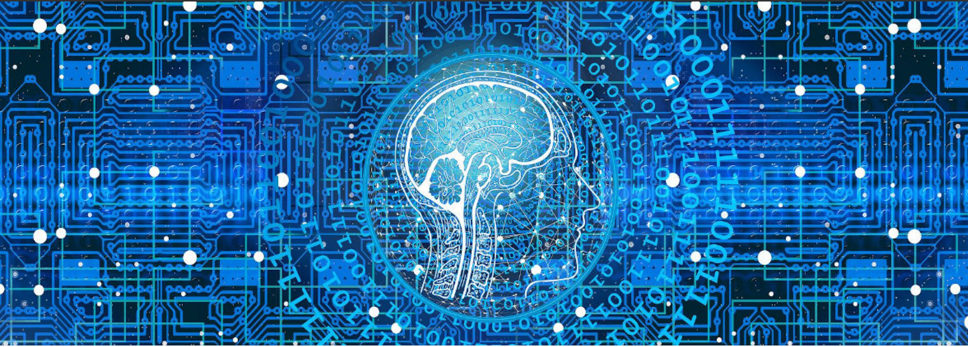


Chapter 10: Fuzzy Logic

Source A: Foundations of Artificial Intelligence by J. Boedecker, W. Burgard, F. Hutter, B. Nebel (Uni. Freiburg)

Content

- What is Fuzziness
- Fuzzy Sets
- Fuzzy Logic
- Reasoning with Fuzzy

The origins of fuzziness

- In contrast to precise, limited and constrained language that we use to describe notions, entities, and concepts while building logical models (so far), the real-life concepts and entities are described in much less rigid way.
- Let's consider the following example. In real life sentence: **John is a tall guy**
  - may mean many things, depending of our perspective, the place we live (meaning of tall is different in e.g. Japan) and so on.
  - But, if we want to feed John's data into computer, we have to determine his height precisely – say 190 cm.
  - But what if do not know John's height exactly?
- In real life we are doing perfectly all right with the sentences like:
- It takes about 40 minutes to reach an airport if the traffic is not too heavy.
- But what if we want a computer to understand such a sentence? How do we represent about and too heavy in a machine?

Fuzzy concepts and fuzzy sets

- In 1965 Lotfi Zadeh proposed a different way of looking at notions such as: set, containment, subset. His target was to make it possible to deal with concepts (sets) and dependencies that by nature are imprecise and vague - so called fuzzy concepts (sets).
- Again, the example of such concept is the natural language sentence:
- **John is a tall guy.**
- If we know, that John is 175 cm tall, we may start to wonder about the validity of the above sentence.
- In classical set theory we are forced to make definite decision whether 175 cm qualifies John as tall or not.
- In the fuzzy set theory we may be more subtle and express to what degree 175 cm of height makes John a tall guy.

Fuzzy sets

In classical set theory and with classical binary logic, that we usually employ when doing things with use of computer, the (contents of a) set A in some universe X can be expressed in the form of its characteristic (containment) function:
Such classical, rigidly defined set we will further call crisp or definite. The key step in defining fuzzy set theory is the replacement of characteristic function χ_A by function $\mu_A : X \Rightarrow [0,1]$
μ_A is called membership function or fuzzy membership.
If

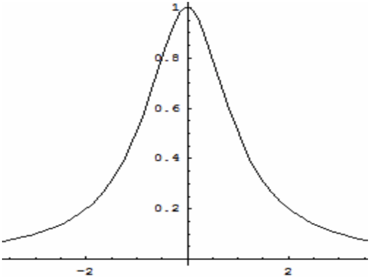
$$\chi_A(x) = \begin{cases} 1 & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A \end{cases}$$

then A is a classical set i.e., crisp (definite) set. If there exists $x \in X$ such that $0 < \mu_A(x) < 1$ - the set A is fuzzy.

Fuzzy sets - examples

A classical example of a fuzzy set near - zero provided by Zadeh for the concept of real number near 0. This set may be defined, for example, by the following membership function:

$$\mu_{near-zero} = \frac{1}{1+x^2}$$

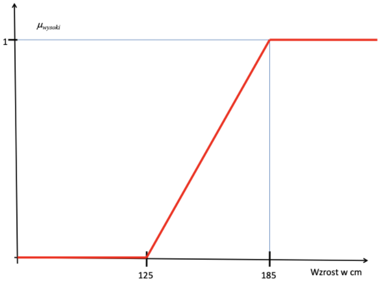


Fuzzy sets - examples

The previously considered notion of tall guy could be given – for height x in centimetres – by membership:

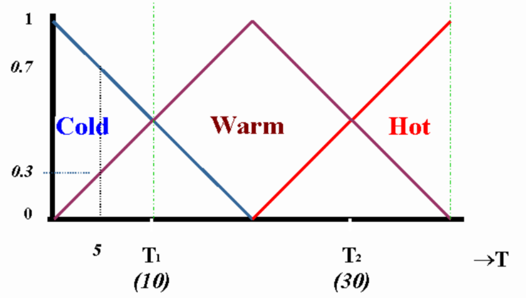
$$\mu_{tall} = \begin{cases} 0 & \text{if} \quad x \leq 125 \\ 1 & \text{if} \quad x \geq 185 \\ \frac{x-185}{2} + 1 & \text{if} \quad 125 < x < 185 \end{cases}$$

:



Fuzzy sets - examples

- Another example of three fuzzy sets for the notion of cold, warm and hot, where x is a temperature.



Height of a fuzzy set

The height of a fuzzy set is the largest degree of membership possessed by all elements of the fuzzy set. Thus, for a fuzzy set μ over a basic set A, its height is defined as:

$$H(F) = \max\{\mu(x) : x \in X\}$$

Accordingly, two-valued logic works with sets whose height is 1.

α-cut

- The α-cut ($\alpha \in [0, 1]$) of a fuzzy set μ is the set of all elements whose membership in μ is at least α.
- With this given α and a fuzzy set μ, their so-called α-cuts can thus be determined.
  - But also from the knowledge of the α-cuts alone a fuzzy set can be approximated, so that it is sufficient to store single α-cuts (usually of 0.1, 0.2, ... 1.0) in the computer. The original function can then be computed from these, which is facilitated by the use of triangular or trapezoidal sets.

4 Operations on fuzzy sets

In order to be able to compute meaningfully with fuzzy sets, one needs known concepts such as average, union and complement. There are two norms with which these three concepts can be formulated in such a general way

- t-norm
- t-conorm
- negation

t-Norm

In order to create an average operator, it must satisfy some minima- l requirements of mathematics. These are:

- Existence of a neutral element: $T(a, 1) = a$
- Monotonicity: $a \leq b \Rightarrow T(a, c) \leq T(b, c)$
- Commutativity: $T(a, b) = T(b, a)$
- Associativity: $T(a, T(b, c)) = T(T(a, b), c)$

The most general average operator that just satisfies these requirements is a function $T : [0, 1]^2 \Rightarrow [0, 1]$ and is called the t-norm. This is also often called the AND operator of fuzzy logic. Here, T is monotone and $T(a, 0) = 0$ holds.

- Point 1 states that the intersection of a fuzzy set with an ordinary set results in the degree of membership being preserved.
- Point 2 guarantees that a degree of membership in the average cannot become smaller when a fuzzy set is intersected with a "larger" fuzzy set.
- Points 3 and 4 should be self-evident for average operators.

t-Conorm

Also a union operator must satisfy certain minimal requirements: Existence of a neutral element: $\perp(a, 0) = a$

- Monotonicity: $a \leq b \Rightarrow \perp(a, c) \leq \perp(b, c)$
- Commutativity: $\perp(a, b) = \perp(b, a)$
- Associativity: $\perp(a, \perp(b, c)) = \perp(\perp(a, b), c)$

The most general union operator that just satisfies these requirements is a function $\perp : [0, 1]^2 \Rightarrow [0, 1]$ and is called a t-conorm. This is also often called the OR operator of fuzzy logic.

Negation

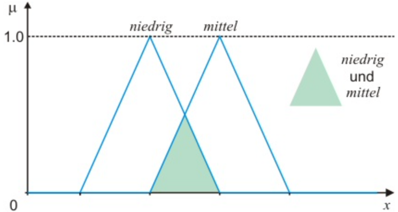
Every constructed negation function is a special case of the generalized negation:

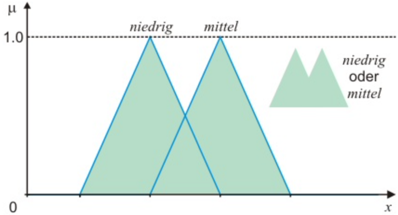
- This is a function n : $[0,1] \Rightarrow [0,1]$ with $\alpha \Rightarrow 1-\alpha$.

By this definition of negation, one obtains from the t-norm the t-conorm and vice versa, so that the dual principle is preserved as in classical logic. That is, from negation and t-norm the t-conorm can be determined, and from negation and t-conorm the t-norm.

Choice of norms

- The simplest functions corresponding to these general operators and their requirements are the "min" and "max" functions corresponding to averaging and unification.
- The following graph is intended to illustrate this and it can be interpreted these terms as the colloquial "and" and "or" respectively.
- Here, the "min" and "max" functions are defined on the membership function as follows:
  - Average = min $(\mu_{Function1}(x), \mu_{Function2(x)})$
  - Association = max $(\mu_{Function1}(x), \mu_{Function2}(x))$
- There are more possibilities, but we leave it with this





Fuzzy relations

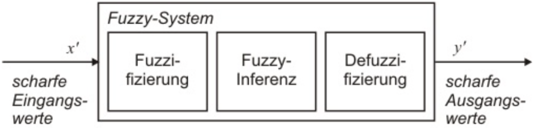
- It is often necessary to determine the degree of membership of not only one value, but several. Thus several basic sets exist, which are mapped by means of a relation to a fuzzy set.
- A fuzzy relation is understood to be fuzzy sets whose basic set represents a Cartesian product of several basic sets. For example, a two-digit fuzzy relation is a mapping $\rho : X \times Y \Rightarrow [0, 1]$. The value $\rho(x, y)$ indicates how strongly x is related ρ to y.
- Through these relations, one can link several conditions together and obtain a common result. For example, the operations for average and union already discussed are nothing but fuzzy relations on themselves $(X \times X)$.

Fuzzy Rules

- Fuzzy relations are common, especially in computer applications. They arise from linking fuzzy sets defined on several basic sets. This is usually done in the form
  - If x = low and y = medium then z = high
- Such a linkage represents a two-digit fuzzy relation and can of course be extended arbitrarily (proof by substitution and complete induction). The linkage of the "and" operator can be realized by means of the MIN operator as shown for operations on fuzzy sets:
  - $\mu R(x, y) = min(\mu_{low}(x), \mu_{medium}(x))$
- Accordingly, for an "or" linkage, one uses the MAX- operator.

Fuzzy systems

Fuzzy systems are by no means systems that work with uncertain information. Rather, they are systems which are based on fuzzy logic.



- As in a normal system, they receive sharp input values (e.g. temperature, pressure, speed) and output sharp output values accordingly (e.g. open pressure valve).
- Only the core of a fuzzy system, the so-called fuzzy inference (see Closing with Fuzzy) works with fuzzy information, which is why a fuzzyfication and then a defuzzification must be switched beforehand in order to convert precise values into fuzzy values and vice versa.

Reasoning with Fuzzy

Now that the basics of fuzzy logic are known, let's look at how to work with the help of this logic. For this purpose we have already learned the structure of a fuzzy system, the individual components of which will now be examined.

- Fuzzyfication
- Fuzzy inference
- Defuzzification

Fuzzyfication

- A fuzzy system first receives a safe value to work with (e.g. sensor values).
- However, the system itself works with fuzzy information, which has to be generated in the first step of the fuzzy system.
- This step is called fuzzyfication
  - For a given linguistic variable [(”strong”, ”fast”,...”)], let membership functions be defined for its linguistic terms.
  - For a crisp input value, the membership degrees with respect to all linguistic terms are determined using the fuzzy sets.
  - Here, fuzzy sets of linguistic terms are usually defined in such a way that for a sharp input value, the sum of the degrees of membership with respect to all linguistic terms of a linguistic variable add up to one. Thus, linguistic variables overlap and ensure that there are no gaps in the description.
  - Thus, for each fixed input value it can be calculated unambiguously to which degree it can be assigned to a linguistic variable.

Fuzzy inference

- Fuzzy inference is the core of a fuzzy system.
- Fuzzy rules are used to work with fuzzy information (otherwise a rule would have to be created for every possible input, which is often not possible).
- To apply a rule of the form already described to fuzzy information, a rule is needed for the case where the premise is more or less true and thus, of course, the conclusion is more or less true. This evaluation is also called fuzzy reasoning or approximate reasoning. There are different methods for this:
  - reasoning with negation and linkage
  - Mamdani implication
  - further implications, which we will not cover here

Reasoning with negation and concatenation

- Implications defined analogously to classical logic (complement, intersection, union).
- In this logic, the following statement holds:
  - $(p \Rightarrow q) \iff (\neg p \vee q)$.
- Thus, using the correspondences of negation ($\neg A := 1 - \mu A(x)$) and union ($A \vee B := max(\mu_A(x), \mu_B(x))$), we can conclude the following implication:
  - $A \Rightarrow B := max(1 - \mu_A(x), \mu_B(x))$
- However, this implication is not quite clever, because if $\mu_A(x) = 0$, one has the problem that the conclusion has the value 1. This has disadvantages when linking with other rules. Therefore there are more sophisticated variants. The best known is the Mamdani implication.

Mamdani implication

- The Mamdani implication is based on the idea that the truth value of the result should not be higher than the truth value of the premise:
  - $\mu A \Rightarrow B(x, y) := min(\mu_A(x), \mu_B(x))$
- Here, problem described above is circumvented.
- There are much more implication possibilities (algebraic product, Zadeh implication, ...) but we do not treat them here.

Rule evaluation

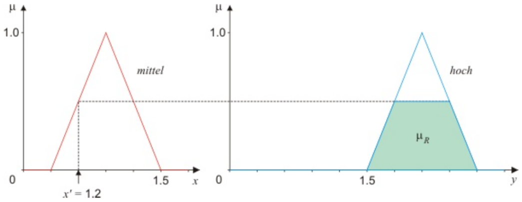
- The rule evaluation is now the secret of the fuzzy system, because here the entire knowledge of the system is stored, which determines how the output values result from the input values.
- However, since soft rules are present and these refer to linguistic variables or their fuzzy sets, we obtain here by the fuzzyfication corresponding values μ of the membership function.
- Here, a rule can consist of premises linked by several operators (e.g., AND or OR), all of which are satisfied to a certain degree μ, thus yielding an overall weight of the degree of satisfaction of the rule premise.
- Moreover, a fuzzy system can also have multiple rules, so that the appropriate rule must first be selected. Therefore, the question arises how this total weight should be calculated or how the valid rule should be determined.

Rule evaluation: Evaluation of individual rules

- In order to calculate the total weight of a rule consisting of several partial premises, a MIN-operator is used, provided that the linkage consists only of AND-operators (this is the rule, since OR-operators can be defined as a separate rule and then the following section comes into play).
- This is the most common variant, but the product can also be formed. However, this ultimately depends on the system to be formed.

Rule evaluation: Evaluation of several rules

- If the fuzzy system has several rules, these are linked with a MAX operator. This corresponds to an OR operation of all rules, which guarantees that all rules are taken into account.
- The Mamdani implication is also called MAX-MIN inference because of the procedure described above, the algebraic product is also called MAX- PROD inference.



This plot shows the evaluation of the rule "If x=mean, then y=high", where you can see that no sharp output value, but a truncated fuzzy set comes out as a result

Defuzzification

As a result of an inference process we get again a fuzzy set, i.e. we get a fuzzy information μ_R. However, a fuzzy system must now convert this information back into sharp output values so that terminal devices (e.g. pressure valves) can be controlled. For this conversion, there are again different methods, of which we will deal with the most important two:

- Maximum method
- Center of gravity method

Maximum method

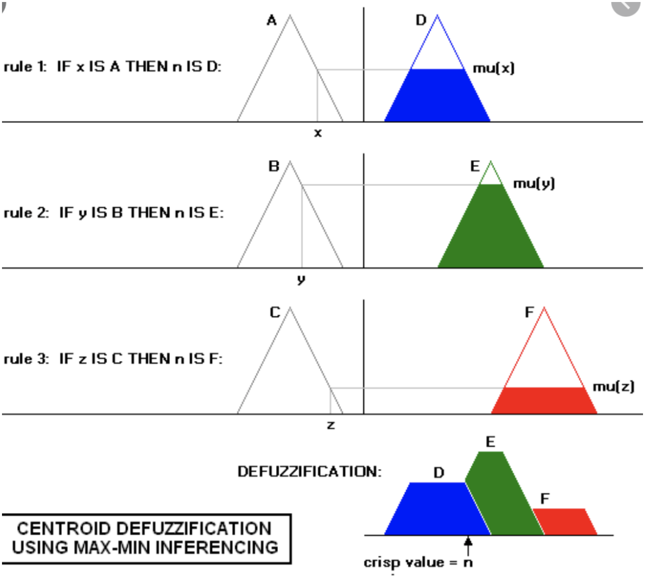
In the maximum method, only the rule with the highest degree of fulfillment is used to determine the sharp initial value. There are three different variants:

- Selection of the mean value: Here, the left and right corner points of the maximum are added and divided by two.
- Choice of the left edge point: Here, only the left corner point is used.
- Selection of the right edge point: Here, the right corner point is used accordingly.

With the MAX-PROD variant, all three variants provide the same result. However, this method has advantages and disadvantages which should be estimated, because:

- The computational effort for a sharp initial value is low
- The calculated value is independent of the actual degree of fulfillment of the Rule with maximum degree of fulfillment
- If there are several rules with the same degree of fulfillment as a maximum, the average value should be chosen. the average value should be chosen, because different conclusions could occur.

Example for Defuzzification



Example

- T.B.D.