL, however, in addition to being only available in Spanish, it works exclusively with Oracle 8. As a result, we have just tested the client and not the server. Figure shows its graphical interface, simple and intuitive. A comprehensive study for another FSQL client, *Visual FSQL*, can be read here: [Oliva, 2003].
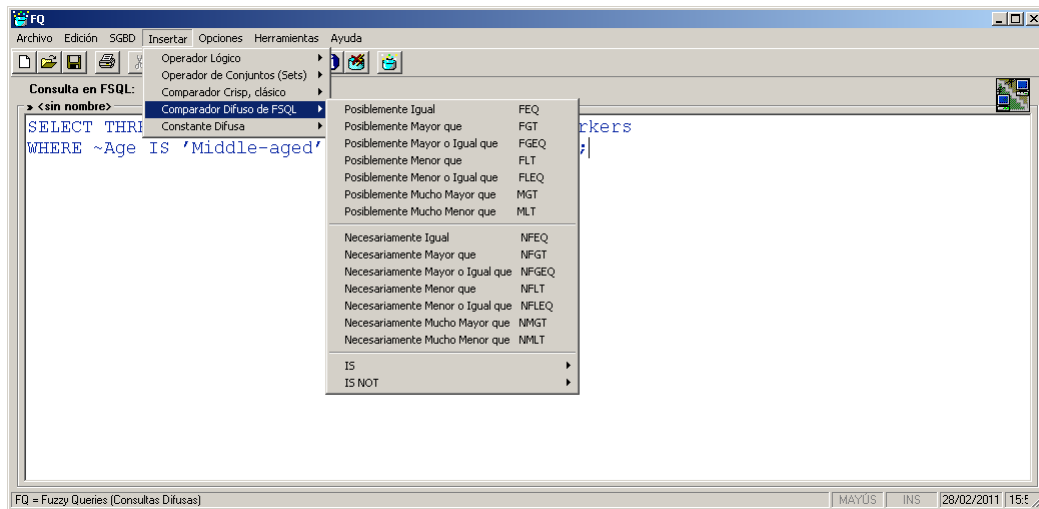
Figure 5.11: Screenshot of the client FQ

This gap between the prolific literature on fuzzy databases and the small number of serious implementations reflects the lack of a scientific or industrial consensus. In this section, we presented an overview of the foundations of fuzzy databases, but there are many theories and different implementations. Several implementation methods have been developed on various DBMS (MySQL, Oracle, Access, etc) and many extensions of SQL have been proposed, such as:

- SQLf [P. Bosc, 1995]

- FQUERY [Zadrozny and Kacprzyk, 1996]

- Summary SQL [Rasmussen and Yager, 1996]

- FSQL [Andreasen et al., 1997]

- SQLf_j [Dembczynski et al., 2004]

- ...

5.3 Potential contributions to decision support systems

Now that we have seen the interest and methods to incorporate fuzzy logic to databases, we will analyze the various applications of fuzzy logic, more generally, in a decision-making system. A decision support system is designed to promote the

decision-making: we have seen that fuzzy logic is a great tool to mimic the decision-making issue of human reasoning. In this section we will present several possible implementations of fuzzy logic.

First, let us recall briefly the various elements of the decision-making chain, presented graphically in Figure 5.12. The heart of a decision support system is the data warehouse, whether virtual or real. The data warehouse is a collection of data from operational databases of the company, which are considered sources of data. The data are extracted from operational sources into the data warehouse: for this purpose, the data is cleaned and integrated to eliminate the numerous inconsistencies within and between sources. Note that the data warehouse is not intended to bring together all the data from the sources: the objective is to only extract the relevant data sources, even in some case aggregate some, and record them perennially: the data in a data warehouse are generally never altered or erased. Once the relevant data are integrated and stored in the data warehouse, they will be used by tools available to policy makers, providing a single, cross-company: personalized dashboards, statistical studies, data mining, etc.
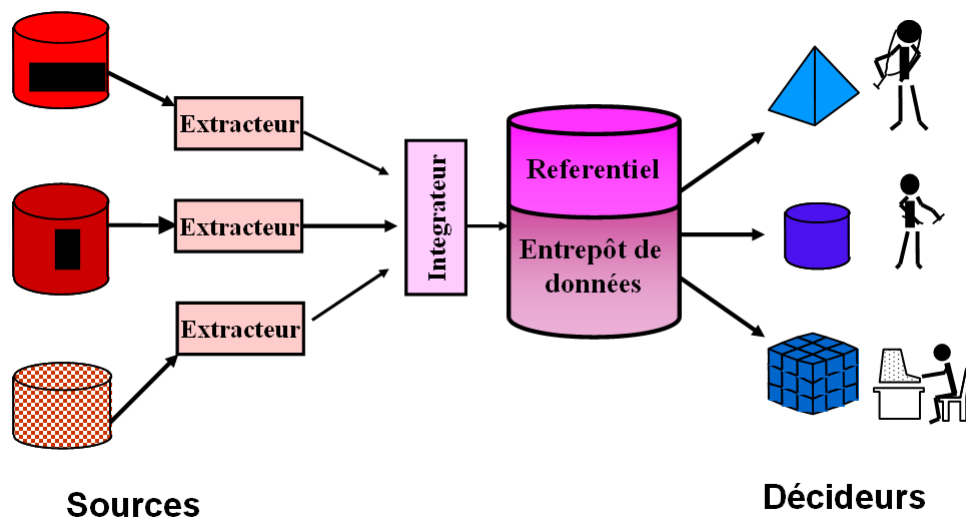
Figure 5.12: The decision-making chain

Source : Encyclopedia Universalis, 2008

As for sources, fuzzy logic would reflect the imperfections of the data, either globally at the level of a source, or for individuals or attributes in a particular source. Currently, it is difficult to record into the data warehouse the differences in data sources perfection. For example, suppose that source 1 is a very reliable source, while source 2 is so only moderately. We could well imagine the numbers to express the source 2 in

fuzzy numbers, choosing membership functions reflecting the degree of imperfection. Once transferred into the data warehouse, these numbers will not be confused with other numbers deemed reliable, thanks to the fuzzy mechanisms that will be set in order to take into account the difference in quality of information.

Fuzzy logic allows for efficient use of natural language, for example to make graduations or implications, or to account for the compositionality of language. This would offer a lot of benefits for decision-making systems:

- When new data arrive, we would like to record certain attributes in natural language, rather than setting arbitrary intervals.

- At the level of the data warehouse and analytical tools used by policy makers, it is very desirable to aggregate the data easily: using linguistic variables makes this feature easy to implement. For example, if we want to analyze the profile of the average customer in all our stores making a very good sales and customers are generally satisfied with the store, it would be interesting to directly express "good" and "satisfied" without having to set thresholds.

- Natural language is inherently understandable by non-technical people in the decision-making system, including policy makers and experts. Consider a case in point: if we add the knowledge of an expert system somewhere in the decision system, such as at the level of data integration which eliminates inconsistencies that occur when data sources are combined, they can be expressed in language natural, thus allowing the expert to transcribe independently his knowledge, which has various benefits (organizational, psychological, financial, etc.).

- Use natural language facilitates links with ontologies or semantic databases.

Data mining can also make very effective use of fuzzy logic. First, it can be used for clustering. The idea is that the membership of an individual to a cluster is not binary but fuzzy. Figure 5.13 shows a set of linearly distributed data. Figure 5.14 shows the binary membership function obtained by a classical clustering algorithm, Figure 5.15 shows the curve of a membership function of fuzzy clusters.

Figure 5.13: Linear data set to classify
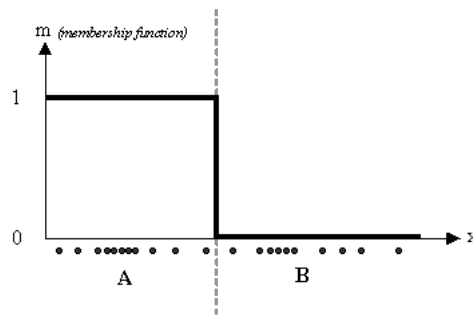
Figure 5.14: Binary clustering
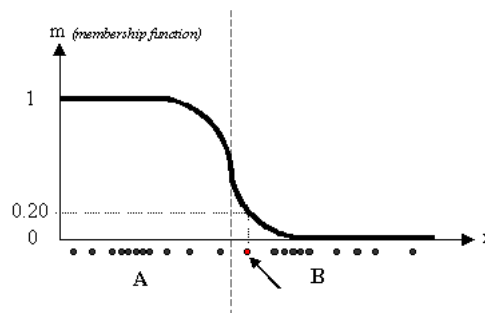
Figure 5.15: Fuzzy clustering with 2 clusters
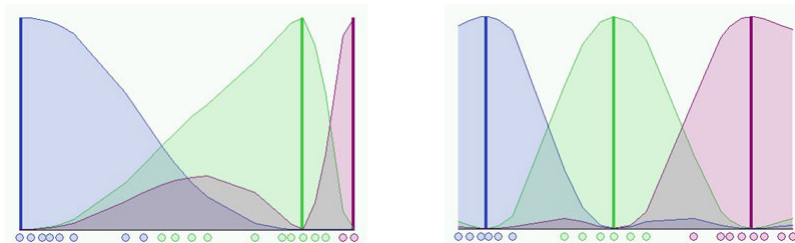
Figure 5.16: Fuzzy clustering with 3 clusters

Also, the use of neuro-fuzzy systems that we saw in Chapter 4.3.4 can show policy-makers how to unveil certain functions of their business as well as hidden patterns in the data. Like our rating experiment on video games, we could very well imagine that a video game publisher examines customer expectations in precisely the same way as we did in our experiment: finding the criteria for customer appreciation of a game, for example whether for a game that looks untechnically great but is technically poor, a significant investment to increase the technical aspect would be useful

or not. According to our test subjects, the answer is that such investment is likely to be unnecessary.

Thus we see that the applications of fuzzy logic for business intelligence and decision support systems are numerous and potentially substantial. Some such as clustering have long been used, others are still in a fledgling state, such as neuro-fuzzy systems.

Chapter 6

Conclusion

In the fields of observation, chance favors only the prepared mind.
Louis Pasteur.

We saw in the first part that fuzzy logic stems from the need to formalize inaccuracies. Despite its simple rules, it is able to mathematically model inference systems much more complex than can classical logic and linear models. Its decision matrix that brings together the fuzzy rules is similar in both form (structure of type *If X and Y then Z*) and content (the concept of linguistic variables) to the type of rules orally expressed by humans.

Besides, fuzzy logic can explain many experiments that had undermined traditional models of human reasoning in the 20th century. We showed how the non-additivity of probability judgments can be expressed in a fuzzy system. We then confronted fuzzy logic with some paradoxes of classical logic when it tries to model human reasoning: the sorites paradox is typically the kind of threshold problem that fuzzy logic reduces and the paradox of entailment does not pose a problem in fuzzy logic. It would be interesting to further explore Hempel's paradox and especially how we could express it in a neuro-fuzzy system. Similarly, Wason selection task would require further analysis, this time by focusing on fuzzy modus ponens and modus tollens. Beyond these paradoxes and non-additivity, fuzzy logic, based on the concept of linguistic variables, is essentially similar to natural language.

Thus fuzzy logic appears as a powerful theoretical framework for studying human reasoning. Surprisingly, we found only one study comparing the decisions made by humans with that of a fuzzy system, whose purpose was essentially to design a system of decision support for medical personnel, not analyze human reasoning as such. We conducted our own experiment and investigated whether a fuzzy system could mimic the results observed in humans. For this purpose, we used a technique for optimizing

fuzzy system using neural networks (neuro-fuzzy), through which we obtained good results, although the correlation between the two criteria for entry is high: a fuzzy system gives results closer to experimental values than those obtained by a polynomial system. This result reinforces the hypothesis that fuzzy logic can be used to explain decisions from human reasoning.

However, we emphasize on the fact on these experiments show at best that fuzzy systems can **imitate** human reasoning: only neuroscience could show that this way of reasoning really exists in the human brain. On this subject, some research articles of neuroscience such as [A., 2005] or [Feng and Capranica, 1978] in 2008 commented by David Olmsted on the website http://www.neurocomputing.org suggest that it is plausible that the concepts of fuzzy logic control means are "biologically compatible".
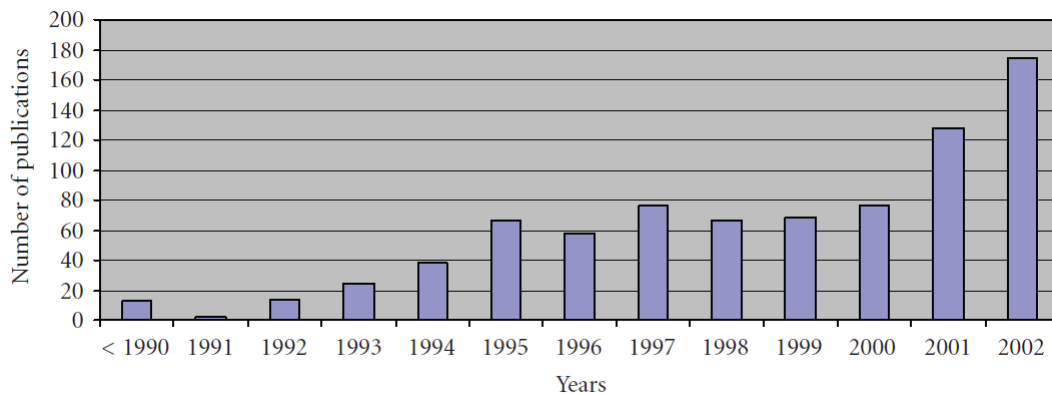
Figure 6.1: Number of publications per year indexed in MEDLINE using fuzzy logic. Source: [Torres A., 2005]

Source : [Torres A., 2005]

This proximity to human reasoning has generated strong enthusiasm of the scientific community, which is reflected by the number of publications, as shown in Figure . One natural application is decision systems: at the heart of the company, bringing together all the interesting information from the operational databases, decision systems could benefit greatly from the fuzzy logic by giving the keys of human reasoning, allowing to refine the decision-making.

Database theorists know what fuzzy logic could bring them in terms of information modeling: queries more intuitive and more powerful on the one hand, the data more consistent with the reality on the other. Many papers have been written, but few significant achievements have followed. The lack of consensus on a standard is probably the main reason behind.

The designers of decision support systems are primarily interested in the development of databases where the emphasis is on quality of information, as opposed to designers of operational databases that have other priorities such as the availability and speed. The significant improvement in the modeling of information permitted by the fuzzy logic should interest them greatly.

In addition, users of decision support systems are the decision makers of the company, who are primarily looking for tools that are easy and powerful way, which fuzzy mechanisms can provide. Thanks to their position, decision-makers can have a significant influence on research themes of decision-making system designers, if the latters manage to convince them of the usefulness of a technology.

Therefore, the designers of decision support systems could play a key role on the distribution of fuzzy systems in the heart of databases: such a role would mainly begin by pedagogically explaining the contribution of such systems to decision makers and to build partnerships with fuzzy database theorists. The latters would then have the opportunity to implement at a large scale some of their theories and start thinking about standards, in order to end the current fragmentation of experimental implementations.

Bibliography

[Atw, 1985] (1985). *An Object-Oriented DBMS for Design Support Applications*, volume 85. [cited at p. 44]

[A., 2005] A., P. (2005). The fuzzy logic of visuomotor control. *Canadian journal of physiology and pharmacology*, 74:456–462. [cited at p. 60]

[Andreasen et al., 1997] Andreasen, T., Christiansen, H., and Larsen, H. (1997). *Flexible query answering systems*. Kluwer Academic Publishers Norwell, MA, USA. [cited at p. 54]

[Boven and Epley, 2003] Boven, L. V. and Epley, N. (2003). The unpacking effect in evaluative judgments: When the whole is less than the sum of its parts. *Journal of Experimental Social Psychology*, 39(3):263 – 269. [cited at p. 19]

[Buckles and Petry, 1982] Buckles, B. P. and Petry, F. E. (1982). A fuzzy representation of data for relational databases. *Fuzzy Sets and Systems*, 7(3):213 – 226. [cited at p. 50]

[Caroff, 2010] Caroff, J. (2010). *Comparaison de la qualité d'image et des doses d'irradiation délivrées en TDM cardiaque avec synchronisation à l'ECG en mode prospectif vs rétrospectif : utilisation de paramètres ajustés à l'IMC des patients*. PhD thesis, Université d'Angers. [cited at p. 34]

[Codd, 1970] Codd, E. F. (1970). A relational model of data for large shared data banks. *Commun. ACM*, 13:377–387. [cited at p. 44]

[Cohen et al., 1956] Cohen, J., Dearnaley, E. J., and Hansel, C. E. M. (1956). The addition of subjective probabilities:the summation of estimates of success and failure. *Acta Psychologica*, 12:371 – 380. [cited at p. 19]

[Dembczynski et al., 2004] Dembczynski, K., Hapke, M., and Przybyl, D. (2004). Sqlfj-a fuzzy querying language and application. Technical report, Research Report RA-010/04, Poznan University of Technology. [cited at p. 54]

[Feng and Capranica, 1978] Feng, A. and Capranica, R. (1978). Sound localization in anurans. ii. binaural interaction in superior olivary nucleus of the green tree frog (hyla cinerea). *J. Comp. Physiol.*, 41:43-5. [cited at p. 60]

[Galindo et al., 1999] Galindo, J., Medina, J. M., and Aranda, M. C. (1999). Querying fuzzy relational databases through fuzzy domain calculus. *International Journal of Intelligent Systems*, 14(4):375–411. [cited at p. 50]

[Gardarin, 2005] Gardarin, G. (2005). *Bases de données*. Eyrolles, New York, NY, USA. [cited at p. 43]

[Hamdi Melih Saraoglu, 2007] Hamdi Melih Saraoglu, S. S. (2007). A fuzzy logic-based decision support system on anesthetic depth control for helping anesthetists in surgeries. *Journal of Medical Systems*, 31:511519. [cited at p. 29]

[Inmon, 1992] Inmon, W. H. (1992). *Building the Data Warehouse*. John Wiley & Sons, Inc., New York, NY, USA. [cited at p. 3]

[Jangi, 1992] Jangi, R. (1992). *Neuro-Fuzzy modeling: Architecture, Analysis and Application*. PhD thesis, University of California, Berkeley. [cited at p. 24, 37]

[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. 12]

[Macchi, 1999] Macchi, L. (1999). A note on superaddive probability judgment. *Psychological Review*, 106. [cited at p. 19, 20, 22]

[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. 14]

[Medina et al., 1994] Medina, J. M., Pons, O., and Vila, M. A. (1994). Gefred: a generalized model of fuzzy relational databases. *Inf. Sci. Inf. Comput. Sci.*, 76:87–109. [cited at p. 51]

[Mélès, 1971] Mélès, J. (1971). *La gestion par les systèmes*. Editions Hommes et Technique. [cited at p. 2]

[Oliva, 2003] Oliva, R. (2003). Visual FSQL: Gestión Visual de Bases de Datos Difusas en ORACLE a través de Internet usando FSQL. *Proyecto Fin de Carrera, dirigido por J. Galindo, de Ingeniería Superior en Informática de la Universidad de Málaga*. [cited at p. 53]

[P. Bosc, 1995] P. Bosc, O. P. (1995). Sqlf : A relational database language for fuzzy quering. *IEEE transactions on Fuzzy Systems*, 3:1–17. [cited at p. 51, 54]

[P. Bosc, 1998] P. Bosc, L. Liétard, O. P. (1998). Bases de données et flexibilité : Les requêtes graduelles. *Techniques et Sciences informatiques*, 17:355–378. [cited at p. 51]

[P. C. Wason, 1966] P. C. Wason, D. S. (1966). Natural and contrived experience in a reasoning problem. *New horizons in Psychology*. [cited at p. 24]

[Rasmussen and Yager, 1996] Rasmussen, D. and Yager, R. R. (1996). Summarysql - a flexible fuzzy query language. volume 17, page 118, Denmark. [cited at p. 54]

[Redelmeier DA, 1995] Redelmeier DA, Koehler DJ, L. V. T. A. (1995). Probability judgement in medicine: discounting unspecified possibilities. *Med Decis Making*, Jul-Sep;15(3):227–30. [cited at p. 18]

[Ribeiro and Moreira, 2003] Ribeiro, R. A. and Moreira, A. M. (2003). Fuzzy query interface for a business database. *Int. J. Hum.-Comput. Stud.*, 58:363–391. [cited at p. 26]

[S. Skrbic, 2008] S. Skrbic, A. T. (2008). On development of fuzzy relational database applications. pages 268–273. [cited at p. 50]

[Shafer G., 2005] Shafer G., V. V. (2005). The origins and legacy of kolmogorov's grundbegriffe. *The Game-Theoretic Probability and Finance Project*, Working Paper 4. [cited at p. 18]

[Torres A., 2005] Torres A., N. J. J. (2005). Fuzzy logic in medicine and bioinformatics. *Journal of Biomedicine and Biotechnology*, 2005:1–7. [cited at p. iv, 60]

[Tversky and Koehler, 1994] Tversky, A. and Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, 101(4):547 – 567. [cited at p. 19]

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

[Zadeh, 1975a] Zadeh, L. A. (1975a). The concept of a linguistic variable and its applications to approximate reasoning i. *Information Sciences*, 8:199–249. [cited at p. 25]

[Zadeh, 1975b] Zadeh, L. A. (1975b). The concept of a linguistic variable and its applications to approximate reasoning ii. *Information Sciences*, 8:301–357. [cited at p. 25]

[Zadeh, 1975c] Zadeh, L. A. (1975c). The concept of a linguistic variable and its applications to approximate reasoning iii. *Information Sciences*, 9:43–80. [cited at p. 25]

[Zadeh, 1978] Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3 – 28. [cited at p. 25]

[Zadrozny and Kacprzyk, 1996] Zadrozny, S. and Kacprzyk, J. (1996). Fquery for access: towards human consistent querying user interface. In *Proceedings of the 1996 ACM symposium on Applied Computing*, SAC '96, pages 532–536, New York, NY, USA. ACM. [cited at p. 54]

[Zhaohao Sun and Sun., 2005] Zhaohao Sun, G. F. and Sun., J. (2005). Four new fuzzy inference rules for experience based reasoning. *IFSA World Congress.* [cited at p. 24]

[Zvieli and Chen, 1986] Zvieli, A. and Chen, P. P. (1986). Entity-relationship modeling and fuzzy databases. In *Proceedings of the Second International Conference on Data Engineering*, pages 320–327, Washington, DC, USA. IEEE Computer Society. [cited at p. 50]

Appendices

Appendix A

Experiment with the anaesthesiologists

Here is the table of correlation calculated with SPSS:

Correlations

		VAR00013	VAR00014
VAR00013	Pearson Correlation	1	,627
	Sig. (2-tailed)		,000
	N	28	28
VAR00014	Pearson Correlation	,627**	1
	Sig. (2-tailed)	,000	
	N	28	28

**. Correlation is significant at the 0.01 level (2-tailed).

Figure A.1: Correlation matrix between the decisions of the anesthesiologist and those of the fuzzy system

The script to compute the correlation matrix is:

```
------------------ SCRIPT SPSS ------------------
DATASET ACTIVATE DataSet0.
CORRELATIONS
/VARIABLES=VAR00013 VAR00014
/PRINT=BOTH TWOTAIL NOSIG
/MISSING=PAIRWISE.
-------------------------------------------------
```

Appendix B

Game scoring experiment

Here are the details of the game scoring experiment. Some results as well as analysis scripts will be placed here as well.

During the first session, subjects are asked to score thirty games of their choice on a scale of 0 to 100, a high rating meaning that the player enjoys the game, on the website Wiki4Games (http://www.wiki4games.com). This website is based on the *MediaWiki* engine, and is used by all projects of the Wikimedia Foundation, including encyclopedic site *Wikipédia*. To record scores, they will use the extension *W4G Rating Bar*, which is a bar allowing each registered user on the site to give a score. Wiki4Games site is administered by the author of this document, and the extension *W4G Rating Bar* was developed by himself for the motor MediaWiki (http://www.mediawiki.org/wiki/Extension:W4G_Rating_Bar). This will allow total control of the scoring process of the experiment.

During the second session, they are asked to note two criteria comprising all aspects of a video game on an Excel spreadsheet. To analyze the results, we used MATLAB and SPSS.

In particular, we used two MATLAB tools:

- The *Surface Fitting Tool* allows us to generate a surface from 3D points of the experimental data.

- The *Fuzzy Loggic Toolbox* (http://www.mathworks.com/products/fuzzylogic/) allows us to study the results we get from an inference system based on fuzzy logic.

To establish the charts in the scores of the subjects, we store the stored the data in a .mat file and we will use the following script: *notesDavid.mat*:

```
------------------ SCRIPT MATLAB ------------------
% load data
load('notesDavid.mat', 'data')
selectedNotes = data;
X = selectedNotes(:, 1); % technical aspects
Y = selectedNotes(:, 2); % non technical aspects
Z = selectedNotes(:, 3); % global note

% draw figure
figure
scatter3(X, Y, Z, 5, 'filled')

% call the Surface Fitting Tool
sftool(X, Y, Z)
-------------------------------------------------------
```

The following pages show the results for subjects 1 and 2.
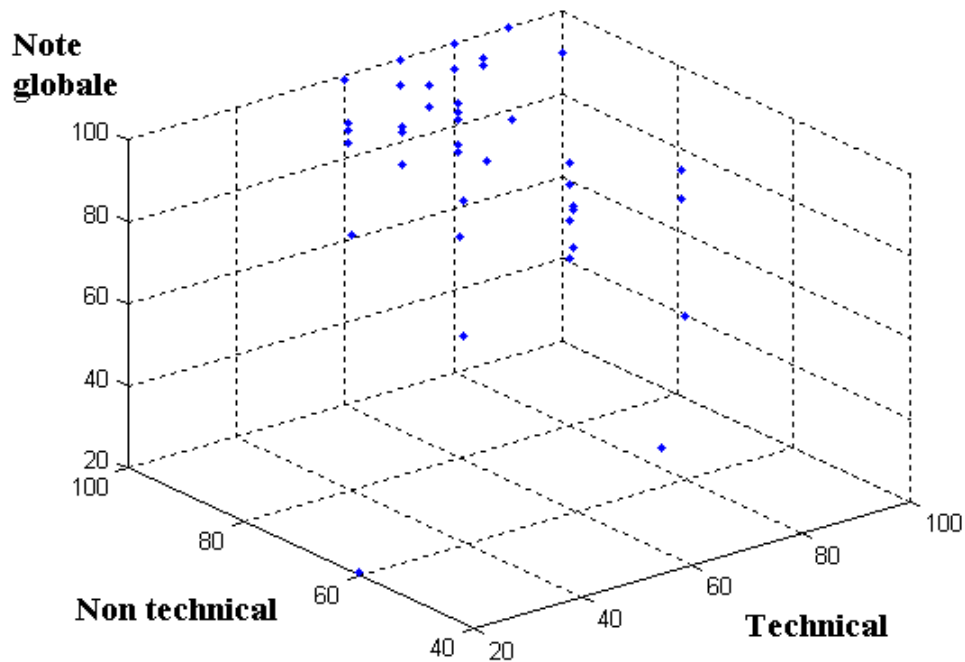
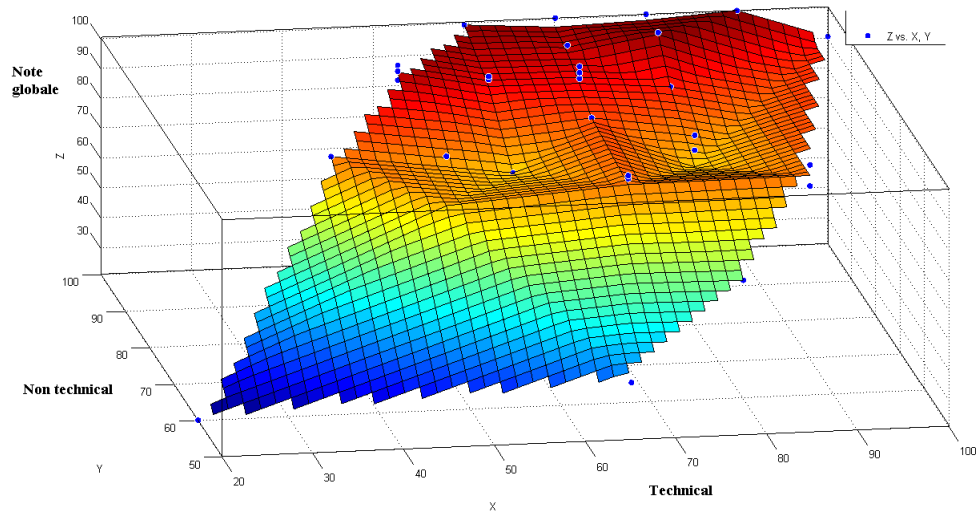Figure B.1: Overall scores of the subject 1 for the 2 criteria

Figure B.2: Surface generated from the overall scores of the subject 1 for the 2 criteria

The RMSE were calculated in MATLAB using the function errperf available here: http://www.mathworks.com/matlabcentral/fileexchange/15130-error-related-performance-metrics.
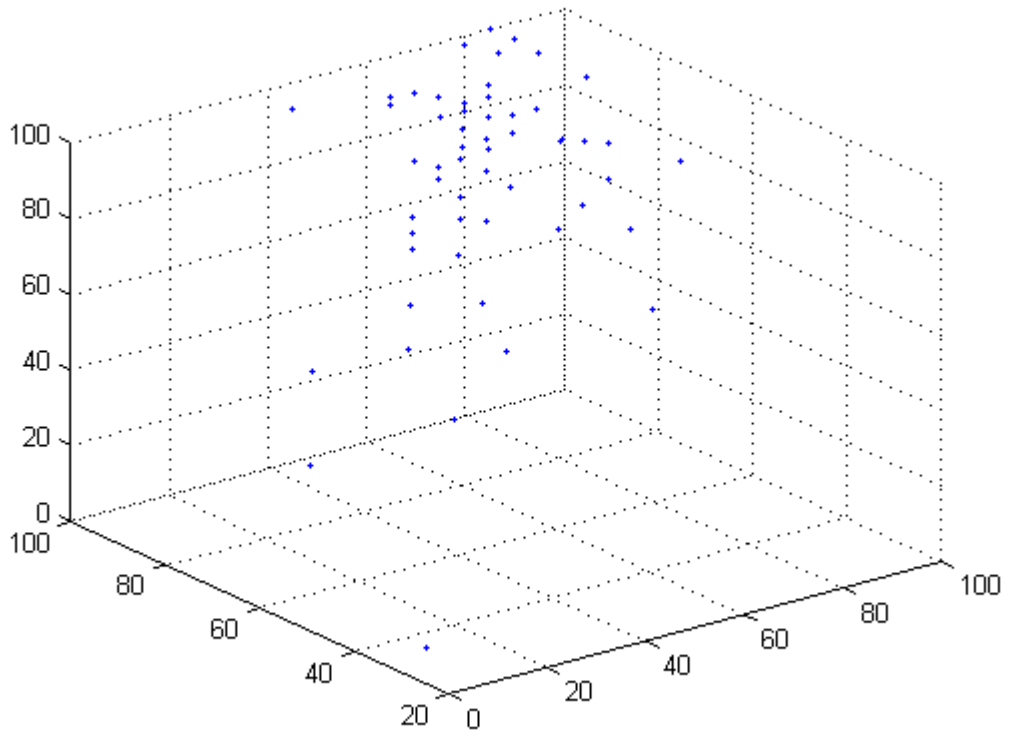
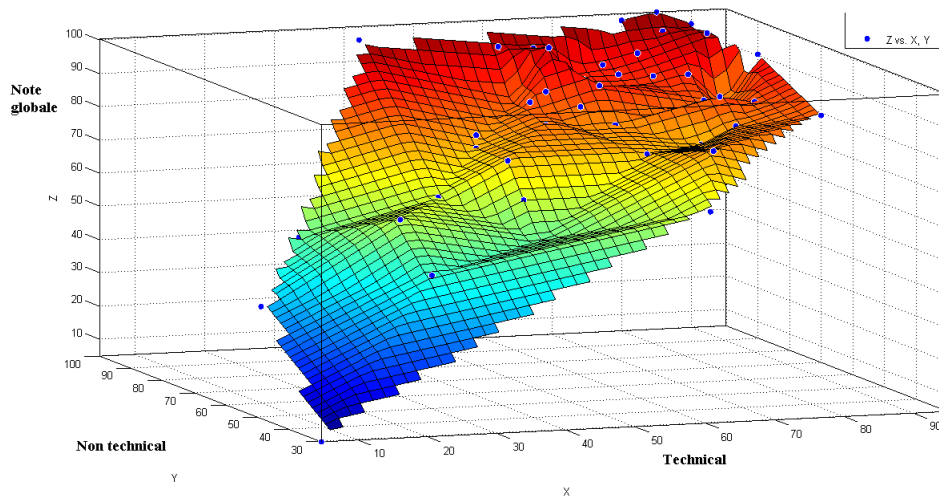Figure B.3: Overall scores of subject 2 for the 2 criteria

Figure B.4: Surface generated from overall scores of subject 2 for the 2 criteria