2

Fundamentals

In this chapter we briefly review the fuzzy logic theory in order to focus the type of fuzzy-rule based systems with which we intend to compute intelligible models. Although all the concepts will be expressed in a formal manner, they will also be clarified with examples in order to facilitate their understanding. Moreover we detail some aspects to be considered in order to assure intelligible models with a satisfactory accuracy.

Most topics have been obtained from the original work of L.A. Zadeh [127] and the summaries written by A. Riid [98], J.M. Mendel [78] and G.J. Klir and T.A. Folger [59]. Some aspects about propositional logic have also been taken from the work of C.B. Allendoerfer and C.O. Oakley [2].

2.1 Fuzzy logic theory

Fuzzy logic was first introduced in 1965 by Lotfi A. Zadeh [125] with the concept of fuzzy sets as an extension of the classical set theory formed by crisp sets. Later he defined a whole algebra, fuzzy logic [127], which uses fuzzy sets to compute with words as an extension of the proper operations of classical logic.

In most cases a fuzzy logic system is, in fact, a nonlinear mapping of an input data vector into a scalar output where this relation is defined by linguistic expressions which are obviously computed with numbers. Thus a fuzzy logic system is unique in that it is able to handle numerical data and linguistic knowledge. The richness of this logic is that there are many possibilities which lead to many different mappings.

In this section we will provide a summary of those necessary parts of the fuzzy logic literature to understand how a fuzzy logic system works.

2.1.1 Universe of scope

The universe of scope, also called the universe of discourse and typically described as U, is the domain of each variable, either input or output. Thus this is the set of allowable values for the variable.

The universe of scope can be a continuous domain or a discrete set of points. And it can also have specific units or not.

Example

For example if we work with the variable *temperature of human body* the universe of scope could be

$$U = \{u \in \mathbb{R} \mid 35 \text{ Celsius degrees} \leq u \leq 42 \text{ Celsius degrees}\}$$

but if we work with the variable *injured players in a football match* the universe of scope could be

$$U = \{u \in \mathbb{N} \mid 0 \leq u \leq 22\}$$

2.1.2 Fuzzy sets versus crisp sets

In classical logic an element u is either a member or non-member of a crisp set X, subset of the universe of scope U. It is typically defined with zero-one membership functions denoted $\mu_X(u)$ such as

$$X \Rightarrow \mu_X(u) = \begin{cases} 1 \text{ if } u \in X \\ 0 \text{ if } u \notin X \end{cases} \tag{2.1}$$

In fuzzy logic theory an element belongs to a fuzzy set X of the universe of scope U with a degree of membership $\mu_X(u) \in [0,1]$ whose value is proportional to the relevance of the element into the set X. They provide a measure of the degree of similarity of the element to the fuzzy subset. The relations between the values of the universe of scope and the degrees of membership are characterized with membership functions which are represented as a set of ordered pairs of each element u and its degree of membership $\mu_X(u)$ such as

$$X = \{(u, \mu_X(u)) \mid u \in U\} \tag{2.2}$$

When U is continuous then X is commonly written as

$$X = \int_U \mu_X(u)/u \tag{2.3}$$

where the integral sign does not denote integration and the slash does not denote a division. It is just a representation of the collection of all points $u \in U$ with associated degrees of membership $\mu_X(u)$.

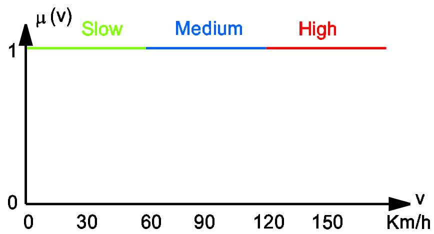
On the other hand, when U is discrete then X is commonly written as

$$X = \sum_{U} \mu_X(u)/u \tag{2.4}$$

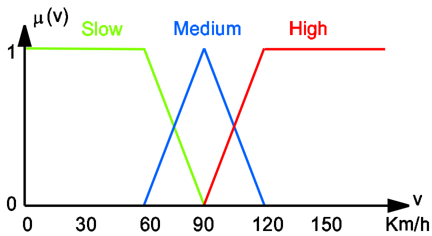
where neither the sum sign nor the slash denote these operations again.

Example

Consider the representation of the speed in a motorway either in classical logic or in fuzzy logic theory as[1]:



Classical theory



Fuzzy logic theory

In this example we have three sets with classical logic theory:

$$\mu_{slow}(v) = \begin{cases} 1 & \text{if } v \leq 60 \\ 0 & \text{otherwise} \end{cases}$$
$$\mu_{medium}(v) = \begin{cases} 1 & \text{if } 60 < v < 120 \\ 0 & \text{otherwise} \end{cases}$$
$$\mu_{high}(v) = \begin{cases} 1 & \text{if } v \geq 120 \\ 0 & \text{otherwise} \end{cases}$$

and also three sets with fuzzy logic theory:

$$\mu_{slow}(v) = \begin{cases} 1 & \text{if } v \leq 60 \\ \frac{90-v}{30} & \text{if } 60 < v \leq 90 \\ 0 & \text{if } v > 90 \end{cases}$$
$$\mu_{medium}(v) = \begin{cases} 0 & \text{if } v \leq 60 \\ \frac{v-60}{30} & \text{if } 60 < v \leq 90 \\ \frac{120-v}{30} & \text{if } 90 < v \leq 120 \\ 0 & \text{if } v > 120 \end{cases}$$
$$\mu_{high}(v) = \begin{cases} 0 & \text{if } v \leq 90 \\ \frac{v-90}{30} & \text{if } 90 < v \leq 120 \\ 1 & \text{if } v > 120 \end{cases}$$

Thus, while in classical logic each element of the universe of scope V, for example $v = 100Km/h$, only belongs to one set, $\mu_{slow}(100) = \mu_{high}(100) = 0$ and $\mu_{medium}(100) = 1$; in fuzzy logic theory this element belongs to more than one set, $\mu_{medium}(100) = 0.66$ and $\mu_{high}(100) = 0.33$.

---

[1] The number of linguistic labels has been arbitrarily chosen. Obviously, it could be less or more.

2.1.3 Properties of fuzzy sets

Recall that fuzzy sets can be defined over a continuous universe of scope or a discrete one. There are some typical functions used to define a fuzzy set: gaussian membership functions, triangular membership functions or trapezoidal membership functions. All these functions are *convex* functions because $\forall u_1, u_2, u_3 \in U \mid u_1 \leq u_2 \leq u_3 \rightarrow \mu_X(u_2) \geq min\,(\mu_X(u_1), \mu_X(u_3))$. Furthermore some properties can be defined about them, plotted in figure 2.1:

- The *height* of a fuzzy set is defined by

  $$height(X) = sup_{u \in U} \mu_A(u)$$

  The fuzzy sets with a height equal to 1 are called *normal* sets.

- The *core* of a fuzzy set is a subset of X defined by

  $$core(X) = \{u \in U \mid \mu_X(u) = 1\}$$

  Normal, convex and piecewise continuous fuzzy sets with only one value in their core are called *fuzzy numbers*.

- The *support* of a fuzzy set is a subset of X defined by

  $$support(X) = \{u \in U \mid \mu_X(u) > 0\}$$

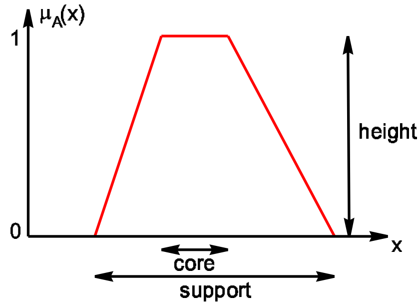


Figure 2.1: Core, support and height of a fuzzy set.

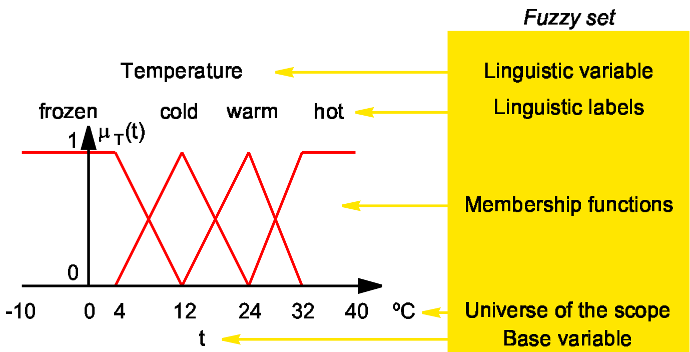
A special case of fuzzy sets must be considered, the fuzzy numbers whose core is equal to its support. These are fuzzy sets with only one value of the universe of scope whose degree of membership is equal to one while the rest of values have a degree of membership equal to zero. This type of fuzzy sets are called *singletons*.

2.1.4 Fuzzy partitions

All models are based on variables. When these models are based on fuzzy logic, for each variable several fuzzy sets are usually defined because a linguistic variable may be usually decomposed into a set of terms which cover its universe of scope. Thus, in most cases either the variables or the fuzzy sets have a linguistic meaning.

Example

Suppose that we must define the variable *temperature*. In this case we could consider the following membership functions:



Observe how given the lingustic variable *temperature* we can define several linguistic labels which are related each one to a fuzzy set.

Like in the last example, it is common that each input value has nonzero degree of membership for at least one fuzzy set. A partition satisfying this premise has *coverage* property. In fact, if S is the number of sets which make up the partition, whose membership functions are defined with $\mu_{X_s}(u)|_{s=1\ldots S}$, usually these partitions assure that

$$\sum_{s=1}^{S} \mu_{X_s}(u) = 1 \quad \forall\ u \in U \tag{2.5}$$

$$\sum_{s=1}^{S} \mu_{X_s}(u) > 0 \quad \forall\ u \in U \tag{2.6}$$

and are called *fuzzy partitions*[2]. In this case, for each input value the sum of the degrees of membership is equal to one and it can belong to two fuzzy sets at most. The element of the universe of scope with a degree of membership equal to 0.5 is called the *crossover point*.

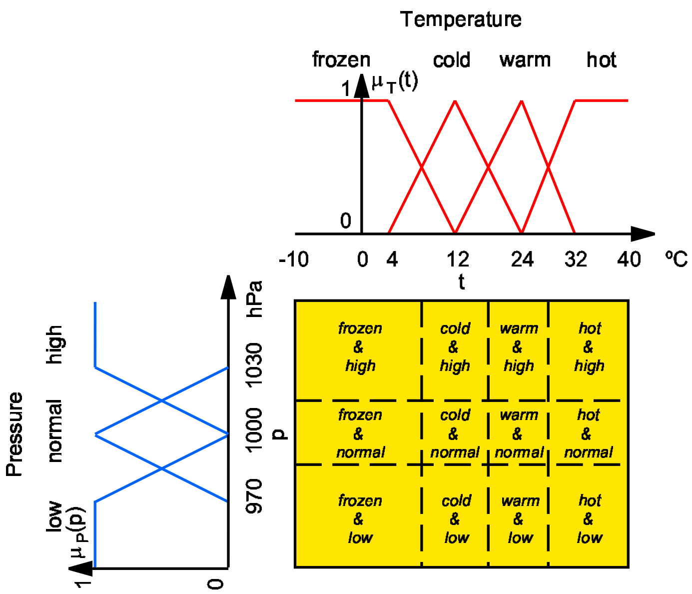
[2]In fact its original name is Ruspini partitions, proposed by E.H. Ruspini.

The number of fuzzy sets and therefore the number of membership functions can differ between different people. Greater resolution is achieved by using more membership functions at the price of greater computational complexity. In fact this will be a very important aspect we will study in this work.

If we have more than one variable when defining a system, we will probably have a fuzzy partition for each one of them. In this case we will split the overall combination of universes of scope in several cases. For each one we will probably define different situations and in this way, fuzzy modeling can be understood as a piecewise modeling where we split the input scope in several linguistic situations.

Example

Observe the following model where we have $4 \times 3 = 12$ different situations:



By working with fuzzy partitions we will have more than one *active* situation (whose both degrees of membership are greater than zero) every time. In this example we will have $2 \times 2 = 4$ *active* situations for each possible input pair.

Thus the advantage of fuzzy sets over crisp ones becomes clearer. Membership functions allow the description of concepts in which the boundary between having a property and not having a property is not sharp. We can describe different situations to be considered more or less (not yes or no) based on the current state. Moreover, by using fuzzy sets and their linguistic labels, we are able to move from numbers to abstractions which is natural for human beings but is otherwise difficult to formulate mathematically.

2.1.5 Basic operations with fuzzy sets

Theoretic operations from classical logic such as the intersection, the union and the complement are extended to fuzzy logic so as to do analogous things with fuzzy sets. Anyway these extensions are not uniquely defined as in classical logic.

To begin, let us briefly review the elementary operations with crisp sets. Let X and Y be two subsets of U. The union denoted $X \cup Y$ contains all of the elements in either X or Y. The intersection denoted $X \cap Y$ contains all of the elements which are simultaneously in X and Y. The complement denoted $\bar{X}$ contains all of the elements which are not in X. In fuzzy logic these operations denote the same concepts.

The general forms of intersection and union are represented by triangular norms (T-norms) and triangular conorms (S-norms), respectively.

A T-norm is a function from $[0,1] \times [0,1] \rightarrow [0,1]$ satisfying the following criteria:

$T(u,1) = u$		One identity
$T(u_1,u_2) \leq T(u_3,u_4)$	whenever $u_1 \leq u_3, u_2 \leq u_4$	Monotonicity
$T(u_1,u_2) = T(u_2,u_1)$		Commutativity
$T(T(u_1,u_2),u_3) = T(u_1,T(u_2,u_3))$		Associativity

The most common T-norms are the minimum and the product.

$$\mu_{X\cap Y}(u) = \min\left(\mu_X(u), \mu_Y(u)\right) \tag{2.7}$$

$$\mu_{X\cap Y}(u) = \mu_X(u)\mu_Y(u) \tag{2.8}$$

These operations are commonly considered when fuzzy sets are joined with the linguistic operator *AND* due to the fact of being a restrictive operator.

A S-norm is a function from $[0,1] \times [0,1] \rightarrow [0,1]$ satisfying the following criteria:

$S(u,0) = u$		Zero identity
$S(u_1,u_2) \leq S(u_3,u_4)$	whenever $u_1 \leq u_3, u_2 \leq u_4$	Monotonicity
$S(u_1,u_2) = S(u_2,u_1)$		Commutativity
$S(S(u_1,u_2),u_3) = S(u_1,S(u_2,u_3))$		Associativity

The most common S-norms are the maximum, the probabilistic sum and the bounded sum:

$$\mu_{X\cup Y}(u) = \max\left(\mu_X(u), \mu_Y(u)\right) \tag{2.9}$$

$$\mu_{X\cup Y}(u) = \mu_X(u) + \mu_Y(u) - \mu_X(u)\mu_Y(u) \tag{2.10}$$

$$\mu_{X\cup Y}(u) = \min\left(1, \mu_X(u) + \mu_Y(u)\right) \tag{2.11}$$

These operations are commonly considered when fuzzy sets are joined with the linguistic operator *OR* due to the fact of being a boundadous operator.

Figure 2.2 shows these basic operations with sets defined over the same universe of scope in order to define the resulting fuzzy set. Anyway T-norms and S-norms can be applied to fuzzy sets (in fact fuzzy relations) defined over different universes of scope as will be described later.

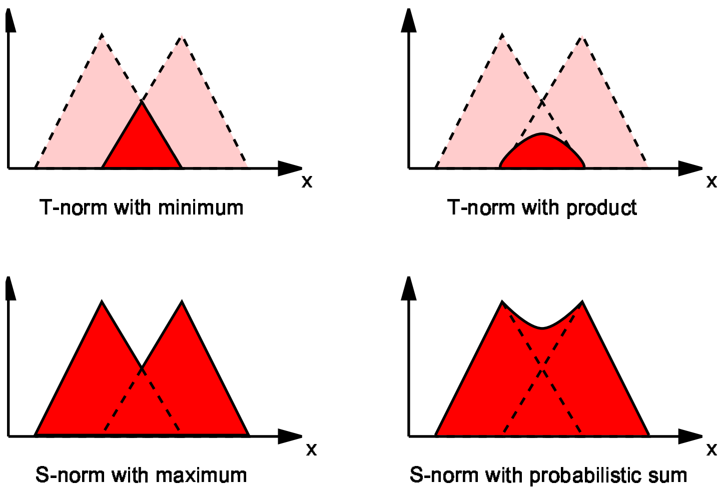


Figure 2.2: Basic operations with fuzzy sets.

Typically, the complement of a fuzzy set is defined as $1 - \mu_X(u)$. This operation is commonly used when the linguistic variable has the linguistic operator *NOT* before the variable itself like in figure 2.3.

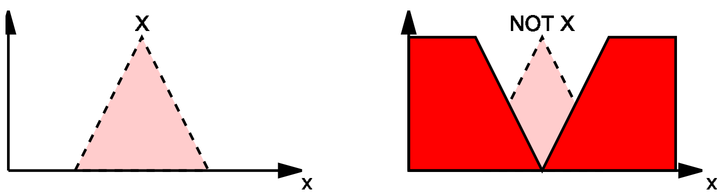


Figure 2.3: Complement.

2.1.6 Fuzzy relations and compositions

A relation can be defined [59] as the presence or absence of association, interaction or interconnectedness between the elements of two or more sets.

Like in the previous operations, the relations can be expressed in fuzzy logic in a similar way as they are expressed in classical logic.

In this sense recall that a binary crisp relation $R(U,V)$ is typically defined as

$$\mu_R(u,v) = \begin{cases} 1 & \text{if and only if } (u,v) \in R(U,V) \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

As an extension from the crisp relations, a fuzzy relation represents *a degree* of the presence or absence of association, interaction or interconnectedness between the elements of two or more fuzzy sets. Thus if U and V are two universes of scope then a binary fuzzy relation $R(U,V)$ is a fuzzy subset of $U \times V$ characterized by a membership function with degrees of membership $\mu_R(u,v) \in [0,1]$ where $u \in U$ and $v \in V$.

$$R(U,V) = \{((u,v), \mu_R(u,v)) \mid (u,v) \in U \times V\} \tag{2.13}$$

Example

Suppose that we want to define the concept *the vehicle x is close to the vehicle y* where both vehicles run in a road of length 100Km. Each position is defined by considering its distance from the origin. If we define this concept in classical logic we first need to decide a clear edge between a close distance and a far distance, for example 30Km. If in this case we consider discrete domains of each vehicle's position with a step of 20Km, the crisp relation could be expressed with a relational matrix such as

$$\begin{array}{c} \\ u_1=0 \\ u_2=20 \\ u_3=40 \\ u_4=60 \\ u_5=80 \\ u_6=100 \end{array} \begin{array}{c} \begin{array}{cccccc} v_1=0 & v_2=20 & v_3=40 & v_4=60 & v_5=80 & v_6=100 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

On the other hand if we used a fuzzy relation we could define this concept by computing each element of the relation with a membership function where $\mu_R(u,v) = 1 - \frac{|u-v|}{100}$.

$$\begin{array}{c} \\ u_1=0 \\ u_2=20 \\ u_3=40 \\ u_4=60 \\ u_5=80 \\ u_6=100 \end{array} \begin{array}{c} \begin{array}{cccccc} v_1=0 & v_2=20 & v_3=40 & v_4=60 & v_5=80 & v_6=100 \end{array} \\ \begin{pmatrix} 1.0 & 0.8 & 0.6 & 0.4 & 0.2 & 0.0 \\ 0.8 & 1.0 & 0.8 & 0.6 & 0.4 & 0.2 \\ 0.6 & 0.8 & 1.0 & 0.8 & 0.6 & 0.4 \\ 0.4 & 0.6 & 0.8 & 1.0 & 0.8 & 0.6 \\ 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 0.8 \\ 0.0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 \end{pmatrix} \end{array}$$

Observe the difference between a crisp relation and a fuzzy relation. A crisp relation $\mu_R(u,v)$ is equal to 0 or 1 while in a fuzzy relation $\mu_R(u,v)$ is a number between 0 and 1. The generalization of the crisp or fuzzy relations to an N-dimensional domain is straightforward.

As a fuzzy relation is in fact a fuzzy set on a product space, the basic operations of fuzzy sets can also be defined with fuzzy relations by using the same operators. Thus, if $R(u,v)$ and $S(u,v)$ are two fuzzy relations on the same product space $U \times V$ then their intersection and union are defined with their composition by using a T-norm and a S-norm, respectively, as

$$\mu_{R\cap S}(u,v) = T\left(\mu_R(u,v), \mu_S(u,v)\right) \tag{2.14}$$
$$\mu_{R\cup S}(u,v) = S\left(\mu_R(u,v), \mu_S(u,v)\right) \tag{2.15}$$

Example

Suppose now that we want to define the concept *the vehicle x is close to the vehicle y AND the vehicle x is closer to the end of the road than the vehicle y.* For the first part of the sentence we recover the fuzzy relation of the last example. The second part may be defined with the following fuzzy relation with the same universes of scope than before:

$$\begin{array}{c} \\ u_1=0 \\ u_2=20 \\ u_3=40 \\ u_4=60 \\ u_5=80 \\ u_6=100 \end{array}
\begin{array}{c} \begin{array}{cccccc} v_1=0 & v_2=20 & v_3=40 & v_4=60 & v_5=80 & v_6=100 \end{array} \\
\begin{pmatrix} 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0.0 \\ 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 \\ 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 \\ 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 \\ 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 \\ 1.0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 \end{pmatrix} \end{array}$$

As both part of the sentence are defined on the same product space $U \times V$ defined by the position of each vehicle on the road, then the whole sentence can be computed with a T-norm (i.e. the minimum) as

$$\begin{array}{c} \\ u_1=0 \\ u_2=20 \\ u_3=40 \\ u_4=60 \\ u_5=80 \\ u_6=100 \end{array}
\begin{array}{c} \begin{array}{cccccc} v_1=0 & v_2=20 & v_3=40 & v_4=60 & v_5=80 & v_6=100 \end{array} \\
\begin{pmatrix} 0.5 & 0.4 & 0.3 & 0.2 & 0.1 & 0.0 \\ 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 \\ 0.6 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 \\ 0.4 & 0.6 & 0.6 & 0.5 & 0.4 & 0.3 \\ 0.2 & 0.4 & 0.6 & 0.6 & 0.5 & 0.4 \\ 0.0 & 0.2 & 0.4 & 0.6 & 0.6 & 0.5 \end{pmatrix} \end{array}$$

In fact most fuzzy relations are build by combining different fuzzy sets. In order to make it possible, the fuzzy sets are first converted to fuzzy relations which are later operated with a norm. The first step is realized with the *cylindrical extension principle* proposed by L.A. Zadeh [127]. This principle is applied when we are interested in adding a new dimension (universe of scope) to a fuzzy relation with n domains to obtain a fuzzy relation with n+1 domains.

In this sense if we have a fuzzy set whose membership function is defined in the universe of scope U, i.e. $\mu_X : U \rightarrow [0,1]$ the *cylindrical extension* of X on $U \times V$ is a fuzzy relation defined by

$$\text{ce of } X \text{ on } U \times V = \{(u,v), \mu_X(u)\} \tag{2.16}$$

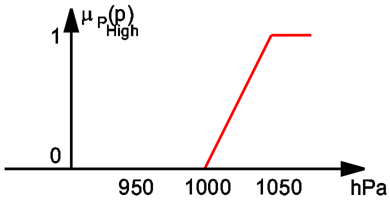
Thus, the degree of membership for each u is just copied to all (u,v) with the same u.

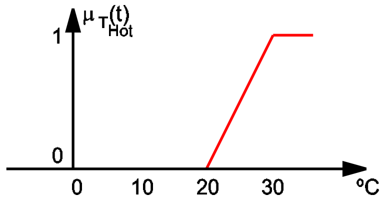
Its complementary operator is the *projection* of $U \times V$ on U which gives a fuzzy set defined by

$$\text{proj of } U \times V \text{ on } U = \left\{u, \max_v(\mu_R(u,v))\right\} \tag{2.17}$$

Example

Suppose that we have the two fuzzy sets called *the (atmospheric) pressure is high* and *the (atmospheric) temperature is hot* from which we want to build the fuzzy relation *the temperature is hot when the pressure is high.*





In order to simplify the example we will work with discrete domains. The pressure will be defined from 925hPa to 1075hPa with a step of 25 and the temperature will defined from 14 to 32 Celsius degress with a step of 2. We first use the *cylindrical extension* to create a fuzzy relation from the set *pressure is high* such as

$$\begin{array}{r c}
press\backslash temp & \begin{array}{cccccccccc} 14 & 16 & 18 & 20 & 22 & 24 & 26 & 28 & 30 & 32 \end{array} \\
\begin{array}{r} 925 \\ 950 \\ 975 \\ 1000 \\ 1025 \\ 1050 \\ 1075 \end{array} &
\left(\begin{array}{cccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\
1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\
1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0
\end{array}\right)
\end{array}$$

and also another *cylindrical extension* to create a fuzzy relation from the set *temperature is cold* such as

$$\begin{array}{r c}
press\backslash temp & \begin{array}{cccccccccc} 14 & 16 & 18 & 20 & 22 & 24 & 26 & 28 & 30 & 32 \end{array} \\
\begin{array}{r} 925 \\ 950 \\ 975 \\ 1000 \\ 1025 \\ 1050 \\ 1075 \end{array} &
\left(\begin{array}{cccccccccc}
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0
\end{array}\right)
\end{array}$$

in order to apply a T-norm (i.e. minimum) and to obtain the final fuzzy relation

$$\begin{array}{r} press\backslash temp \\ 925 \\ 950 \\ 975 \\ 1000 \\ 1025 \\ 1050 \\ 1075 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 14 & 16 & 18 & 20 & 22 & 24 & 26 & 28 & 30 & 32 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \\ 0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \end{pmatrix} \end{array}$$

Now we consider the composition of relations on different product spaces which share a common set, namely $P(U,V)$ and $Q(V,W)$ which is typically defined as

$$R(U,W) = P(U,V) \circ Q(V,W) \tag{2.18}$$

In classical logic $R(U,W)$ is defined as a subset of $U \times W$ such that $(u,w) \in R$ if and only if there exists at least one $v \in V$ such that $(u,v) \in P$ and $(v,w) \in Q$. In this case the composition of two binary crisp relations may be defined as

$$\mu_R(u,w) = \Big\{(u,w), \max_v \left[\min\left(\mu_P(u,v), \mu_Q(v,w)\right)\right]\Big\} \tag{2.19}$$

or

$$\mu_R(u,w) = \Big\{(u,w), \max_v \left[\mu_P(u,v)\mu_Q(v,w)\right]\Big\} \tag{2.20}$$

which are called *max-min composition* and *max-product composition*, respectively. These compositions are not the only ones which represent correctly $R(U,W)$ but they seem to be the most widely used ones.

In the case of fuzzy relations, the composition on different product spaces which share a common set is defined again analogously to the crisp composition but in this case each degree of membership may be any real number between 0 and 1. By using a T-norm and a S-norm we could define the fuzzy composition between two fuzzy relations $P(U,V)$ and $Q(V,W)$ as

$$R(U,W) = \{((u,w), \mu_R(u,w)) \mid (u,w) \in U \times W\} \tag{2.21}$$

where

$$\mu_R(u,w) = S_v\left(T\left(\mu_P(u,v), \mu_Q(v,w)\right)\right) \tag{2.22}$$

In fact these compositions are a simplified expression of two fuzzy relations on the same product space build with the *cylindrical extension* (ce) and then simplified with the *projection* (proj):

$$R(U,W) = proj(T(ce(P(U,V)\ on\ W), ce(Q(V,W)\ on\ U)))\ on\ R\times W \tag{2.23}$$

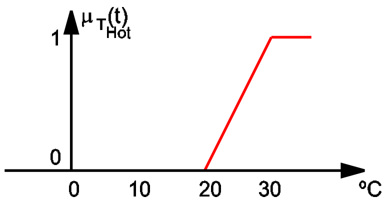
Recall that the projection is in fact a S-norm (typ. maximum). Anyway this expression is rarely used because it is not necessary to build the fuzzy relation on the same product space in order to obtain the final result correctly.

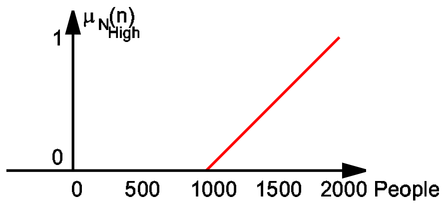
Although it is permissible to use many different T-norms and S-norms, most compositions of fuzzy relations are computed with the maximum as S-norm and either the minimum or the product as T-norm. In this case these compositions are called *max-min composition* or *max-product composition* like in the composition of crisp relations. Many people also call them *sup-star compositions* where the *star* operator is any T-norm.

$$\mu_R(u, w) = \sup_v \left[\mu_P(u, v) \star \mu_Q(v, w)\right] \tag{2.24}$$

Example

We could consider the fuzzy relation of the last example *the temperature is hot when the pressure is high* and also the fuzzy relation *the number of people in the beach is usually high when the temperature is hot* build from the following fuzzy sets, in order to compute the fuzzy relation *the number of people in the beach is usually high when the pressure is also high* with the max-min composition.





$$
\begin{array}{cc}
press\backslash temp & \begin{array}{cccccccccc} 14 & 16 & 18 & 20 & 22 & 24 & 26 & 28 & 30 & 32 \end{array} \\
\begin{array}{r} 925 \\ 950 \\ 975 \\ 1000 \\ 1025 \\ 1050 \\ 1075 \end{array} &
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.5 & 0.5 & 0.5 & 0.5 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 & 1.0
\end{pmatrix}
\end{array} \circ
$$

$$
\begin{array}{cc}
temp\backslash people & \begin{array}{ccccc} 0 & 500 & 1000 & 1500 & 2000 \end{array} \\
\begin{array}{r} 14 \\ 16 \\ 18 \\ 20 \\ 22 \\ 24 \\ 26 \\ 28 \\ 30 \\ 32 \end{array} &
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.2 & 0.2 \\
0 & 0 & 0 & 0.4 & 0.4 \\
0 & 0 & 0 & 0.5 & 0.6 \\
0 & 0 & 0 & 0.5 & 0.8 \\
0 & 0 & 0 & 0.5 & 1.0 \\
0 & 0 & 0 & 0.5 & 1.0
\end{pmatrix}
\end{array} =
$$

$$
\begin{array}{r c}
 & \begin{array}{ccccc} 0 & 500 & 1000 & 1500 & 2000 \end{array} \\
press \backslash people & \\
\begin{array}{r} 925 \\ 950 \\ 975 \\ 1000 \\ 1025 \\ 1050 \\ 1075 \end{array} &
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.5 & 0.5 \\
0 & 0 & 0 & 0.5 & 1.0 \\
0 & 0 & 0 & 0.5 & 1.0
\end{pmatrix}
\end{array}
$$

A special case appears when we must compute the composition between a fuzzy set and a fuzzy relation (not two fuzzy relations). This operation will be called a *fuzzy implication*. As we will point out later, this is probably the most significant operation in most fuzzy logic systems due to the fact of being similar to the composition between an input (the fuzzy set) and a system (the fuzzy relation) in order to obtain the output of the system. See figure 2.4.

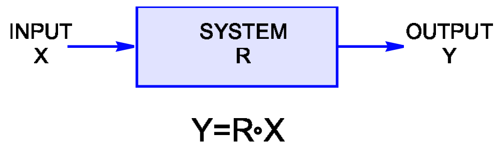


Figure 2.4: The fuzzy implication.

Let X be a fuzzy set of the universe of scope U with a degree of membership $\mu_X(u)$ and R be a fuzzy relation of $U \times V$ defined with the membership function $\mu_R(u, v)$. The result of the *fuzzy implication* of X and R is a fuzzy set $Y = X \circ R$ in the universe of scope V computed with a *sup-star composition*, typically a *max-min composition* or *max-product composition*. Thus, the membership function for Y would be

$$Y(v) = \left\{ \left(v, \sup_u \left[\mu_X(u) \star \mu_R(u, v) \right] \right) \mid v \in V \right\} \tag{2.25}$$

The fuzzy implication is in fact just the name of the relation R, typically denoted by $\mu_{X \to Y}(u, v)$, but nowadays its name is used for either the relation R or the composition $X \circ R$. By being the case with which we will study the models in this work, we will analyze it later in detail.

2.1.7 Fuzzy logic versus propositional logic

Most systems based on fuzzy logic are expressed (and computed) with *if ... then ...* statements, i.e. *if u is X then v is Y* where $u \in U$ and $v \in V$.

In order to operate these rules we need to compute an *implication* which is usually denoted by $\mu_{X \to Y}(u, v)$.

This operation resides within a branch of mathematics known as logic and so far we have been discussing a set theory. Fortunately, as stated in [59] *it is well established that propositional logic is isomorphic to set theory under the appropiate correspondence between components of these two mathematical systems. Furthermore, both of these systems are isomorphic to a Boolean algebra ... The isomorphisms between Boolean algebra, set theory and propositional logic guarantee that every theorem in any one of these theories has a counterpart in each one of the other two theories.*

Consequently, by observing the previous extension from crisp sets, which are the core of Boolean algebra, to fuzzy sets, we might suppose that we will not have many problems to find the extension from the implication between crisp sets to the implication between fuzzy sets. Unfortunately this extension will no be so clear and it has been so far the most criticized aspect about fuzzy logic. The fact is that propositional logic combines unrelated propositions into an implication *without assuming any cause or effect relation to exist* while most fuzzy systems need to express related propostitions between inputs and outputs.

Anyway we will show how there is a *reasonable* manner to compute the implication with fuzzy sets which has been commonly accepted by giving a satisfactory relation between the fuzzy logic and the human reasoning. First, we will give a short review about propositional logic in order to explain the necessary differences with fuzzy logic.

In propositional logic the rules are just a form of proposition and they must be *true* or *false*. In fact any proposition may be combined with other propositions in a process which is called a *logical reasoning*.

There are basically five fundamental combinations or operations which are the core of logical reasoning: the *conjunction* denoted $X \wedge Y$, the *disjunction* denoted $X \vee Y$, the *implication* denoted $X \to Y$, the *equivalence* denoted $X \leftrightarrow Y$ and the *negation* denoted $\neg X$. Their truth tables are given in table 2.1.

X	**Y**	$X \wedge Y$	$X \vee Y$	$X \to Y$	$X \leftrightarrow Y$	$\neg X$
True	True	True	True	True	True	False
True	False	False	True	False	False	False
False	True	False	True	True	False	True
False	False	False	False	True	True	True

Table 2.1: Truth table of the elementary operations in logical reasoning.

A *tautology* is a proposition formed by combining other propositions. The most important tautologies for our work are those related with the implication such as

$$X \to Y \quad \leftrightarrow \quad \neg\,[X \wedge (\neg Y)] \tag{2.26}$$

$$X \to Y \quad \leftrightarrow \quad (\neg X) \vee Y \tag{2.27}$$

They let us express the membership function for $X \to Y$ in terms of membership functions of either propositions X, Y, $\neg X$ and $\neg Y$ which is the main objective of this section.

These tautologies can be easily demonstrated for example in the case of Boolean algebra. Recall first that the mathematical equivalences between logic and set theory (and Boolean algebra) are the next given in table 2.2.

Logic	*Set theory*
$\wedge$	$\cap$
$\vee$	$\cup$
$\neg$	$(\,\bar{}\,)$
True	1
False	0

Table 2.2: Equivalences between logic and set theory.

Thus, by using the equivalences between logic and set theory we can validate the membership functions for $\mu_{X\to Y}(u,v)$ as can be observed in table 2.3.

μ_X	μ_Y	$\mu_{X\to Y}$	$\neg Y$	$X \wedge (\neg Y)$	$\neg\,[X \wedge (\neg Y)]$	$\neg X$	$(\neg X) \vee Y$
1	1	**1**	0	0	**1**	0	**1**
1	0	**0**	1	1	**0**	0	**0**
0	1	**1**	0	0	**1**	1	**1**
0	0	**1**	1	0	**1**	1	**1**

Table 2.3: Validation of the tautologies for $\mu_{X\to Y}$.

Finally recall that in propositional logic there are two very important inference rules, *Modus Ponens* and *Modus Tollens*, which are based on the following criteria given in table 2.4.

Observe the resulting truth tables in the case of Boolean algebra in tables 2.5 and 2.6 in order to validate these inference rules.

		Modus Ponens	**Modus Tollens**
Premise 1	$\Rightarrow$	u is X	v is not Y
Premise 2	$\Rightarrow$	if u is X then v is Y	if u is X then v is Y
Consequence	$\Rightarrow$	v is Y	u is not X
Propositional logic	$\Rightarrow$	$(X \wedge (X \to Y)) \to Y$	$((\neg Y) \wedge (X \to Y)) \to (\neg X)$

Table 2.4: Modus Ponens and Modus Tollens.

Modus Ponens

μ_X	μ_Y	$X \to Y$	$X \wedge (X \to Y)$	$(X \wedge (X \to Y)) \to Y$
1	1	1	1	**1**
1	0	0	0	**1**
0	1	1	0	**1**
0	0	1	0	**1**

Table 2.5: Validation of Modus Ponens with Boolean algebra.

Modus Tollens

μ_X	μ_Y	$\neg Y$	$X \to Y$	$(\neg Y) \wedge (X \to Y)$	$\neg X$	$((\neg Y) \wedge (X \to Y)) \to (\neg X)$
1	1	0	1	0	0	**1**
1	0	1	0	0	0	**1**
0	1	0	1	0	1	**1**
0	0	1	1	1	1	**1**

Table 2.6: Validation of Modus Tollens with Boolean algebra.

Whereas Modus Ponens plays a very important role in engineering applications due to its *cause and effect* form, Modus Tollens is rarely considered. For this reason we will just study the Modus Ponens reasoning when we consider fuzzy logic.

Once we have observed the main aspects of propositional logic, basically focused on the implication and the reasoning rules, we can try to develop the implication of fuzzy sets for fuzzy logic as an extension from the implication of crisp sets for propositional logic. As we have already explained, this extension will not be so clear because most fuzzy systems need to express related propostitions between inputs and outputs while propositional logic works with unrelated propositions. Ultimately, this will cause us to define a different implication operator for fuzzy logic.

Like in our extension from crisp sets to fuzzy sets, our extension from crisp logic to fuzzy logic should be made by replacing the bivalent membership functions of crisp logic with fuzzy membership functions. Thus, the necessary statement *if u is X then v is Y* in order to evaluate an inference rule has a membership function $\mu_{X \to Y}(u, v)$ that might be computed by taking the previous tautologies for the implication of crisp sets and by replacing the conjunction, disjunction and negation operators with S-norms, T-norms and complements in the case of fuzzy sets. If we consider the the minimum and the product as T-norm and the maximum as S-norm, then we could compute the fuzzy implication (apparently) as

$$\mu_{X \to Y}(u, v) = 1 - \min\left[\mu_X(u), 1 - \mu_Y(v)\right] \tag{2.28}$$

$$\mu_{X \to Y}(u, v) = 1 - \mu_X(u)\left(1 - \mu_Y(v)\right) \tag{2.29}$$

by being the extension from the tautology $(X \to Y) \leftrightarrow \neg\left[X \wedge (\neg Y)\right]$ or

$$\mu_{X \to Y}(u, v) = \max\left[1 - \mu_X(u), \mu_Y(v)\right] \tag{2.30}$$

by being the extension from the tautology $(X \to Y) \leftrightarrow (\neg X) \vee Y$.

Once we have (apparently) the definitions for the fuzzy implication, we can evaluate a fuzzy rule as the extension from the inference rules with crisp sets, basically with the Modus Ponens form. In fact in fuzzy logic the Modus Ponens takes the form given in table 2.7 and it is called *Generalized Modus Ponens.*

Premise 1	$\Rightarrow$	u is X*
Premise 2	$\Rightarrow$	if u is X then v is Y
Consequence	$\Rightarrow$	v is Y*

Table 2.7: Generalized Modus Ponens.

The difference is that the fuzzy set in the antecedent of the rule X may be different from the fuzzy set of the premise X^*, although they will be defined on the same domain, as the fuzzy set in the consequent of the rule Y may be different from the fuzzy set of the consequence Y^*, although they will also be defined on the same domain, because now we do not have *true and false* values but degrees of membership.

In order to clarify it suppose that we have the inference rule *if a man is short then he will not be a very good basketball player* and the premise *This man is 168cm.* We can observe how the fuzzy sets involved within this operation are clearly not the same. In other words in crisp logic an inference rule is *fired* only if the first premise is exactly the same as the antecedent of the rule and the result of such operation is the rule's consequent. But in fuzzy logic a rule is *fired* so long as there is a degree of membership different from zero.

Now we come back to the (incoherent) use of the previous fuzzy implications in order to evaluate an inference rule given with the Generalized Modus Ponens form. In this case we will have a fuzzy relation $X \to Y$ defined as

$$X \to Y = \{(u,v), \mu_{X\to Y}(u,v)) \mid (u,v) \in U \times V\} \tag{2.31}$$

a fuzzy set defined with $\mu_{X^*}(u)$ and we want to obtain a fuzzy set defined with $\mu_{Y^*}(v)$. As we have explained before, this must be performed with a composition. Thus, by using the sup-star composition we have that

$$Y^*(v) = \left\{ \left(v, \sup_u \left[\mu_{X^*}(u) \star \mu_{X\to Y}(u,v) \right] \right) \mid v \in V \right\} \tag{2.32}$$

The problem arise when we define the fuzzy relation $\mu_{X\to Y}$ from the previous tautologies because then the necessary *cause and effect* definition is broken. Observe how the output fuzzy set Y^* will display high degrees of membership in spite of having an input fuzzy set X^* with low degrees of membership. This is due to the fact that the definition for the implication in propositional logic considers that a proposition is false if and only if the antecedent is false and the consequent is true.

E.H. Mamdani [75] seems to have been the first one to observe this problem. When he designed a fuzzy logic system to control a process, he used a definition for the implication operation according to the *cause and effect* requirement for engineering applications. Thus, he proposed the use of the *minimum implication* as

$$R_M = \mu_{X\to Y}(u,v) = \min\left[\mu_X(u), \mu_Y(v)\right] \tag{2.33}$$

The definition of the implication in fuzzy logic is not unique. For example P.M. Larsen [67] proposed later the *product implication* as

$$R_P = \mu_{X\to Y}(u,v) = \mu_X(u)\mu_Y(v) \tag{2.34}$$

These definitions have been clearly the most widely used inferences in the engineering applications, in spite of not being according to the definition of the implication with propositional logic. Although there are many other tautologies for $\mu_{X\to Y}$ in crisp logic and there are many other norms appart from the minimum, product and maximum in fuzzy logic, there has not been found any combination able to compute the fuzzy implication as an extension from the crisp implication which could express in a satisfactory manner the *cause and effect* relation so necessary in most fuzzy logic applications [46]. For this reason, a different *engineering* implication was defined in fuzzy logic, reason why fuzzy logic has been sometimes very criticized.

Anyway, many *not engineering* implications have also found its place in fuzzy logic. D. Dubois and H. Prade [25, 26] have done a lot of research on this field. For example when several rules are combined together to give an unique relation which defines the overall system in contrast to the analysis of each rule alone[3], a better interpretation of the system can be sometimes obtained with *not engineering* implications (i.e. Gödel implication). Some alternatives are:

- *Zadeh implication*

  $$R_Z = \mu_{X\to Y}(u,v) = \max\left(\min\left(\mu_X(u), \mu_Y(v)\right), 1 - \mu_X(u)\right)$$

  based on $(X \to Y) \leftrightarrow (X \wedge Y) \vee \neg X$

- *Kleene-Dienes or Dienes-Rescher implication*

  $$R_D = \mu_{X\to Y}(u,v) = \max\left(1 - \mu_X(u), \mu_Y(v)\right)$$

  based on the original $(X \to Y) \leftrightarrow (\neg X) \vee Y$

- *Stochastic implication*

  $$R_S = \mu_{X\to Y}(u,v) = \min\left(1, 1 - \mu_X(u) + \mu_X(u)\mu_Y(v)\right)$$

  based on the probabilistic equality $P(Y|X) = 1 - P(X) + P(X)P(Y)$

- *Lukasiewicz implication*

  $$R_L = \mu_{X\to Y}(u,v) = \min\left(1, 1 - \mu_X(u) + \mu_Y(v)\right)$$

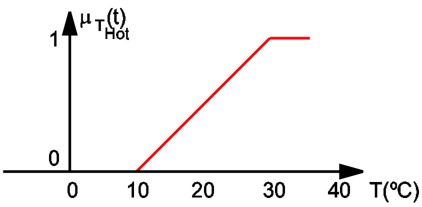
  based on the original $(X \to Y) \leftrightarrow (\neg X) \vee Y$ with the bounded sum as S-norm

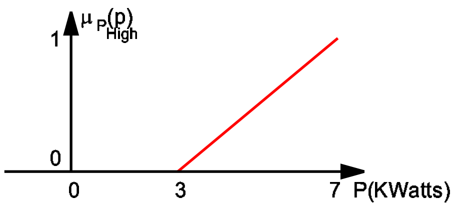
[3]Later we will comment these alternatives which are called *composition based inference* and *individual-rule based inference*, respectively.

These implications satisfy the following relation $R_P \subseteq R_M \subseteq R_Z \subseteq R_D \subseteq R_S \subseteq R_L$. Thus the most *restrictive* implications that satisfy the *cause and effect* requirement are the product (Larsen) and the minimum (Mamdani), as we have pointed out before. The list of possible fuzzy implications goes on (Goguen implication, Gödel implication, Sharp implication, ...). See [21] for more details.

Example

We can consider the following example in order to observe the need of an *engineering* definition for the implication in the case of fuzzy logic. Suppose that we want to regulate a climate system and one of the linguistic rules says *if the temperature is hot then the demanded power must be high.* The linguistic variables are defined with the following fuzzy sets:





In order to compare the propositional implication and the fuzzy implication we will consider for the first one the relation

$$\mu_{T_{Hot} \to P_{High}}(t, p) = 1 - \min\left[\mu_{T_{Hot}}(t), 1 - \mu_{P_{High}}(p)\right]$$

while we will consider the Mamdani implication for the second one

$$\mu_{T_{Hot} \to P_{High}}(t, p) = \min\left[\mu_{T_{Hot}}(t), \mu_{P_{High}}(p)\right]$$

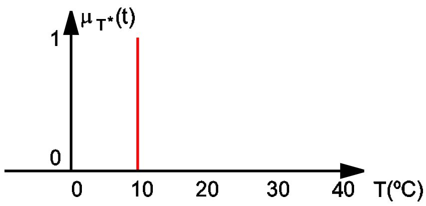
If we consider discretes universes of scope such that the temperature has a step of 5 Celsius degrees and the pressure has a step of 1KW, the resulting relation in the case of the propositional logic is

$$\begin{array}{rc} pow\backslash temp & \begin{array}{cccccccc} 0 & 5 & 10 & 15 & 20 & 25 & 30 & 35 \end{array} \\ \begin{array}{r} 0\\1\\2\\3\\4\\5\\6\\7 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0.75 & 0.50 & 0.25 & 0.25 & 0.25 \\ 1 & 1 & 1 & 0.75 & 0.50 & 0.50 & 0.50 & 0.50 \\ 1 & 1 & 1 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

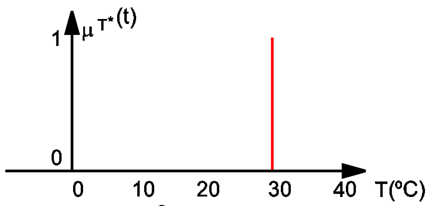
while the fuzzy relation is

$$\begin{array}{c|cccccccc} pow\backslash temp & 0 & 5 & 10 & 15 & 20 & 25 & 30 & 35 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 0.25 & 0.25 \\ 5 & 0 & 0 & 0 & 0.25 & 0.50 & 0.50 & 0.50 & 0.50 \\ 6 & 0 & 0 & 0 & 0.25 & 0.50 & 0.75 & 0.75 & 0.75 \\ 7 & 0 & 0 & 0 & 0.25 & 0.50 & 0.75 & 1 & 1 \end{array}$$

These relations show how the propositional implication does not agree with the *cause and relation* requirement while on the other hand, the fuzzy implication gives a relation which is according to it.
This phenomenon can be clearly observed by considering an input value. In fact we will consider two cases. For the first one the temperature will be 10 Celsius degrees while it will be 30 Celsius degrees for the second one.
First, we should create a fuzzy set which represents each situation (the fuzzy set called A^* in the theory). In fuzzy logic the best way to represent a single value is a singleton set:



First case



Second case

Thus, we can compute the *max-min composition* for each case between these sets given in discrete form as

$$\mu_{T^*} = \frac{0}{0} + \frac{0}{5} + \frac{1}{10} + \frac{0}{15} + \frac{0}{20} + \frac{0}{25} + \frac{0}{30} + \frac{0}{35}$$

$$\mu_{T^*} = \frac{0}{0} + \frac{0}{5} + \frac{0}{10} + \frac{0}{15} + \frac{0}{20} + \frac{0}{25} + \frac{1}{30} + \frac{0}{35}$$

and the previous relations (propositional and fuzzy).
For the first case when the temperature is 10 Celsius degrees, the output sets would be

$$\mu_{P^*} = \frac{1}{0} + \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} \text{ with propositional implication}$$

$$\mu_{P^*} = \frac{0}{0} + \frac{0}{1} + \frac{0}{2} + \frac{0}{3} + \frac{0}{4} + \frac{0}{5} + \frac{0}{6} + \frac{0}{7} \text{ with fuzzy implication}$$

and show how the propositional implication does not agree with the *cause and effect* requirement.

For the second case when the temperature is 30 Celsius degrees, the output sets would be

$$\mu_{P^*} = \frac{1}{0} + \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{0.25}{4} + \frac{0.5}{5} + \frac{0.75}{6} + \frac{1}{7} \text{ with propositional implication}$$

$$\mu_{P^*} = \frac{0}{0} + \frac{0}{1} + \frac{0}{2} + \frac{0}{3} + \frac{0.25}{4} + \frac{0.5}{5} + \frac{0.75}{6} + \frac{1}{7} \text{ with fuzzy implication}$$

and show how the fuzzy implication gives a clearer result of the demanded power. In fact in most control systems we will need a single output value and not a fuzzy value. This is normally computed with the center of gravity of the output set. Anyway this aspect will be treated later.

2.1.8 Fuzzy systems

Although fuzzy logic can work with many forms of linguistic reasoning, with the tolerance of the values[4], with different operators, ... most fuzzy systems and almost all fuzzy control systems, are given and are computed with few differences between them. In this section we will review them because these are the ones we will consider in our work.

Fuzzy logic provides the means for constructing fuzzy systems which consist of several rules by explaining how the linguistic labels are related. Each fuzzy rule is a statement where the antecedent and the consequent consist of fuzzy propositions that in turn are statements which join the linguistic variables with linguistic operators like *and, or, ...* Nevertheless, in the majority of fuzzy modeling and control problems, only the linguistic operator *and* is used to join the linguistic labels of the antecedent whereas the consequent is formed by only one linguistic label (MISO systems). For this reason we will just consider this case. Thus most rules are based on statements like

if < antecedent fuzzy proposition > then < consequent fuzzy proposition >

and in general are statements like[5]

$$\text{if } input_1 \text{ is } X_{1_r} \text{ and } input_2 \text{ is } X_{2_r} \text{ and ... } input_N \text{ is } X_{N_r} \text{ then } output \text{ is } Y_r$$

We could compute these statements as we have done before but there are two main reasons to demand a simpler method:

[4]If we recover the last example we could work with a fuzzified set different from a singleton, i.e. $\mu_{T^*} = \frac{0}{0} + \frac{0}{5} + \frac{0}{10} + \frac{0}{15} + \frac{0}{20} + \frac{0.5}{25} + \frac{1}{30} + \frac{0.5}{35}$ when the temperature is about 30 Celsius degrees.

[5]$input_1$, $input_2$, ... , $input_N$ are the input variables; X_{1_r}, X_{2_r}, ... , X_{N_r} are the fuzzy sets of each input variable; *ouptut* is the output variable; Y_r is a fuzzy set.

⇒ In most applications we will have single values for each input variable (not fuzzy numbers) and we will expect a single value for the output (not a fuzzy number).

⇒ It is interesting to avoid working with large matrices in order to simplify the computation. Furthermore the output of the system normally depends on a small part of these matrices for a given combination of input values.

Consequently, most fuzzy systems relate the linguistic statements with numerical values by using a simple algorithm in order to implement the mapping between input-output variables, the *fuzzy inference engine*. Here we will explain it. Anyway, we will first detail it in a formal manner and later we will review it in a graphical manner to emphasize its simplicity.

Consider a MISO fuzzy system with R *if ... then ...* rules where each one is defined with a fuzzy relation between the N fuzzy input sets in $U = U_1 \times U_2 \times \cdots \times U_N$ and the output fuzzy set in V.

For each rule, the antecedent has N fuzzy sets $X = \{X_1, X_2, \ldots, X_N\}$ whose membership functions are $\mu_X(u) = \{\mu_{X_1}(u_1), \mu_{X_2}(u_2), \ldots, \mu_{X_N}(u_N)\}$ while the consequent has one fuzzy set Y with $\mu_Y(v)$.

As the input sets are joined by AND's, the antecedent is in fact a fuzzy relation in U where for each possible input value $u = (u_1, u_2, \ldots, u_N)$ there will be a degree of membership such that $\mu_X(u) = \mu_{X_1}(u_1) \cap \mu_{X_2}(u_2) \cap \cdots \cap \mu_{X_N}(u_N)$ where the $\cap$ operator is any T-norm. This procedure is called the *conjunction*.

The fuzzy rule is defined by computing the implication between the antecedent X (fuzzy relation) and the consequent Y (fuzzy set). Consequently $X \to Y$ is defined as a fuzzy relation in $U_1 \times U_2 \times \cdots \times U_N \times V$ having a degree of membership $\mu_{X\to Y}(u_1, u_2, \ldots, u_N, v)$ for each possible $u_1 \times u_2 \times \cdots \times u_N \times v$. Recall that the fuzzy implication, in contrast with the propositional logic, is defined in general with a T-norm such as the minimum or the product. Thus $\mu_{X\to Y}(u_1, u_2, \ldots, u_N, v) = \mu_{X_1}(u_1) \cap \mu_{X_2}(u_2) \cap \cdots \cap \mu_{X_N}(u_N) \cap \mu_Y(v)$ where the $\cap$ operator is any T-norm.

As we have explained, the fuzzy inference engine gives an output value v^* for a given input value $u^* = (u_1^*, u_2^*, \ldots, u_N^*)$ applied to the fuzzy system. Thus, for each rule we must compute the composition between the input relation $X^* = \{X_1^*, X_2^*, \ldots, X_N^*\}$ and the fuzzy relation given by the implication of the rule $X \to Y$ as $Y^* = X^* \circ (X \to Y)$.

Before computing the composition, observe that the input relation X^* must be the fuzzy representation of the input sample $u^* = (u_1^*, u_2^*, \ldots, u_N^*)$. The most widely used form is the singleton. Thus, X^* has a degree of membership such that $\mu_{X^*}(u_1, u_2, \ldots, u_N)$ is equal to 1 if $u_i = u_i^* \; \forall \; i = 1 \ldots N$ and 0 otherwise. This procedure is called the *fuzzification*.

If the max-min composition or the max-product composition are used, the result of the composition $Y^* = X^* \circ (X \rightarrow Y)$ always gives a fuzzy set Y^* whose degrees of memberhsip are given as $\mu_{Y^*}(v) = \mu_{X \rightarrow Y}(u_1^*, u_2^*, \ldots, u_N^*, v)$ because the minimum or the product between the zeros of $\mu_{X^*}(u_1, u_2, \ldots, u_N)$ when $u_i \neq u_i^* \; \forall \; i = 1 \ldots N$ and any $\mu_{X \rightarrow Y}(u_1, u_2, \ldots, u_N, v)$ is always zero. Thus, the degrees of membership of $\mu_{X \rightarrow Y}(u_1^*, u_2^* \ldots, u_N^*, v)$ prevail over the rest when the maximum is computed.

Finally we desire an output single value from the R fuzzy sets Y^* of each rule which have been obtained previously. This procedure is called *defuzzification*. First, we will deal with the defuzzification of a single output set and then we will consider the R output sets. For this purpose, many alternatives have been proposed in the literature and have been deeply compared [21, 68]. However there are no scientific bases for any of them and its choice is an art rather than a science. In the case of engineering applications the main criterion is its computational simplicity. Here we can comment some of the most popular techniques:

- *Maximum defuzzification.* The output value v^* is the point of v for which μ_{Y^*} is maximum. A conflict arise if there is more than one value and furthermore it does not take into account the distribution of $\mu_{Y^*}(v)$.

- *Mean of maxima defuzzification.* The output value v^* is the mean of those v for which μ_{Y^*} is maximum. Some other conflicts may arise if for instance there are two points with the maximum μ_{Y^*} which are not together and the method gives as result a mid point whose degree of membership is low. Anyway, this problem only occurs if there is an incoherent definition of the system.

- *Center of gravity, center of area or centroid defuzzification.* The output value v^* is the center of gravity of the output set given by:

$$v^* = \frac{\int_V v \mu_{Y^*}(v) dv}{\int_V \mu_{Y^*}(v) dv}$$

  Despite considering all the information of the fuzzy set, it reveals two problems: the higher computational cost and the fact that the extreme values of the universe of scope are rarely obtained.

If the R output sets must be considered there are basically two options:

- The output sets are aggregated with a S-norm, normally the maximum, in order to apply a defuzzification technique to the resulting set.

- Each set is defuzzified alone and then the R results are weighted with the *relevance* of each rule given by the activation degree $\alpha = \mu_{X_1}(u_1^*) \cap \mu_{X_2}(u_2^*) \cap \cdots \cap \mu_{X_N}(u_N^*)$. Thus, if the defuzzified value of each set is $v_r^*|_{r=1\ldots R}$ and the activation degree of the corresponding rule is $\alpha_r|_{r=1\ldots R}$ then,
$$v^* = \frac{\sum_r \alpha_r v_r^*}{\sum_r \alpha_r}$$

The last alternative, usually called the *sum-product* operator, is in general faster by avoiding the aggregation. Furthermore, sometimes the values for v_r^* are computed ahead of time because most fuzzy sets are symmetric and its most *representative* value is its center, no matter the defuzzification method we choose. If the sets are not symmetric then the more *representative* value of the universe of scope can vary based on the application. For instance the sets placed in the extremes of the universe of scope are represented by the value with the maximum degree of membership, and thus the output range is usually bigger.

If the most *representative* value of each output set is computed every time (not stored at the beginning) as the center of gravity of the set, the procedure is called *center of sums defuzzification*. If these values are computed every time as the mean of maxima of the set, this procedure is called *height defuzzification*.

A comparison between all these methods can be found in [21]. The main conclusion says that the *height method* and the *center of sums* satisfy the same number of criteria and more than other methods. These criteria are continuity[6], disambiguity[7], plausibility[8], low computational complexity and the possiblity of weight counting[9].

Furthermore, later we will deal with a fuzzy structure (output singletons) where the sum-product with center of sums or height defuzzification, gives the same result than the original centroid method.

We have explained how a fuzzy system can be computed by evaluating each rule individually. However it is possible to find only a fuzzy relation describing the meaning of the overall set of rules. The first method is called

[6]A small change in the input should not result in a large change in the output.

[7]The result can not be ambiguous.

[8]The result should lie approximately in the middle of the support and with a high degree of membership.

[9]The result can be computed as a weighted average of the defuzzified sets of the rules.

individual-rule based inference in contrast to the second one called *composition based inference*. Nevertheless, the first method is preferred since it is more comprehensible and computationally very efficient by saving a lot of memory. For this reason all the fuzzy systems we consider will be individual-rule type.

Although all these operations can finally derive a compact equation able to obtain v^*, the overall method can be clarified with the steps given in figure 2.5, which will be explained in a graphical manner.

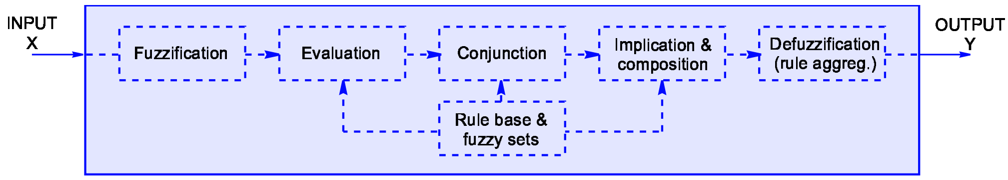


Figure 2.5: Basic steps in a fuzzy inference engine.

1. Fuzzification

   Normally the inputs of the fuzzy system are crisp values, reason why they have to be converted to fuzzy sets. If the input values are vague then these fuzzy sets can be modeled with fuzzy numbers such as triangular membership functions but they are normally modeled with singletons.

   Obviously the singleton fuzzifier is simpler and more used than its nonsingleton counterpart in order to avoid a higher computational complexity. Furthermore, the need for a more complex function has not been well justified until the moment.

   Figure 2.6 shows how the crisp value u_1 of the input variable $input_1$ is modeled with a singleton membership function.

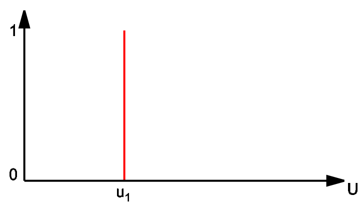


Figure 2.6: Fuzzification with singleton.

2. Evaluation of the input fuzzy sets
   In order to evaluate the antecedent proposition of the rule r in numerical terms, we first compute the degrees of membership of the fuzzy set related to the variable i (μ_{i_r}) for $1 \leq i \leq N$. This is computed by using a T-norm between the resulting fuzzy set of the fuzzification step and the fuzzy set μ_{i_r}.

   If the fuzzification has been done with singletons whose height is 1 then the evaluation of the input fuzzy set gives directly the degree of membership of the fuzzy set evaluated at u_i as shows figure 2.7.

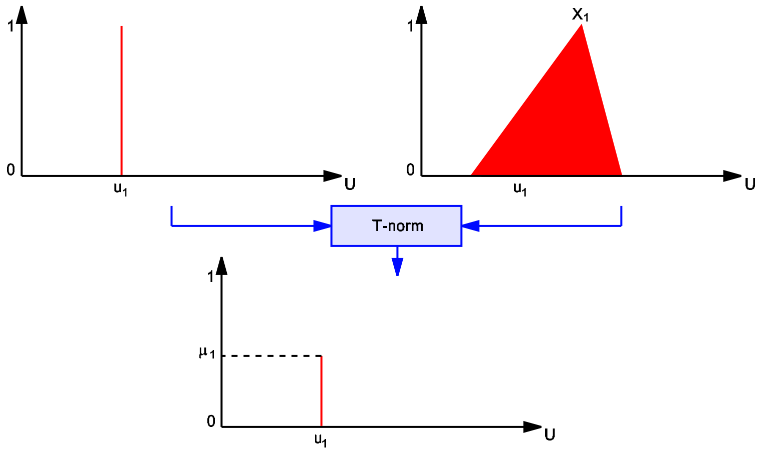


Figure 2.7: Evaluation of the input fuzzy sets.

3. Conjunction
   If the linguistic operator *and* is used to join all the input labels then a T-norm is applied to the degrees of membership (μ_{i_r}) obtained in previous step in order to obtain the activation degree α_r of the rule (also called the degree of fulfillment, the firing strength, ...) as exemplifies figure 2.8 where 2 input variables are considered. This procedure is called *conjunction*.

4. Implication & Composition
   This step computes the *then* operation. Recall that the most widely used implications in fuzzy logic are obtained by computing with a T-norm (mininum or product) the input-output relation build from the input-output sets.

   Thus, if we have fuzzified with singleton sets then the output fuzzy set is defined by applying a T-norm to each value of the output membership

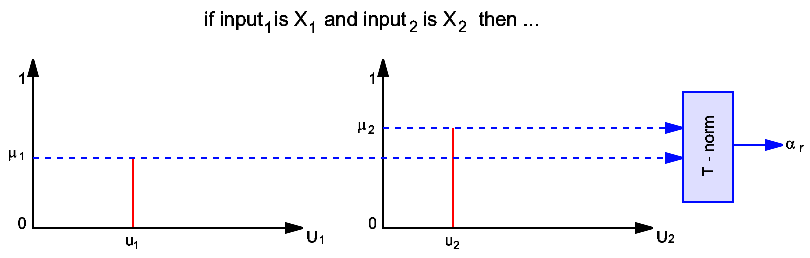


Figure 2.8: Conjunction.

function and the previous α_r value in order to obtain an output fuzzy set with a height equal to α_r.

Results can obvioulsy differ according to the selected T-norm (min, product, ...) as shows figure 2.9.

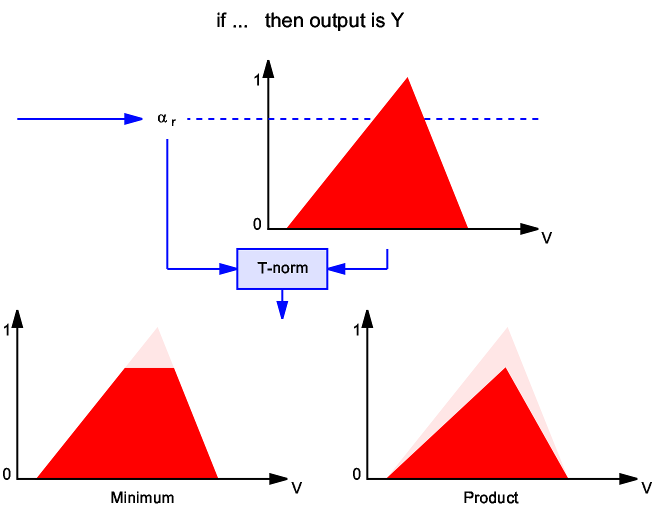


Figure 2.9: Implication.

5. Rule aggregation
   Once all the rules have been computed with the previous steps they are aggregated by using a S-norm as if they had been joined with the linguistic connector *or* (Zadeh said *else*) in order to obtain the final output fuzzy set of the overall system.

   In most cases the maximum is used as S-norm as shows figure 2.10 where we have considered a system with only 3 rules. Anyway this step can be avoided if one uses the product-sum operator for the defuzzification.

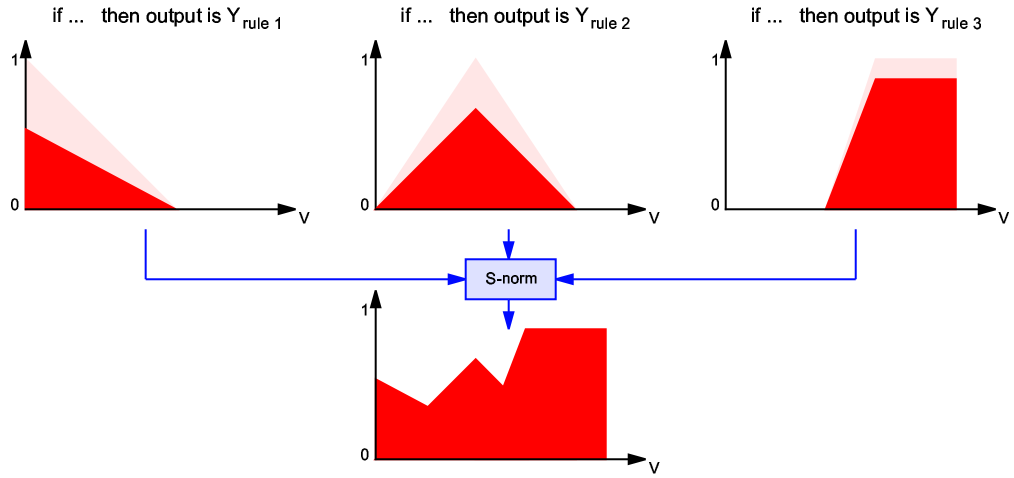


Figure 2.10: Rule aggregation.

6. Defuzzification
   Recall that the two common methods are the *center-of-gravity* (CoG) and *mean-of-maxima* (MoM).
   The first one is the same method employed to calculate the center of gravity of a mass with the difference that the points of the mass are replaced by the degrees of membership of the output set. The second one discriminate the part of the output fuzzy set whose degrees of membership are under a certain level, normally one, and thus, only the mean of those points with a degree of membership equal to one is considered. The CoG was the first one which was proposed whereas the MoM had its origin in the search of a clearly faster option. Despite giving different results as it is observed in figure 2.11, both methods give a similar overall performance.

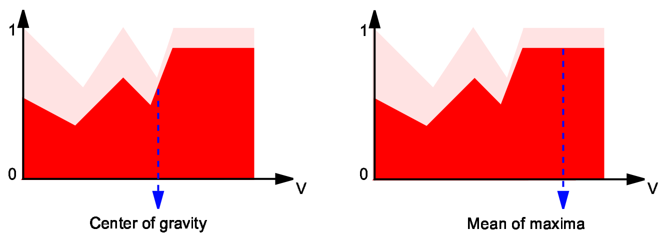


Figure 2.11: Defuzzification.

   Nevertheless and in order to accelerate the whole procedure, sometimes the rule aggregation procedure is not computed and then a simplified defuzzification procedure (with the sum-product operator) can be considered. In fact it is possible to obtain a crisp output value after the antecedent conjunction because the implication and the defuzzification steps are computed together.

The first thing we must consider is that the output fuzzy sets of the model are replaced by a representative real number, typically its center of gravity or its mean of maxima, which is stored at the beginning and thus, there is no need to compute them further more. Then the product operator is used as T-norm to perform the antecedent conjunction and also to weigh the output real values of the consequent in order to find the final output. This option, exemplified in figure 2.12, is the one we will consider in this work by being the fastest technique without loosing intelligibility properties.

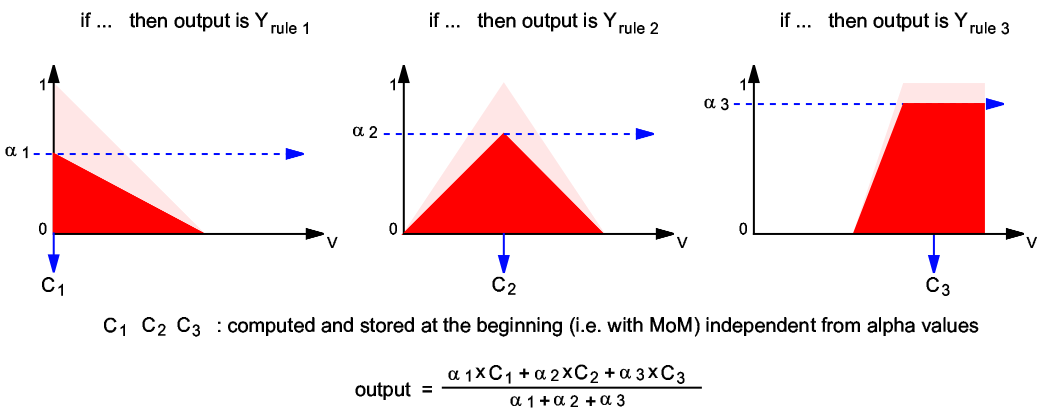


Figure 2.12: Centroid method with sum-product operator.

2.2 Types of fuzzy rule-based systems

By being interested in getting intelligible fuzzy models it seems necessary to remember the different kinds of fuzzy rule-based systems (FRBS) usually employed in order to compare them. This is a very significant aspect to be considered because nowadays there are basically three different fuzzy structures with different objectives and a different trade-off between intelligibility and accuracy.

2.2.1 Mamdani type

Ebrahim H. Mamdani was one of the responsible for the recognizion of fuzzy logic as an interesting alternative to classical control methods [75, 74]. He was the first to apply fuzzy logic to a real control problem and his efforts to reveal the features of fuzzy control were granted with a name in the fuzzy literature. Mamdani-type FRBS, also known as linguistic FRBS, are the main tool in order to develop intelligible fuzzy models. Main reason is that any element of the model is based on a linguistic label, for what the rules are expressed with the following structure:

if $input_1$ is S_{1_r} and $input_2$ is S_{2_r} and ... $input_N$ is S_{N_r} then $output$ is S_{O_r}

where $input_1$, $input_2$, ... , $input_N$ are the input variables; S_{1_r}, S_{2_r}, ... , S_{N_r} are the fuzzy sets of each input variable represented by a linguistic label; *ouptut* is the output variable; and S_{O_r} is a fuzzy set represented again by a linguistic label.

As all of terms are linguistic labels it appears particularly suitable when the human-machine interface is under observation because its information is clearly understandable.

2.2.2 Takagi-Sugeno type

On the other hand, T. Takagi and M. Sugeno [112] were not satisfied with the fact that Mamdani-type FRBS were not efficient with data-driven modeling algorithms. Therefore and in order to automate the tuning of model's parameters and also to reduce the necessary rules to model the system, they came up with an alternative rule format:

if $input_1$ is S_{1_r} and $input_2$ is S_{2_r} and ... $input_N$ is S_{N_r} then ...
$output$ is $K_{0_r} + K_{1_r} input_1 + K_{2_r} input_2 + \cdots + K_{N_r} input_N$

where K_{0_r}, K_{1_r}, K_{2_r}, ... ,K_{N_r} are real numbers.

Thus, TS-type FRBS, also known as Takagi-Sugeno-Kang models or linear fuzzy models, replace the consequent fuzzy proposition formed by a fuzzy set with a first order linear function of inputs. In this way each rule can be considered as a local linear model which are blended together by means of aggregation to define the overall output. In some cases also second order linear functions or even higher order linear functions have been considered in literature. Despite reducing significantly the necessary rules to model the system with a satisfactory error, the intelligibility is very poor by not considering linguistic labels in their consequents. For this reason this type of FRBS will not be considered in this work.

2.2.3 Output singleton type

A special case of FRBS is given when a TS-type FRBS is considered but with a zero order function in the consequent. This gives models whose rules are expressed as:

if $input_1$ is S_{1_r} and $input_2$ is S_{2_r} and ... $input_N$ is S_{N_r} then *output* is K_{0_r}

being K_{0_r} just a real number which can also be analyzed as a singleton fuzzy set placed at K_{0_r}.

Either in terms of intelligibility or accuracy, this type of FRBS is placed between Mamdani-type FRBS and TS-type FRBS.

Furthermore, as we have explained before, sometimes the defuzzification is computed considering only a representative number of the output fuzzy sets and thus, this is more or less equivalent to replace the Mamdani-type FRBS with the corresponding output singleton type because the values of the output membership functions other than crisp representations have no influence to the system output. This is why they are very used in many control systems.

This type of FRBS can be computed with a very compact equation. Consider this system with R rules and N input variables for each rule, whose sets are given with $\mu_{X_n^{(r)}}(u_n)$ $1 \leq n \leq N$, $1 \leq r \leq R$. The output singleton of each rule is placed at $Y^{(r)}$ $1 \leq r \leq R$. The input sets for each rule are connected with the linguistic operator AND. If we consider the product for the T-norm and the sum-product operator for the defuzzification, the output of the FRBS can be obtained with the following equation,

$$v(u) = \frac{\sum_{r=1}^{R} Y^{(r)} \left(\prod_{n=1}^{N} \mu_{X_n^{(r)}}(u_n) \right)}{\sum_{r=1}^{R} \left(\prod_{n=1}^{N} \mu_{X_n^{(r)}}(u_n) \right)} \tag{2.35}$$

In fact the choice of the product as T-norm instead of the minimum can be well justified because it retains more information. Suppose for instance a system with two inputs. When the product operator is used to compute a rule then both inputs will have an effect on the output. On the other hand, if the minimum operator is used, only one input will have an effect on the output. Several authors have shown how the interpolation properties of a fuzzy system can be better understood if the product is used [8, 98]. Anyway the method we will explain later can be applied with any of both operators, although we also prefer the product.

By being interested in seeking intelligible fuzzy models but if possible in a fast manner and also with a low computational cost, we will consider this type of FRBS in our work.

Its main benefits are that

⇒ they retain linguistic intelligibility in the manner of Mamdani-type FRBS

⇒ they simplify the model's computation

⇒ they also guarantee the attractive properties of TS-type FRBS which open the way for automated determination of the system's parameters from input-output data.

For example R. Jang [47] demonstrated how the output singleton type FRBS and the radial basis neural networks were equivalent and thus, many techniques developed for neural networks can also be applied to fuzzy systems.

2.3 Criteria about intelligibility and accuracy

2.3.1 Trade-off between intelligibility and accuracy

As Zadeh stated in his *Principle of incompatibility* [127]: "as the complexity of a system increases, our ability to make precise and yet significant statements about its behavior diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics".

Consequently, it turns out to be contradictory to obtain high degrees of intelligibility and accuracy. Thus, normally only one of these two properties prevails and therefore, fuzzy modeling has been a discipline with two major lines of work [13]:

- *Precise fuzzy modeling:* its main objective is to obtain precise models and therefore to obtain very similar results between the real system and the final model. This is an objective property which can be clearly defined in terms of error.

- *Linguistic fuzzy modeling:* its main objective is to obtain intelligible models and therefore to express the behavior of the system in an understandable way. This is a subjective property which depends on many factors. Nevertheless, there are several criteria which can be found in literature and will be considered in this work.

2.3.2 Intelligibility criteria

Some researchers have become aware of the intelligibility of fuzzy models and have proposed interesting properties to be analyzed. One must take into account that these properties must not be considered as an obligation as a whole and by depending on the problem some of them might be nonsense. Here we will expose them just to introduce how they are treated and we will recall them when explaining the method in the next chapter.

1. *Distinguishability of fuzzy partitions*
   When designing an intelligible fuzzy system, the fuzzy sets of each variable should be distinguishable with a clearly defined range of the universe of scope in order to associate linguistic terms.

   For this purpose, several techniques have been considered in literature: the constraint of the membership function parameters, the merging of similar fuzzy sets or the establishment of a semantic order among the linguistic terms.

   We will consider fuzzy partitions (or Ruspini partitions) by being distinguishable partitions as shows figure 2.13 where for each point of the universe of scope the total degree of membership should be equal to one and it could belong to two fuzzy sets at most.

   Furthermore, we will not limit the number of fuzzy sets because the maximum number of distinguishable linguistic terms is obviously a subjective condition and this constraint could degrade seriously the performance of the model.

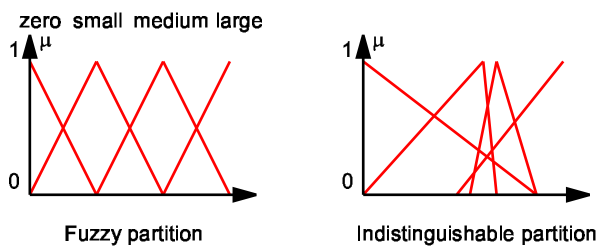


Figure 2.13: Distinguishability property.

2. *Justifiable number of labels*
   The number of fuzzy sets for each variable should be compatible with the number of conceptual entities. This property is justified with the *Principle of incompatibility* and also with some studies [79] suggesting

that the typical number of different labels which can be handled at the short-time memory is 7 ± 2 depending on each individual person.

For this reason many fuzzy modeling techniques bound the number of sets. Nevertheless and as we have explained, we will leave our method to use as many sets as necessary. Anyway, we will use a hierarchical method from which users can stop the process at any time if the intelligibility is degraded.

Furthermore and referring basically to the output sets, we will also use clustering techniques in order to reduce the number of fuzzy sets and to achieve a justifiable number of labels.

3. *Completeness of fuzzy partitions*
A partition is complete if for any point of the universe of scope there is at least one fuzzy set with a degree of membership greater than zero as shows figure 2.14. It avoids the fact that the system does not respond to a given input value.

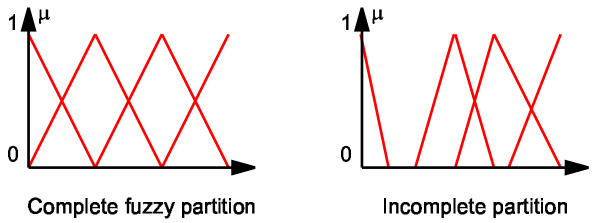


Figure 2.14: Completeness property.

By using the former fuzzy partitions we satisfy this property. Anyway, in some cases we will eliminate some of the final rules obtained with our method at the end of the process, if the input-output data can not validate them, and then we may break the completeness of fuzzy partitions if finally we do not have any rule related to a fuzzy set, as we will explain later in the method.

4. *Coverage of the universe of scope*
The membership functions must cover the entire universe of scope in order to facilitate a *linguistic representation* to each possible input. This property is defined in a way very similar to the completeness and in fact both properties are considered the same in most cases.

We fulfill it by placing two extreme fuzzy sets in the upper limit and in the lower limit of the universe of scope, and by working with fuzzy partitions whose sets are overlapped covering the entire universe of scope.

5. *Normality property*
   Fuzzy sets must be normal [116] because to consider non-normal membership functions is semantically to question the meaning of the represented linguistic term. We will work always with normal fuzzy sets.

6. *Natural zero positioning*
   Only if the nature of the problem requires it, one of the membership functions should be centered at zero in order to represent the linguistic term *nearly zero*. This is very useful in many control problems in order to track a reference signal.

   In fact our method will consider a special treatment when working with odd functions in order to assure this property.

7. *Compactness of fuzzy rules*
   The number of rules must be as small as possible. In general the typical full grid structure is not compact and the number of fuzzy rules increases exponentially with the input labels as shows figure 2.15.

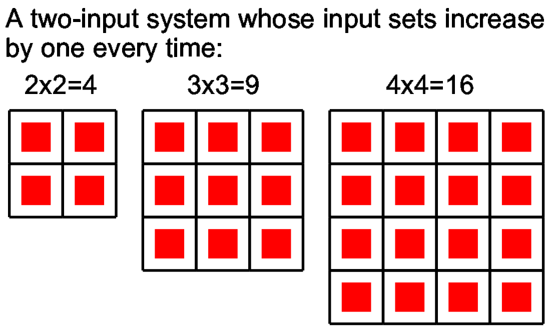
   

   Figure 2.15: Compactness property.

   We will not be able to avoid this grid structure by working with fuzzy partitions. Nevertheless, if the number of rules is very large we should at least try that the number of rules which are fired simultaneously for any input must remain as low as possible in order to furnish a simple local view of the behavior [92]. By working with fuzzy partitions we will always fire $2^{\text{number of inputs}}$ rules at most.

8. *Consistency of fuzzy rules*
   The rule base must be consistent with the knowledge it represents. This is accomplished by considering the following points [54]:

(a) If two rules have the same antecedent but different consequent, they are inconsistent.

(b) If two rules have similar antecedents, but fully different consequent, they are inconsistent.

(c) If the antecedents of two rules are different, they are always considered to be consistent, no matter what consequents they are.

(d) If two rules have similar antecedents and similar consequents, they are consistent.

In fact [54] developed a consistency measure by using a fuzzy similarity measure basically in order to control the previous second point because the other three are in general satisfied. Nevertheless we will not control it by believing that consistency must be assured with the input-output data set. If the fuzzy model is according to the information available, it must be as consistent as the input data.

9. *Transparency*
One should understand the influence of each system's parameter on the system output. Transparency was deeply analyzed by A. Riid [98] who developed a complete study of 0^{th} order Takagi-Sugeno systems (output singleton FRBS) for the same fuzzy partitions we use. The most significant conclusions to be considered are [98, 99]:

(a) With 50% overlapped fuzzy sets, the interval of the output values which is consequence of the explicit contribution of a given rule is defined with single points called transparency checkpoints.

(b) Consider triangular membership functions defined with three points being l the lower limit, c the center and u the upper limit. In order to assure the input transparency with triangular membership functions, that is to guarantee the existence of transparency checkpoints, the following condition must be applied: $u_{\text{set i-1}} \leq c_{\text{set i}} \leq l_{\text{set i+1}}$

(c) The interpolation between transparency checkpoints is linear in a SISO system if the centroid defuzzification is considered. Otherwise the output due to neighboring rules (transparency checkpoints) is not linear because a planar surface is defined by three freely chosen points.

In this way and since we will work with output singleton FRBS, we can consider these aspects in order to understand the performance of the models.

2.3.3 Accuracy criteria

Accuracy refers to the capacity of the model to faithfully represent the modeled system by giving output values very similar to the original data. This property is clearly objective in contrast with the intelligibility criteria which were very subjective.

The common technique able to quantify the accuracy is the error defined as the difference between the output value obtained with the model and the expected value from the original data. Anyway the *error* can be defined by taking into account different alternatives which will be studied in the next chapter.

2.4 Summary

In this chapter we have reviewed the basics of fuzzy theory and fuzzy logic. The presented results are only a very small part of the body of fuzzy sets and fuzzy logic but we have considered only those necessary aspects to understand the rest of the thesis.

We have first defined the terms related to fuzzy logic to be used in the thesis. Then we have reviewed and justified the type of system we will work with. We have shown how in spite of the existence of many alternatives in fuzzy logic, the use of FRBS with output singletons and the use of the singleton fuzzyfication with the sum-product operator, simplify the overall computation without degrading its most desirable properties.

Furthermore we have explained the criteria which should be considered if we are interested in seeking intelligible fuzzy models, which are the core of this work. They have been introduced in order to be applied as much as possible in the method we propose and in fact most refinements we will describe later have been adopted in order to satisfy as many properties as possible.