Tutorial On Fuzzy Logic

Jan Jantzen <jj@oersted.dtu.dk>[1]

Abstract

Fuzzy logic is based on the theory of fuzzy sets, where an object's membership of a set is gradual rather than just member or not a member. Fuzzy logic uses the whole interval of real numbers between zero (*False*) and one (*True*) to develop a logic as a basis for rules of inference. Particularly the fuzzified version of the modus ponens rule of inference enables computers to make decisions using fuzzy reasoning rather than exact.

Contents

1 Introduction **2**

2 Fuzzy Set Theory **2**

2.1 Fuzzy Sets 2

2.2 Fuzzy Set Operations 8

3 Fuzzy Logic **13**

3.1 Propositions 13

3.2 Inference 20

4 Fuzzy Rules **23**

4.1 Linguistic Variables 24

4.2 Modus Ponens Inference 25

5 Summary **27**

References **28**

[1] Technical University of Denmark, Oersted-DTU, Automation, Bldg 326, 2800 Kongens Lyngby, DENMARK. Tech. report no 98-E 868 (logic), revised 17 Jun 2008.

1 Introduction

An assertion can be more or less true in fuzzy logic. In classical logic an assertion is either true or false — not something in between — and fuzzy logic extends classical logic by allowing intermediate truth values between zero and one. Fuzzy logic enables a computer to interpret a linguistic statement such as

’if the washing machine is half full, then use less water.’

It adds intelligence to the washing machine since the computer infers an action from a set of such if-then rules. Fuzzy logic is ’computing with words,’ quoting the creator of fuzzy logic, Lotfi A. Zadeh.

The objective of this tutorial is to explain the necessary and sufficient parts of the theory, such that engineering students understand how fuzzy logic enables fuzzy reasoning by computers.

2 Fuzzy Set Theory

Fuzzy sets are a further development of mathematical set theory, first studied formally by the German mathematician Georg Cantor (1845-1918). It is possible to express most of mathematics in the language of set theory, and researchers are today looking at the consequences of ’fuzzifying’ set theory resulting in for example fuzzy logic, fuzzy numbers, fuzzy intervals, fuzzy arithmetic, and fuzzy integrals. Fuzzy logic is based on fuzzy sets, and with fuzzy logic a computer can process words from natural language, such as ’small’, ’large’, and ’approximately equal’.

Although elementary, the following sections include the basic definitions of classical set theory. This is to shed light on the original ideas, and thus provide a deeper understanding. But only those basic definitions that are necessary and sufficient will be presented; students interested in delving deeper into set theory and logic, can for example read the book by Stoll (1979[10]); it provides a precise and comprehensive treatment .

2.1 Fuzzy Sets

According to Cantor a *set* $\mathcal{X}$ is a collection of definite, distinguishable objects of our intuition which can be treated as a whole. The objects are the *members* of $\mathcal{X}$. The concept ’objects of our intuition’ gives us great freedom of choice, even sets with infinitely many members. Objects must be ’definite’: given an object and a set, it must be possible to determine whether the object is, or is not, a member of the set. Objects must also be ’distinguishable’: given a set and its members, it must be possible to determine whether any two members are different, or the same.

The members completely define a set. To determine membership, it is necessary that the sentence ’x is a member of $\mathcal{X}$’, where x is replaced by an object and $\mathcal{X}$ by the name of a set, is either true or false. We use the symbol $\in$ and write $x \in \mathcal{X}$ if object x is a member of the set $\mathcal{X}$. The assumption that the members determine a set is equivalent to saying: Two sets $\mathcal{X}$ and $\mathcal{Y}$ are equal, $\mathcal{X} = \mathcal{Y}$, iff (if and only if) they have the same members. The set whose members are the objects $x_1, x_2, \ldots, x_n$ is written

$$\{x_1, x_2, \ldots, x_n\}.$$

In particular, the set with no members is the *empty set* symbolized by $\emptyset$. The set $\mathcal{X}$ is included in $\mathcal{Y}$,

$$\mathcal{X} \subseteq \mathcal{Y}$$

iff each member of $\mathcal{X}$ is a member of $\mathcal{Y}$. We also say that $\mathcal{X}$ is a subset of $\mathcal{Y}$, and it means that, for all x, if $x \in \mathcal{X}$, then $x \in \mathcal{Y}$. The empty set is a subset of every set.

Almost anything called a set in ordinary conversation is acceptable as a mathematical set, as the next example indicates.

Example 1 *Classical sets*
The following are lists or collections of definite and distinguishable objects, and therefore sets in the mathematical sense.

(a) The set of non-negative integers less than 3. This is a finite set with three members $\{0, 1, 2\}$.

(b) The set of live dinosaurs in the basement of the British Museum. This set has no members, it is the empty set $\emptyset$.

(c) The set of measurements greater than 10 volts. Even though this set is infinite, it is possible to determine whether a given measurement is a member or not.

(d) The set $\{0, 1, 2\}$ *is the set from (a). Since* $\{0, 1, 2\}$ *and* $\{2, 1, 0\}$ *have the same members, they are equal sets. Moreover,* $\{0, 1, 2\} = \{0, 1, 1, 2\}$ *for the same reason.*

(e) The members of a set may themselves be sets. The set

$$\mathcal{X} = \{\{1, 3\}, \{2, 4\}, \{5, 6\}\}$$

is a set with three members, namely, $\{1, 3\}$, $\{2, 4\}$, *and* $\{5, 6\}$. *Matlab supports sets of sets, or nested sets, in cell arrays. The notation in Matlab for assigning the above sets to a cell array* `x` *is the same.*

(f) It is possible in Matlab to assign an empty set, for instance: `x = {[]}`.

Although the brace notation $\{\cdot\}$ is practical for listing sets of a few elements, it is impractical for large sets and impossible for infinite sets. How do we then define a set with a large number of members?

An answer requires a few more concepts. A *proposition* is an assertion (declarative statement) which can be classified as either true or false. By a *predicate* in x we understand an assertion formed using a formula in x. For instance, '$0 < x \leq 3$', or '$x > 10$ volts' are predicates. They are not propositions, however, since they are not necessarily true or false. Only if we assign a value to the variable x, each predicate becomes a proposition. A predicate $P(x)$ in x defines a set $\mathcal{X}$ by the convention that the members of $\mathcal{X}$ are exactly those objects a such that $P(a)$ is true. In mathematical notation:

$$\{x \mid P(x)\},$$

read 'the set of all x such that $P(x)$.' Thus $a \in \{x \mid P(x)\}$ iff $P(a)$ is a true proposition.

A system in which propositions must be either true or false, but not both, uses a two-valued logic. As a consequence, what is not true is false and vice versa; that is the *law of the excluded middle*. This is only an approximation to human reasoning, as Zadeh observed:

> Clearly, the "class of all real numbers which are much greater than 1," or "the class of beautiful women," or "the class of tall men," do not constitute classes or sets in the usual mathematical sense of these terms. (Zadeh, 1965[12])

Zadeh's challenge we might call it, because for instance 'tall' is an elastic property. To define the set of tall men as a classical set, one would use a predicate $P(x)$, for instance $x \geq 176$, where x is the height of a person, and the right hand side of the inequality a threshold value in centimeters (176 centimeters $\simeq$ 5 foot 9 inches). This is an abrupt approximation to the concept 'tall'. From an engineering viewpoint, it is likely that the measurement is uncertain, due to a source of noise in the equipment. Thus, measurements within the narrow band $176 \pm \varepsilon$, where ε expresses variation in the noise, could fall on either side of the threshold randomly.

Following Zadeh a *membership grade* allows finer detail, such that the transition from membership to non-membership is gradual rather than abrupt. The membership grade for all members defines a *fuzzy set* (Fig. 1). Given a collection of objects $\mathcal{U}$, a fuzzy set $\mathcal{A}$ in $\mathcal{U}$ is defined as a set of ordered pairs

$$\mathcal{A} \equiv \{\langle x, \mu_{\mathcal{A}}(x)\rangle \mid x \in \mathcal{U}\} \tag{1}$$

where $\mu_{\mathcal{A}}(x)$ is called the *membership function*for the set of all objects x in $\mathcal{U}$ — for the symbol '$\equiv$' read 'defined as'. The membership function relates to each x a membership grade $\mu_{\mathcal{A}}(x)$, a real number in the closed interval $[0, 1]$. Notice it is now necessary to work with pairs $\langle x, \mu_{\mathcal{A}}(x)\rangle$ whereas for classical sets a list of objects suffices, since their membership is understood. An *ordered pair* $\langle x, y\rangle$ is a list of two objects, in which the object x is considered first and y second (note: in the set $\{x, y\}$ the order is insignificant).

The term 'fuzzy' (indistinct) suggests an image of a boundary zone, rather than an abrupt frontier. Indeed, fuzzy logicians speak of classical sets being *crisp sets*,to distinguish them from fuzzy sets. As with crisp sets, we are only guided by intuition in deciding which objects are members and which are not; a formal basis for how to determine the membership grade of a fuzzy set is absent. The membership grade is a precise, but arbitrary measure: it rests on personal opinion, not reason.

The definition of a fuzzy set extends the definition of a classical set, because membership values μ are permitted in the interval $0 \leq \mu \leq 1$, and the higher the value, the higher the membership. A classical set is consequently a special case of a fuzzy set, with membership values restricted to $\mu \in \{0, 1\}$.

A single pair $\langle x, \mu(x)\rangle$ is a fuzzy *singleton*; thus the whole set can be viewed as the union of its constituent singletons.

Example 2 *Fuzzy sets*
The following are sets which could be described by fuzzy membership functions.

(a) The set of real numbers $x \gg 1$ (x much greater than one).

(b) The set of high temperatures, the set of strong winds, or the set of nice days are fuzzy sets in weather reports.

(c) The set of young people. A one year old baby will clearly be a member of the set of young people, and a 100-year-old person will not be a member of this set. A person aged 30

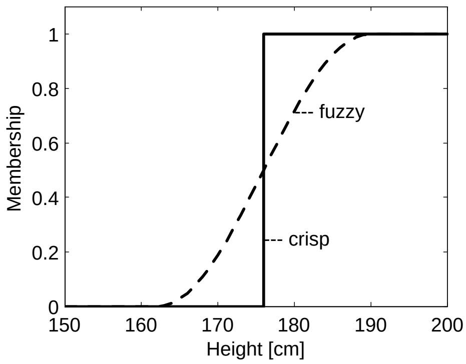

Figure 1. Two definitions of the set of "tall men", a crisp set and a fuzzy set. (figtall.m)

might be young to the degree 0.5.

(d) The set of adults. The Danish railways allow children under the age of 15 to travel at half price. An adult is thus defined by the set of passengers aged 15 or older. By their definition the set of adults is a crisp set.

(e) A predicate may be crisp, but perceived as fuzzy: a speed limit of 60 kilometres per hour is by some drivers taken to be an elastic range of more or less acceptable speeds within, say, $60 - 70$ *kilometres per hour (*$\simeq 37 - 44$ *miles per hour). Notice how, on the one hand, the traffic law is crisp while, on the other hand, those drivers's understanding of the law is fuzzy.*

Members of a fuzzy set are taken from a *universe of discourse*, or *universe*for short. The universe is all objects that can come into consideration, confer the set $\mathcal{U}$ in (1). The universe depends on the context, as the next example shows.

Example 3 *Universes*

(a) The set $x \gg 1$ *could have as a universe all real numbers, alternatively all positive integers.*

(b) The set of young people could have all human beings in the world as its universe. Alternatively it could have the numbers between 0 *and* 100*; these would then represent age in years.*

(c) The universe depends on the measuring unit; a duration in time depends on whether it is measured in hours, days, or weeks.

(d) A non-numerical quantity, for instance taste*, must be defined on a* $\mathrm{psychological\ continuum}$*; an example of such a universe is* $\mathcal{U} = \{bitter, sweet, sour, salt, hot\}$.

A programmer can exploit the universe to suppress faulty measurement data, for instance negative values for a duration of time, by making the program consult the universe.

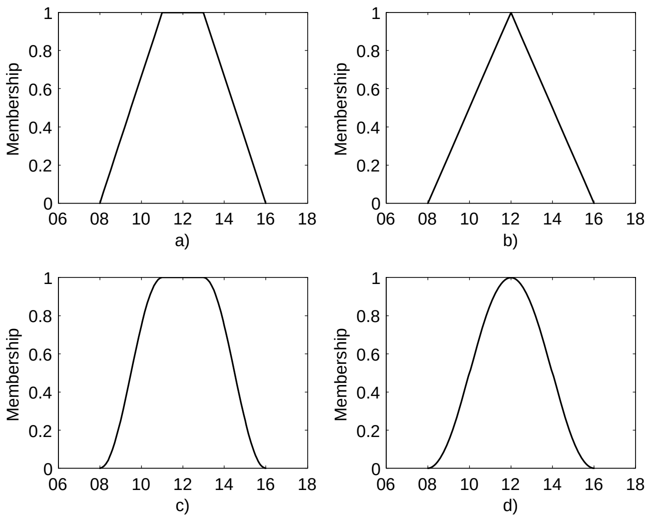

Figure 2. Around noon. Four possible membership functions representing the time 'around noon': a) trapeziodal, b) triangular, c) smooth trapezoid, and d) smooth triangular. The universe is the hours of the day in 24-hour format. (figmf0.m)

There are two alternative ways to represent a membership function: continuous or discrete. A continuous fuzzy set $\mathcal{A}$ is defined by means of a continuous membership function $\mu_{\mathcal{A}}(x)$. A *trapezoidal* membership function is a piecewise linear, continuous function, controlled by four parameters $\{a, b, c, d\}$ (Jang et al., 1997[4])

$$\mu_{trapezoid}(x; a, b, c, d) = \left\{ \begin{array}{ll} 0 & , x \leq a \\ \frac{x-a}{b-a} & , a \leq x \leq b \\ 1 & , b \leq x \leq c \\ \frac{d-x}{d-c} & , c \leq x \leq d \\ 0 & , d \leq x \end{array} \right\}, x \in \mathbb{R} \tag{2}$$

The parameters $a \leq b \leq c \leq d$ define four breakpoints, here designated: left footpoint (a), left shoulderpoint (b), right shoulderpoint (c), and right footpoint (d). Figure 2 (a) illustrates a trapezoidal membership function.

A *triangular* membership function is piecewise linear, and derived from the trapezoidal membership function by merging the two shoulderpoints into one, that is, setting $b = c$, Fig. 2 (b).

Smooth, differentiable versions of the trapezoidal and triangular membership functions can be obtained by replacing the linear segments corresponding to the intervals $a \leq x \leq b$

and $c \leq x \leq d$ by a nonlinear function, for instance a half period of a cosine function,

$$\mu_{STrapezoid}(x;a,b,c,d) = \left\{ \begin{array}{ll} 0 & ,x \leq a \\ \frac{1}{2} + \frac{1}{2}\cos(\frac{x-b}{b-a}\pi) & ,a \leq x \leq b \\ 1 & ,b \leq x \leq c \\ \frac{1}{2} + \frac{1}{2}\cos(\frac{x-c}{d-c}\pi) & ,c \leq x \leq d \\ 0 & ,d \leq x \end{array} \right\}, x \in \mathbb{R}$$

We call it here *STrapezoid* for '*smooth trapezoid*' or '*soft trapezoid*'. Figures 2 (c-d) illustrate the smooth membership functions. Other possibilities exist for generating smooth trapezoidal functions, for example Gaussian, generalized bell, and sigmoidal membership functions (Jang et al., 1997[4]).

Discrete fuzzy sets are defined by means of a discrete variable x_i $(i = 1, 2, \ldots)$. A discrete fuzzy set $\mathcal{A}$ is defined by ordered pairs,

$$\mathcal{A} = \{\langle x_1, \mu(x_1)\rangle, \langle x_2, \mu(x_2)\rangle, \ldots \mid x_i \in \mathcal{U}, i = 1, 2, \ldots\}$$

Each membership value $\mu(x_i)$ is an evaluation of the membership function μ at a discrete point x_i in the universe $\mathcal{U}$, and the whole set is a collection, usually finite, of pairs $\langle x_i, \mu(x_i)\rangle$.

Example 4 *Discrete membership function*
To achieve a discrete triangular membership function from the trapezoid (2) assume the universe is a vector $\mathbf{u}$ *of 7 elements. In Matlab notation,*

```
u=[9 10 11 12 13 14 15]
```

Assume the defining parameters are $a = 10$*,* $b = 12$*,* $c = 12$*, and* $d = 14$ *then, by (2), the corresponding membership values are a vector of* 7 *elements,*

```
0 0 0.5 1 0.5 0 0
```

Each membership value corresponds to one element of the universe, more clearly displayed as

0	0	0.5	1	0.5	0	0
9	10	11	12	13	14	15

with the universe in the bottom row, and the membership values in the top row. When this is impractical, in a program, the universe and the membership values can be kept in separate vectors.

As a crude rule of thumb, the continuous form is more computing intensive, but less storage demanding than the discrete form.

Zadeh writes that a fuzzy set induces a *possibility distribution* on the universe, meaning, one can interpret the membership values as possibilities. How are then possibilities related to probabilities? First of all, probabilities must add up to one, or the area under a density curve must be one. Memberships may add up to anything (discrete case), or the area under the membership function may be anything (continuous case). Secondly, a probability distribution concerns the likelihood of an event to occur, based on observations, whereas a possibility distribution (membership function) is subjective. The word 'probably' is synonymous with presumably, assumably, doubtless, likely, presumptively. The word 'possible' is synonymous with doable, feasible, practicable, viable, workable. Their relationship is best described in the sentence: what is probable is always possible, but not vice versa. This is illustrated next.

Example 5 *Probability versus possibility*
a) Consider the statement 'Hans ate x eggs for breakfast', where $x \in \mathcal{U} = \langle 1, 2, ..., 8 \rangle$ (Zadeh in Zimmermann, 1993[16]*). We may associate a probability distribution p by observing Hans eating breakfast for 100 days,*

.1	.8	.1	0	0	0	0	0
1	2	3	4	5	6	7	8

A fuzzy set expressing the grade of ease with which Hans can eat x eggs may be the following possibility distribution π,

1	1	1	1	.8	.6	.4	.2
1	2	3	4	5	6	7	8

Where the possibility $\pi(3) = 1$, the probability $p(3) = 0.1$ only.
b) Consider a universe of four car models

$$\mathcal{U} = \{Trabant, Fiat\ Uno, BMW, Ferrari\}\ .$$

We may associate a probability $p(x)$ of each car model driving 100 miles per hour (161 kilometres per hour) on a motorway, by observing cars for 100 days,

$$p(Trabant) = 0,\ p(Fiat\ Uno) = 0.1,\ p(BMW) = 0.4,\ p(Ferrari) = 0.5$$

The possibilities may be

$$\pi(Trabant) = 0,\ \pi(Fiat\ Uno) = 0.5,\ \pi(BMW) = 1,\ \pi(Ferrari) = 1$$

Notice that each possibility is at least as high as the corresponding probability.

Equality and inclusion are defined by means of membership functions. Two fuzzy sets $\mathcal{A}$ and $\mathcal{B}$ are equal, iff they have the same membership function,

$$\mathcal{A} = \mathcal{B} \equiv \mu_{\mathcal{A}}(x) = \mu_{\mathcal{B}}(x) \tag{3}$$

for all x. A fuzzy set $\mathcal{A}$ is a subset of (included in) a fuzzy set $\mathcal{B}$, iff the membership of $\mathcal{A}$ is less than equal to that of $\mathcal{B}$,

$$\mathcal{A} \subseteq \mathcal{B} \equiv \mu_{\mathcal{A}}(x) \leq \mu_{\mathcal{B}}(x) \tag{4}$$

for all x.

2.2 Fuzzy Set Operations

In order to generate new sets from existing sets we define two operations, in certain respects analogous to addition and multiplication. The (classical) union of the sets $\mathcal{X}$ and $\mathcal{Y}$, symbolized by $\mathcal{X} \cup \mathcal{Y}$ and read '$\mathcal{X}$ union $\mathcal{Y}$', is the set of all objects which are members of

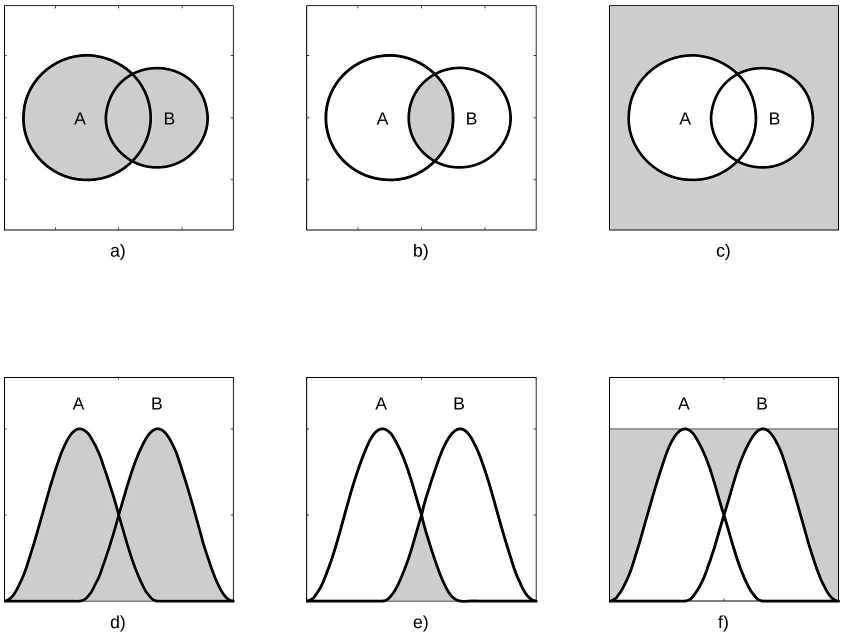

Figure 3. Set operations. The top row are classical Venn diagrams; the universe is represented by the points within the rectangle, and sets by the interior of the circles. The bottom row their fuzzy equivalents; the universal set is represented by a horisontal line at membership $\mu = 1$, and sets by membership functions. The shaded areas are: a) and d) union $A \cup B$, b) and e) intersection $A \cap B$, and c) and f) complement $\overline{A \cup B}$. (figvenn2.m)

$\mathcal{X}$ or $\mathcal{Y}$, or both. That is,

$$\mathcal{X} \cup \mathcal{Y} \equiv \{x \mid x \in \mathcal{X} \text{ or } x \in \mathcal{Y}\}$$

Thus, by definition, $x \in \mathcal{X} \cup \mathcal{Y}$ iff x is a member of at least one of $\mathcal{X}$ and $\mathcal{Y}$. For example,

$$\{1, 2, 3\} \cup \{1, 3, 4\} = \{1, 2, 3, 4\}$$

The (classical) intersection of the sets $\mathcal{X}$ and $\mathcal{Y}$, symbolized by $\mathcal{X} \cap \mathcal{Y}$ and read '$\mathcal{X}$ intersection $\mathcal{Y}$', is the set of all objects which are members of both $\mathcal{X}$ and $\mathcal{Y}$. That is,

$$\mathcal{X} \cap \mathcal{Y} \equiv \{x \mid x \in \mathcal{X} \text{ and } y \in \mathcal{Y}\}$$

For example,

$$\{1, 2, 3\} \cap \{1, 3, 4\} = \{1, 3\}$$

The (classical) complement of a set $\mathcal{X}$, symbolized by $\overline{\mathcal{X}}$ and read 'the complement of $\mathcal{X}$' is

$$\overline{\mathcal{X}} \equiv \{x \mid x \notin \mathcal{X}\}$$

That is, the set of those members of the universe which are not members $(\notin)$ of $\mathcal{X}$. Venn diagrams clearly illustrate the set operations, Fig. 3 (a-c).

When turning to fuzzy sets, the gradual membership complicates matters. Figure 3 (d-f) shows an intuitively acceptable modification of the Venn diagrams. The following fuzzy set operations are defined accordingly:

Let $\mathcal{A}$ and $\mathcal{B}$ be fuzzy sets defined on a mutual universe $\mathcal{U}$. The fuzzy union of $\mathcal{A}$ and $\mathcal{B}$ is

$$\mathcal{A} \cup \mathcal{B} \equiv \{\langle x, \mu_{\mathcal{A}\cup\mathcal{B}}(x)\rangle \mid x \in \mathcal{U} \text{ and } \mu_{\mathcal{A}\cup\mathcal{B}}(x) = \max(\mu_{\mathcal{A}}(x), \mu_{\mathcal{B}}(x))\}$$

The fuzzy intersection of $\mathcal{A}$ and $\mathcal{B}$ is

$$\mathcal{A} \cap \mathcal{B} \equiv \{\langle x, \mu_{\mathcal{A}\cap\mathcal{B}}(x)\rangle \mid x \in \mathcal{U} \text{ and } \mu_{\mathcal{A}\cap\mathcal{B}}(x) = \min(\mu_{\mathcal{A}}(x), \mu_{\mathcal{B}}(x))\}$$

The fuzzy complement of $\mathcal{A}$ is

$$\overline{\mathcal{A}} \equiv \{\langle x, \mu_{\overline{\mathcal{A}}}(x)\rangle \mid x \in \mathcal{U} \text{ and } \mu_{\overline{\mathcal{A}}}(x) = 1 - \mu_{\mathcal{A}}(x)\}$$

While the notation may look cumbersome, it is in practice easy to apply the fuzzy set operations $\max$, $\min$, and $1-\mu$.

Example 6 *Buying a house (after Zimmermann, 1993*[16]*)*
A four-person family wishes to buy a house. An indication of their level of comfort is the number of bedrooms in the house. But they also wish a large house. The universe $\mathcal{U} = \langle 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\rangle$ *is the set of houses to be considered described by their number of bedrooms. The fuzzy set Comfortable may be described as a vector* $\mathbf{c}$, *or in Matlab*

$$\texttt{c} = \begin{bmatrix} 0.2 & 0.5 & 0.8 & 1 & 0.7 & 0.3 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Let $\mathbf{l}$ *describe the fuzzy set Large, defined as*

$$\texttt{l} = \begin{bmatrix} 0 & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The intersection of Comfortable and Large is then `min(c,l)`,

$$0 \quad 0 \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.3 \quad 0 \quad 0 \quad 0 \quad 0$$

To interpret, five bedrooms is optimal having the largest membership value 0.6. *It is, however, not fully satisfactory, since the membership is less than* 1. *The second best solution is four bedrooms, membership* 0.4. *If the market is a buyer's market, the family will probably wait until a better offer comes up, thus hoping to achieve full satisfaction (membership 1).*

The union of Comfortable and Large is `max(c,l)`

$$0.2 \quad 0.5 \quad 0.8 \quad 1 \quad 0.7 \quad 0.8 \quad 1 \quad 1 \quad 1 \quad 1$$

Here four bedrooms is fully satisfactory (membership 1) because it is comfortable. Also 7-10 bedrooms are satisfactory, because the house is large. If the market is a seller's market, the family might buy the house, being content that at least one of the criteria is fulfilled.

If the children are about to move away from the family within the next couple of years, the parents may wish to buy a house that is Comfortable and Not Large, or `min(c, 1-l)`

$$0.2 \quad 0.5 \quad 0.8 \quad 0.6 \quad 0.4 \quad 0.2 \quad 0 \quad 0 \quad 0 \quad 0$$

In that case, three bedrooms is satisfactory to the degree 0.8. *The example indicates how fuzzy sets can be used for computer aided decision support.*

In mathematics the word 'relation' is used in the sense of relationship, for example the predicates: x is less than y, or y is a function of x. A *binary relation* $\mathcal{R}$ is a set of ordered pairs. We may write it $x\mathcal{R}y$ which reads: 'x is related to y.' There are established symbols for various relations, for example $x = y$, $x < y$. One simple relation is the set of all pairs $\langle x, y \rangle$, such that x is a member of a set $\mathcal{X}$ and y is a member of a set $\mathcal{Y}$. This is the (classical) *cartesian product* of $\mathcal{X}$ and $\mathcal{Y}$,

$$\mathcal{X} \times \mathcal{Y} = \{\langle x, y \rangle \mid x \in \mathcal{X}, y \in \mathcal{Y}\}$$

In fact, any binary relation $x\mathcal{R}y$ is a subset of the cartesian product $\mathcal{X} \times \mathcal{Y}$, and we can think of those instances of $\mathcal{X} \times \mathcal{Y}$, that are members of $\mathcal{R}$ as having membership 1.

By analogy, a binary *fuzzy relation* consists of pairs $\langle x, y \rangle$ with an associated fuzzy membership value. For example, given $\mathcal{X} = \mathcal{Y} = \{1, 2, 3\}$ we can set up a relation 'approximately equal' between all pairs of the three numbers, most clearly displayed in a tabular arrangement,

$\mathcal{X}$ \ $\mathcal{Y}$	1	2	3
1	1	0.8	0.3
2	0.8	1	0.8
3	0.3	0.8	1

In the fuzzy cartesian product each object is defined by a pair: the object, which is a pair itself, and its membership. Let $\mathcal{A}$ and $\mathcal{B}$ be fuzzy sets defined on $\mathcal{X}$ and $\mathcal{Y}$ respectively, then the cartesian product $\mathcal{A} \times \mathcal{B}$ is a fuzzy set in $\mathcal{X} \times \mathcal{Y}$ with the membership function

$$\mathcal{A} \times \mathcal{B} = \left\{ \left\langle \langle x, y \rangle, \mu_{\mathcal{A}\times\mathcal{B}}(x, y) \right\rangle \mid x \in \mathcal{X}, y \in \mathcal{Y}, \mu_{\mathcal{A}\times\mathcal{B}}(x, y) = \min\left(\mu_{\mathcal{A}}(x), \mu_{\mathcal{B}}(y)\right) \right\}$$

For example, assume $\mathcal{X}$ and $\mathcal{Y}$ are as above, and $\mu_{\mathcal{A}}(x_i) = \langle 0, 0.5, 1 \rangle$, with $i = 1, 2, 3$, and $\mu_{\mathcal{B}}(y_j) = \langle 1, 0.5, 0 \rangle$, with $j = 1, 2, 3$, then $\mathcal{A} \times \mathcal{B}$ is a two-dimensional fuzzy set

$\mathcal{A}$ \ $\mathcal{B}$	1	0.5	0
0	0	0	0
0.5	0.5	0.5	0
1	1	0.5	0

The element at row i and column j is the intersection of $\mu_{\mathcal{A}}(x_i)$ and $\mu_{\mathcal{B}}(y_j)$. Again we note that to each object $\langle x_i, y_j \rangle$ is associated a membership $\mu_{\mathcal{A}\times\mathcal{B}}(x_i, y_j)$, whereas the classical cartesian product consists of objects $\langle x_i, y_j \rangle$ only.

In order to see how to combine relations, let us look at an example from the cartoon Donald Duck.

Example 7 *Donald Duck's family*
Assume that Donald Duck's nephew Huey resembles nephew Dewey to the grade 0.8, and Huey resembles nephew Louie to the grade 0.9. We have therefore a relation between two subsets of the nephews in the family. This is conveniently represented in a matrix, with one row and two columns (and additional headings),

$\mathbf{R}_1 =$

	Dewey	*Louie*
Huey	*0.8*	*0.9*

Let us assume another relation between nephews Dewey and Louie on the one side, and uncle Donald on the other, a matrix with two rows and one column,

$$\mathbf{R}_2 = \begin{array}{c|c|} & \textit{Donald} \\ \cline{2-2} \textit{Dewey} & \textit{0.5} \\ \cline{2-2} \textit{Louie} & \textit{0.6} \\ \cline{2-2} \end{array}$$

We wish to find out how much Huey resembles Donald by combining the information in the two matrices. Observe that

(i) Huey resembles Dewey ($\mathbf{R}_1(1,1) = 0.8$*), and Dewey in turn resembles Donald (*$\mathbf{R}_2(1,1) = 0.5$*), or*

(ii) Huey resembles Louie ($\mathbf{R}_1(1,2) = 0.9$*), and Louie in turn resembles Donald (*$\mathbf{R}_2(2,1) = 0.6$*).*

Assertion (i) contains two relationships combined by 'and'; it seems reasonable to take the intersection. With our definition, this corresponds to choosing the smallest membership value for the (transitive) Huey-Donald relationship, or $\min(0.8, 0.5) = 0.5$*. Similarly with statement (ii). Thus from chains (i) and (ii), respectively, we deduce that*

(iii) Huey resembles Donald to the degree 0.5*, or*

(iv) Huey resembles Donald to the degree 0.6*.*

Although the results in (iii) and (iv) differ, we are equally confident in either result; we have to choose either one or the other, so it seems reasonable to take the union. With our definition, this corresponds to choosing the largest membership value, or $\max(0.5, 0.6) = 0.6$*. Thus, the answer is that Huey resembles Donald to the degree* 0.6*.*

Generally speaking, this was an example of *composition* of relations. Let $\mathcal{R}$ and $\mathcal{S}$ be two fuzzy relations defined on $\mathcal{X} \times \mathcal{Y}$ and $\mathcal{Y} \times \mathcal{Z}$ respectively. Their composition is a fuzzy set defined by

$$\mathcal{R} \circ \mathcal{S} = \left\{ \left\langle \langle x, z \rangle, \bigcup_y \mu_{\mathcal{R}}(x, y) \cap \mu_{\mathcal{S}}(y, z) \right\rangle \mid x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z} \right\}$$

When $\mathcal{R}$ and $\mathcal{S}$ are expressed as matrices $\mathbf{R}$ and $\mathbf{S}$, the composition is equivalent to an *inner product*. The inner product is similar to an ordinary matrix (dot) product, except multiplication is replaced by any function and summation by any function. Suppose $\mathbf{R}$ is an $m \times\ p$ matrix and $\mathbf{S}$ is a $p \times\ n$ matrix. Then the inner $\cup - \cap$ product (read 'cup-cap product') is an $m\ \times\ n$ matrix $\mathbf{T} = (t_{ij})$ whose ij-entry is obtained by combining the *i*th row of $\mathbf{R}$ with the *j*th column of $\mathbf{S}$, such that

$$t_{ij} = (r_{i1} \cap s_{1j}) \cup (r_{i2} \cap s_{2j}) \cup \ldots \cup (r_{ip} \cap s_{pj}) = \bigcup_{k=1}^{p} r_{ik} \cap s_{kj} \tag{5}$$

With our definitions of the set operations, the composition is specifically called *max-min composition* (Zadeh in Zimmermann, 1993[16]). Sometimes the min operation is replaced by * for multiplication, then it is called *max-star composition*.

Example 8 *Inner product*

For the tables $\mathbf{R}_1$ *and* $\mathbf{R}_2$ *above, the inner product yields,*

$$\mathbf{R}_1 \circ \mathbf{R}_2 = \boxed{0.8 \mid 0.9} \circ \begin{bmatrix} 0.5 \\ 0.6 \end{bmatrix} = (0.8 \cap 0.5) \cup (0.9 \cap 0.6) = 0.5 \cup 0.6 = 0.6$$

which agrees with the previous result.

3 Fuzzy Logic

Logic started as the study of language in arguments and persuasion, and it can be used to judge the correctness of a chain of reasoning — in a mathematical proof for instance. The goal of the theory is to reduce principles of reasoning to a code. The 'truth' or 'falsity' assigned to a proposition is its *truth-value*. In *fuzzy logic* a proposition may be true or false, or an intermediate truth-value such as *maybe true*. The sentence 'John is a tall man' is an example of a fuzzy proposition. For convenience we shall here restrict the possible truth-values to a discrete domain $\{0, 0.5, 1\}$ for *false*, *maybe true*, and *true*. In doing so we are in effect restricting the theory to *multi-valued* logic, or rather three-valued logic to be specific. In practice one would subdivide the unit interval into finer divisions, or work with a continous truth-domain. Nevertheless, much of what follows is valid even in a continuous domain as we shall see.

3.1 Propositions

In daily conversation and mathematics, sentences are connected with the words *and*, *or*, *if-then* (or *implies)*, and *if and only if*. These are called *connectives*. A sentence which is modified by the word 'not' is called the *negation* of the original sentence.The word 'and' is used to join two sentences to form the *conjunction* of the two sentences. The word 'or' is used to join two sentences to form the *disjunction* of the two sentences. From two sentences we may construct one, of the form 'If ... then ...'; this is called an *implication*. The sentence following 'If' is the *antecedent*, and the sentence following 'then' is the *consequent*. Other idioms which we shall regard as having the same meaning are 'p implies q', 'p only if q', 'q if p', etc.

Letters and special symbols make the connective structure stand out. Our choice of symbols is

$\neg$	for 'not'
$\wedge$	for 'and'
$\vee$	for 'or'
$\Rightarrow$	for 'if-then' (implication)
$\Leftrightarrow$	for 'if and only if' (equivalence)

The next example illustrates how the symbolic forms expose the underlying logical structure.

Example 9 *Baseball betting (Stoll, 1979*[10]*)*

Consider the assertion about four baseball teams: If either the Pirates or the Cubs loose and the Giants win, then the Dodgers will be out of first place, and I will loose a bet. Since it is an implication, it may be symbolised in the form $r \Rightarrow s$. The antecedent is composed from the three sentences p *(The Pirates lose),* c *(The Cubs lose), and* g *(The Giants win). The consequent is the conjunction of* d *(The Dodgers will be out of first place) and* b *(I will lose a bet). The original sentence may thus be represented by* $((p \vee c) \wedge g) \Rightarrow (d \wedge b)$.

An assertion which contains at least one propositional variable is called a *propositional form*. The main difference between *proposition* and propositional *form* is that every proposition has a truth-value, whereas a propositional form is an assertion whose truth-value cannot be determined until propositions are substituted for its propositional variables. But when no confusion results, we will refer to propositional forms as propositions.

A *truth-table* summarises the possible truth-values of an assertion. Take for example the truth-table for the two-valued propositional form $p \vee q$. The truth-table (below, left) lists all possible combinations of truth-values — the Cartesian product — of the arguments p and q in the two leftmost columns. The rightmost column holds the truth-values of the proposition. Alternatively, the truth-table can be rearranged into a two-dimensional array, a so-called Cayley table (below, right).

p	q	$p \vee q$
0	0	0
0	1	1
1	0	1
1	1	1

is equivalent to

Or

$$\begin{array}{c|c|c|l} \multicolumn{4}{c}{p \vee q} \\ & 0 & 1 & \rightarrow q \\ \hline 0 & 0 & 1 & \\ \hline 1 & 1 & 1 & \\ \hline \downarrow & & & \\ p & & & \end{array}$$

Along the vertical axis in the Cayley table, symbolized by arrow $\downarrow$, are the possible values 0 and 1 of the first argument p. Along the horizontal axis, symbolized by arrow $\rightarrow$, are the possible values 0 and 1 of the second argument q. Above the table, the proposition $p \vee q$ reminds us that the table concerns disjunction. At the intersection of row i and column j (only counting the inside of the box) is the truth-value of the expression $p_i \vee q_j$. By inspection, one entry renders $p \vee q$ false, while three entries render $p \vee q$ true. Truth-tables for binary connectives are thus given by two-by-two matrices. A total of 16 such tables can be constructed, and each has been associated with a connective.

We can derive the truth-table for 'nand' ('not and') from 'or'. By the definition $(\neg p) \vee (\neg q)$ we negate each variable of the previous truth-table, which is equivalent to reversing the axes and permuting the entries back in the usual ascending order on the axes,

Nand

$$\begin{array}{c|c|c|l} \multicolumn{4}{c}{(\neg p) \vee (\neg q)} \\ & 0 & 1 & \rightarrow q \\ \hline 0 & 1 & 1 & \\ \hline 1 & 1 & 0 & \\ \hline \downarrow & & & \\ p & & & \end{array}$$

From ’nand’ (’not and’) we derive ’and’ by negating the entries of the table (’not not and’ = ’and’),

And

$\neg\left((\neg p) \vee (\neg q)\right)$

	0	1	$\rightarrow q$
0	0	0	
1	0	1	
$\downarrow p$			

Notice that it was possible to define ’and’ by means of ’or’ and ’not’, rather than assume its definition given as an axiom.

By means of ’or’ and ’not’ we can proceed to define ’implication’. Classical logic defines implication $\neg p \vee q$, which is called *material implication*. We negate the p-axis of the ’or’ table, which is equivalent to reversing the axis and permuting the entries back in the usual ascending order, thus

Implication

$\neg p \vee q$

	0	1	$\rightarrow q$
0	1	1	
1	0	1	
$\downarrow p$			

Equivalence is taken to mean $(p \Rightarrow q) \wedge (q \Rightarrow p)$, which is equivalent to the conjunction of each entry of the implication table with the elements of the transposed table, element by element, or

Equivalence

$(p \Rightarrow q) \wedge (q \Rightarrow p)$

	0	1	$\rightarrow q$
0	1	0	
1	0	1	
$\downarrow p$			

It is possible to evaluate, in principle at least, a logical proposition by an exhaustive test of all combinations of truth-values of the variables, and this is the idea behind *array based logic* (Franksen, 1979[1]). The next example illustrates an application of array based logic.

Example 10 *Array based logic*

In the baseball example, we had the relation $((p \vee c) \wedge g) \Rightarrow (d \wedge b)$. *The proposition contains five variables, and each variable can take two truth-values. There are therefore* $2^5 = 32$ *possible combinations. Only* 23 *are legal, in the sense that the proposition is true for these combinations, and* $32 - 23 = 9$ *cases are illegal, in the sense that the proposition is false for those combinations. If we are interested only in the legal combinations for which*

'I win the bet' $(b = 0)$*, then the following table results*

p	c	g	d	b
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
1	0	0	0	0
1	0	0	1	0
1	1	0	0	0
1	1	0	1	0

There are thus 10 *winning outcomes out of* 32 *possible.*

By analogy we can define similar truth-tables for fuzzy logic connectives. If we start again by defining negation and disjunction, we can derive the truth-tables of other connectives from that point of departure. Let us define *disjunction* as set union, that is,

$$p \vee q \equiv \max(p, q). \tag{6}$$

We can then build the truth-table for the fuzzy connective 'or',

Or

$p \vee q$	0	0.5	1 →q
0	0	0.5	1
0.5	0.5	0.5	1
1	1	1	1
↓ p			

Like before, the p-axis is vertical and the q-axis horizontal. At the intersection of row i and column j (only regarding the inside of the box) is the value of the expression $\max(p_i, q_j)$ in accordance with (6). When looking for definitions of fuzzy connectives, we will require that such connectives should agree with their classical counterparts for the truth-domain $\{0, 1\}$. In terms of truth-tables, the values in the four corners of the fuzzy Cayley table, should agree with the Cayley table for the classical connective.

Let us define *negation* as set complement, that is,

$$\neg p \equiv 1 - p.$$

The truth-table for 'nand' is derived from 'or' by the definition $(\neg p) \vee (\neg q)$ by negating the variables. That is equivalent to reversing the axes in the Cayley table. And moving further, 'and' is the negation of the entries in the truth-table for 'nand', thus

Nand $(\neg p) \vee (\neg q)$	0	0.5	1 → q
0	1	1	1
0.5	1	0.5	0.5
1	1	0.5	0
↓ p			

And $\neg((\neg p) \vee (\neg q))$	0	0.5	1 → q
0	0	0	0
0.5	0	0.5	0.5
1	0	0.5	1
↓ p			

(7)

It is reassuring to observe that even though the truth-table for 'and' is derived from the truth-table for 'or', the truth-table for 'and' is identical to one generated using the $\min$ operation, set intersection.

The *implication*, however, has always troubled fuzzy logicians. If we define it as *material implication*, $\neg p \vee q$, then we get a fuzzy truth-table which is unsuitable, as it causes several useful logical laws to break down. It is important to realise, that we must make a design choice at this point, in order to proceed with the definition of implication and equivalence. The choice is which logical laws we wish to apply.

Not all laws known from two-valued logic can be valid in fuzzy logic. Take for instance the propositional form

$$p \vee \neg p \Leftrightarrow 1 \tag{8}$$

which is equivalent to the law of the excluded middle. Testing with the truth-value $p = 0.5$ (fuzzy logic) the left hand side yields

$$0.5 \vee \neg 0.5 = \max(0.5, 1 - 0.5) = 0.5.$$

This is different from the right hand side, and thus the law of the excluded middle is invalid in fuzzy logic. If a proposition is true with a truth-value of 1, for any combination of truth-values assigned to the variables, we shall say it is *valid*. Such a proposition is a *tautology*. If the proposition is true for some, but not all combinations, we shall say it is *satisfiable*. Thus (8) is satisfiable, since it is true in two-valued logic, but not in three-valued logic.

One tautology that we definitely wish to apply in fuzzy logic applications is

$$\text{Tautology 1:} \quad [p \wedge (p \Rightarrow q)] \Rightarrow q \tag{9}$$

Or in words: If p, and p implies q, then q. We have labelled it tautology 1, because we need it later in connection with the *modus ponens* rule of inference. Another tautology that we definitely wish to apply in fuzzy logic applications is the transitive relationship

$$\text{Tautology 2:} \quad [(p \Rightarrow q) \wedge (q \Rightarrow r)] \Rightarrow (p \Rightarrow r) \tag{10}$$

Or in words from left to right: if p implies q which in turn implies r, then p implies r. Whether these propositions are valid in fuzzy logic depends on how we define the connectives. Or rather, we must define the connectives, implication in particular, such that those propositions become valid (Jantzen, 1995[5]).

Many researchers have proposed definitions for the implication connective (e.g., Zadeh, 1973; Mizumoto, Fukami & Tanaka, 1979; Fukami, Mizumoto & Tanaka, 1980; Wenstøp, 1980[13, 9, 2, 11]; see also the survey by Lee, 1990[7]). In fact Kiszka, Kochanska &

Sliwinska list 72 alternatives to choose from (1985[6]). One candidate, not necessarily the best, is the so-called *sharp implication*. It can be defined

$$p \Rightarrow q \equiv p \leq q \tag{11}$$

This choice is motivated by the following argument. When the crisp set $\mathcal{X}$ is included in the crisp set $\mathcal{Y}$, then $\mathcal{X}$ is said to imply $\mathcal{Y}$. The fuzzy set $\mathcal{A}$ is included in the fuzzy set $\mathcal{B}$, iff $\mu_{\mathcal{A}}(x) \leq \mu_{\mathcal{B}}(x)$ by (4), therefore we define implication by (11).

Once it is agreed that $p \Leftrightarrow q$ is the same as $(p \Rightarrow q) \wedge (q \Rightarrow p)$, the truth-table for equivalence ($\Leftrightarrow$) is determined from implication and conjunction,

Implication $p \leq q$

p ↓ \ →q	0	0.5	1
0	1	1	1
0.5	0	1	1
1	0	0	1

Equivalence $(p \Rightarrow q) \wedge (q \Rightarrow p)$

p ↓ \ →q	0	0.5	1
0	1	0	0
0.5	0	1	0
1	0	0	1

(12)

Example 11 *Fuzzy baseball*
The baseball example illustrates what difference three-valued logic makes. The proposition

$$((p \vee c) \wedge g) \Rightarrow (d \wedge b)$$

contains five variables, and now each can take three truth-values. This implies $3^5 = 243$ *possible combinations;* 148 *of these are legal in the sense that the proposition is true (truth-value* 1*). If we are interested again in the legal combinations for which 'I win the bet'* $(b = 0)$*, then there are* 33 *winning outcomes out of* 148*. Instead of listing all, we show one for illustration,*

$$(p, c, g, d, b) = (0.5, 0.5, 0, 1, 0) .$$

With two-valued logic we found 10 *winning outcomes out of* 32 *possible.*

The example indicates that fuzzy logic provides more solutions, compared to two-valued logic, and it requires more computational effort.

It is straight forward to test whether tautologies 1 and 2 are valid, because it is possible to perform an exhaustive test of all combinations of truth-values of the variables, provided the domain of truth-values is discrete and limited.

Example 12 *Testing tautology 1*
Tautology 1 is,

$$[p \wedge (p \Rightarrow q)] \Rightarrow q$$

Since the proposition contains two variables p *and* q*, and each variable can take three truth-*

values, there will be $3^2 = 9$ *possible combinations to test. The truth-table has 9 rows,*

p	q	$p \Rightarrow q$	$[p \wedge (p \Rightarrow q)]$	$[p \wedge (p \Rightarrow q)] \Rightarrow q$
0	0	1	0	1
0	0.5	1	0	1
0	1	1	0	1
0.5	0	0	0	1
0.5	0.5	1	0.5	1
0.5	1	1	0.5	1
1	0	0	0	1
1	0.5	0	0	1
1	1	1	1	1

Columns 1 and 2 are the input combinations of p *and* q*. Column 3 is the result of sharp implication, and column 4 is the left hand side of tautology 1. Since the rightmost column is all* 1*'s, the proposition is a tautology, even in three-valued logic.*

Example 12 suggests a new tautology, in fact. If we study closely the truth-values for $[p \wedge (p \Rightarrow q)]$ (column 4 in the example), and compare them with the truth-table for conjunction, we discover they are identical. We can thus postulate,

$$\text{Tautology 3:} \quad [p \wedge (p \Rightarrow q)] \Leftrightarrow p \wedge q$$

We shall use this result later in connection with fuzzy inference.

The main objective of the approach taken here, is not to show that one implication is better than another, but to reduce the proof of any tautology to a test that can be programmed on a computer. The scope is so far limited to the chosen truth domain $\{0, 0.5, 1\}$; this could be extended with intermediate values, however, and the test performed again in case a higher resolution is required. Nevertheless, it suffices to check for all possible combinations of $\{0, 0.5, 1\}$ if we wish to check the equality of two propositions involving variables connected with $\wedge, \vee$, and $\neg$ (Gehrke, Walker & Walker, 2003[3]). For propositions involving sharp implication, which cannot be transcribed into a definition involving the said three connectives only, we have to conjecture the results for fuzzy logic.

Since implication can be defined in many possible ways, one has to determine a design criterion first, namely the tautologies, before choosing a proper definition for the implication connective.

Originally, Zadeh interpreted a truth-value in fuzzy logic as a fuzzy set, for instance *Very true* (Zadeh, 1988[15]). Thus Zadeh based fuzzy (linguistic) logic on treating *Truth* as a linguistic variable that takes words or sentences as values rather than numbers (Zadeh, 1975[14]). Please be aware that our approach differs, as it is built on scalar truth-values rather than vectors.

Example 13 *Mamdani implication*
The so-called Mamdani implication, *is often used in fuzzy control (Mamdani, 1977*[8]*). Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be fuzzy sets defined on* $\mathcal{X}$ *and* $\mathcal{Y}$ *respectively, then the Mamdani implication is a*

fuzzy set in $\mathcal{X} \times \mathcal{Y}$ with the membership function

$$\{\langle\langle x, y\rangle, \mu_{\mathcal{A}' \Rightarrow' \mathcal{B}}(x, y)\rangle \mid x \in \mathcal{X}, y \in \mathcal{Y}, \mu_{\mathcal{A}' \Rightarrow' \mathcal{B}}(x, y) = \min\left(\mu_{\mathcal{A}}(x), \mu_{\mathcal{B}}(y)\right)\}$$

Notice the definition is similar to the definition of the fuzzy cartesian product. Its Cayley table, using the minimum operation, is

Mamdani 'implication'

$\min(p, q)$

$p \downarrow$ \ $q \rightarrow$	0	0.5	1
0	0	0	0
0.5	0	0.5	0.5
1	0	0.5	1

One discovers by inspection that only one out of the four corner-values matches the truth-table for two-valued implication. Therefore, it is not an implication, and Mamdani 'inference' would be a more appropriate name, as we shall see later.

3.2 Inference

Logic provides principles of reasoning, by means of *inference*, the drawing of conclusions from assertions. The verb 'to infer' means to conclude from evidence, deduce, or to have as a logical consequence (do not confuse 'inference' with 'in*ter*ference'). *Rules of inference* specify conclusions drawn from assertions known or assumed to be true.

One such rule of inference is *modus ponens*. It is often presented in the form of an *argument*:

$$\begin{array}{l} P \\ P \Rightarrow Q \\ \hline Q \end{array}$$

In words: If 1) P is known to be true, and 2) we assume that $P \Rightarrow Q$ is true, then 3) Q must be true. Restricting for a moment to two-valued logic, we see from the truth-table for implication,

Implication

$p \Rightarrow q$

$p \downarrow$ \ $q \rightarrow$	0	1
0	1	1
1	0	1

, (13)

whenever $P \Rightarrow Q$ and P are true then so is Q; by P true we consider only the second row, leaving Q true as the only possibility. In such an argument the assertions above the line are the *premises*, and the assertion below the line the *conclusion*. Notice the premises are assumed to be true, we are not considering *all* possible truth combinations as we did when proving tautologies.

On the other hand, underlying modus ponens is tautology 1, which expresses the same, but is valid for *all* truth-values. Therefore modus ponens is valid in fuzzy logic, if tautology 1 is valid in fuzzy logic.

Example 14 *Four useful rules of inference*
There are several useful rules of inference, which can be represented by tautological forms. Four are presented here by examples from medicine.

(a) Modus ponens. *Its tautological form is*

$$[p \wedge (p \Rightarrow q)] \Rightarrow q.$$

Let p stand for 'altitude sickness,' and let $p \Rightarrow q$ stand for 'altitude sickness causes a headache'. If it is known that John suffers from altitude sickness, p is true, and $p \Rightarrow q$ is assumed to be true for the case of illustration, then the conclusion q is true, that is, John has a headache.

(b) Modus tollens. *Its tautological form is*

$$[\neg q \wedge (p \Rightarrow q)] \Rightarrow \neg p.$$

Let p and q be as in (a). Thus, if John does not have a headache, then we may infer that John does not suffer from altitude sickness.

(c) Disjunctive syllogism. *Its tautological form is*

$$[(p \vee q) \wedge \neg p] \Rightarrow q.$$

Let p stand for 'altitude sickness' as previously, but let q stand for 'dehydration'. Thus, if it is known for a fact that John's headache is due to either altitude sickness or dehydration, and it is not altitude sickness, then we may infer that John suffers from dehydration.

(d) Hypothetical syllogism. *Its tautological form is tautology 2:*

$$[(p \Rightarrow q) \wedge (q \Rightarrow r)] \Rightarrow (p \Rightarrow r)\,.$$

Let p stand for 'high altitude and fast ascent', let q stand for 'altitude sickness', and let r stand for 'a headache'. Further assume that high altitude and fast ascent together cause altitude sickness, and in turn that altitude sickness causes a headache. Thus, we may infer that John will get a headache in high altitude if John ascends fast.

Provided the tautological forms are valid in fuzzy logic, the inference rules may be applied in fuzzy logic as well.

The inference mechanism in modus ponens can be generalised. The pattern is: given a relation $\mathcal{R}$ connecting logical variables p and q, we infer the possible values of q given a particular instance of p. Switching to vector-matrix representation, to emphasize the computer implementation, with $\mathbf{p}$ a (column) vector and $\mathbf{R}$ a two-dimensional truth-table, with the p-axis vertical, the inference is defined

$$\mathbf{q}^t = \mathbf{p}^t \circ \mathbf{R}$$

The operation $\circ$ is an inner $\vee - \wedge$ product. The $\wedge$ operation is the same as in $p \wedge (p \Rightarrow q)$ and the $\vee$ operation along the columns yields what can possibly be implied about q, confer the rightmost implication in $[p \wedge (p \Rightarrow q)] \Rightarrow p$. Assuming p is true corresponds to setting

$$\mathbf{p} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

But the scheme is more general, because we could also assume $\mathbf{p}$ is false, compose with $\mathbf{R}$ and study what can be inferred about $\mathbf{q}$. Take for instance modus ponens, thus

$$\mathbf{R} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

which is the truth-table for $p \Rightarrow q$. Assigning $\mathbf{p}$ as above,

$$\mathbf{q}^t = \mathbf{p}^t \circ \mathbf{R} = \begin{pmatrix} 0 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

The outcome $\mathbf{q}^t$ is a truth-vector pointing at q true as the only possible conclusion, as expected. Testing for instance with $\mathbf{p} = (1\ 0)^t$ yields

$$\mathbf{q}^t = \mathbf{p}^t \circ \mathbf{R} = \begin{pmatrix} 1 & 0 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

Thus q could be anything, true or false, as expected. The inference could even proceed in the reverse direction, from q to p, but then we must compose from the right side of $\mathbf{R}$ to match the axes. Assume for instance q is true, or $\mathbf{q} = (1\ 0)^t$, then

$$\mathbf{p} = \mathbf{R} \circ \mathbf{q} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

To interpret: if q is false and $p \Rightarrow q$, then p is false (modus tollens).

The array based inference mechanism is even more general, because $\mathbf{R}$ can be any dimension n ($n > 0$ and integer). Given values of $n - 1$ variables, the possible outcomes of the remaining variable can be inferred by an n-dimensional inner product. Furthermore, given values of $n - d$ variables (d integer and $0 < d < n$), then the truth-array connecting the remaining d variables can be inferred.

Generalising to three-valued logic $\mathbf{p}$ and $\mathbf{q}$ are vectors of three elements, $\mathbf{R}$ a 3-by-3 matrix, and the inner $\vee - \wedge$ product interpreted as the inner max-min product. Assuming p is true, corresponds to assigning

$$\mathbf{p} = \begin{pmatrix} 0 \\ 0.5 \\ 1 \end{pmatrix}$$

In modus ponens

$$\mathbf{R} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

which is the truth-table for sharp implication. With $\mathbf{p}$ as above,

$$\mathbf{q}^t = \mathbf{p}^{\mathbf{t}} \circ \mathbf{R} = \begin{pmatrix} 0 & .5 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & .5 & 1 \end{pmatrix}$$

To interpret, p true and $p \Rightarrow q$ true, implies q true, as expected because tautology 1 is valid in three valued logic.

The inner product $\mathbf{p^t} \circ \mathbf{R}$ can be decomposed into three operations.

1. A cartesian product

$$\mathbf{R}^1 = \mathbf{p} \times \mathbf{1}^t = \begin{pmatrix} 0 \\ 0.5 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0.5 & 0.5 & 0.5 \\ 1 & 1 & 1 \end{pmatrix}$$

called the *cylindrical extension*. The result is a matrix whose columns are the vector $\mathbf{p}$ repeated as many times as necessary to fit the size of $\mathbf{R}$.

2. An element-wise intersection

$$\mathbf{R}^2 = \mathbf{R}^1 \wedge \mathbf{R} = \begin{pmatrix} 0 & 0 & 0 \\ 0.5 & 0.5 & 0.5 \\ 1 & 1 & 1 \end{pmatrix} \wedge \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0 & 1 \end{pmatrix}$$

This operation does not have a name in the literature.

3. An $\vee$-operation along the columns of $\mathbf{R}^2$,

$$\mathbf{q}^t = \bigvee_p r^2_{pq} = \begin{pmatrix} 0 & 0.5 & 1 \end{pmatrix}$$

which is the projection of $\mathbf{R}$ on the q-axis.

In general, the max-min composition

$$\mathbf{b}^t = \mathbf{a}^t \circ \mathbf{R}$$

consists of three operations: 1) the cylindrical extension $\mathbf{R}^1 = \mathbf{a}^t \times 1_b$, 2) the elementwise intersection $\mathbf{R}^2 = r^1_{ab} \wedge r_{ab}$, and 3) the projection onto the b-axis $\mathbf{b}^t = \bigvee_b r^2_{ab}$. This constitutes the *compositional rule of inference*, where $\mathbf{b}$ is inferred from $\mathbf{a}$ by means of $\mathbf{R}$.

4 Fuzzy Rules

A fuzzy rule has the form

If x is $\mathcal{A}$ then y is $\mathcal{B}$,

in which $\mathcal{A}$ and $\mathcal{B}$ are fuzzy sets, defined on universes $\mathcal{X}$ and $\mathcal{Y}$, respectively. This is an implication, where the antecedent is 'x is $\mathcal{A}$', and the consequent is 'y is $\mathcal{B}$'. Examples of such rules in everyday conversation are

1. If it is dark, then drive slowly
2. If the tomato is red, then it is ripe
3. If it is early, then John can study
4. If the room is cold, then increase the heat
5. If the washing machine is half full, then wash shorter time

Other forms can be transcribed into the if-then form, for example 'when in Rome, do like the Romans' becomes 'if in Rome, then do like the Romans'. Examples 4 and 5 could appear in a computer embedded in a heating unit or a washing machine.

To understand how a computer can infer a conclusion, take rule 3,

If it is early, then John can study

Assume that 'early' is a fuzzy set defined on the universe

$$\mathcal{U} = \langle 4, 8, 12, 16, 20, 24 \rangle .$$

These are times t of the day, in steps of four hours, in 24 hour format to distinguish night from day numerically. Define 'early' as a fuzzy set on $\mathcal{U}$,

$$early = \{\langle 4, 0 \rangle, \langle 8, 1 \rangle, \langle 12, 0.9 \rangle, \langle 16, 0.7 \rangle, \langle 20, 0.5 \rangle, \langle 24, 0.2 \rangle\}$$

For simplicity define 'can study' as a singleton fuzzy set $\mu_{study} = 1$. If the hour is truly early, for instance 8 o'clock in the morning, then $\mu_{early}(8) = 1$, and thus John can study to the fullest degree, that is $\mu_{study} = 1$. However, at 20 (8 pm) then $\mu_{early}(20) = 0.5$, and accordingly John can study to the degree 0.5. The degree of fulfillment of the antecedent (the *if*-side) weights the degree of fulfillment of the conclusion — a useful mechanism, that enables one rule to cover a range of hours. The procedure is: given a particular time instant t_0, the resulting truth-value is computed as $\min(\mu_{early}(t_0), \mu_{study})$.

4.1 Linguistic Variables

Whereas an algebraic variable takes numbers as values, a *linguistic variable* takes words or sentences as values (Zadeh in Zimmermann, 1993[16]). The name of such a linguistic variable is its *label*. The set of values that it can take is called its *term set*. Each value in the term set is a *linguistic value* or *term* defined over the universe. In summary: A linguistic variable takes a linguistic value, which is a fuzzy set defined on the universe.

Example 15 *Term set Age*
Let x be a linguistic variable labelled 'Age'. Its term set $\mathcal{T}$ could be defined as

$$\mathcal{T}(age) = \{young,\ very\ young,\ not\ very\ young,\ old,\ more\ or\ less\ old\}$$

Each term is defined on the universe, for example the integers from 0 to 100 years.

A *hedge* is a word that acts on a term and modifies its meaning. For example, in the sentence 'very near zero', the word 'very' modifies the term 'near zero'. Examples of other hedges are 'a little', 'more or less', 'possibly', and 'definitely'. In fuzzy reasoning a hedge operates on a membership function, and the result is a membership function.

Even though it is difficult precisely to say what effect the hedge 'very' has, it does have an intensifying effect. The hedge 'more or less' (or 'morl' for short) has the opposite effect. Given a fuzzy term with the label $\mathcal{A}$ and membership function $\mu_{\mathcal{A}}(x)$ defined on the universe

$\mathcal{X}$, the hedges 'very' and 'morl' are defined

$$very\ \mathcal{A} \equiv \left\{\left\langle x, \mu_{very\ \mathcal{A}}(x)\right\rangle \mid \mu_{very\ \mathcal{A}}(x) = \mu^2_{\mathcal{A}}(x),\ x \in \mathcal{X}\right\}$$
$$morl\ \mathcal{A} \equiv \left\{\langle x, \mu_{morl\ \mathcal{A}}(x)\rangle \mid \mu_{morl\ \mathcal{A}}(x) = \mu^{\frac{1}{2}}_{\mathcal{A}}(x),\ x \in \mathcal{X}\right\}$$

We have applied squaring and square root, but a whole family of hedges is generated by $\mu^k_{\mathcal{A}}$ or $\mu^{\frac{1}{k}}_{\mathcal{A}}$ (with integer k). For example

$$extremely\ \mathcal{A} \equiv \left\{\left\langle x, \mu_{very\ \mathcal{A}}(x)\right\rangle \mid \mu_{very\ \mathcal{A}}(x) = \mu^3_{\mathcal{A}}(x),\ x \in \mathcal{X}\right\}$$
$$slightly\ \mathcal{A} \equiv \left\{\left\langle x, \mu_{very\ \mathcal{A}}(x)\right\rangle \mid \mu_{very\ \mathcal{A}}(x) = \mu^{\frac{1}{3}}_{\mathcal{A}}(x),\ x \in \mathcal{X}\right\}$$

A derived hedge is for example $somewhat\ \mathcal{A}$ defined as $morl\ \mathcal{A}\ and\ not\ slightly\ \mathcal{A}$. For the special case where $k = 2$, the operation μ^2 is *concentration* and $\mu^{\frac{1}{2}}$ is *dilation*. With $k = \infty$ the hedge $\mu^k_{\mathcal{A}}$ could be named $exactly$, because it would suppress all memberships lower than 1.

Example 16 *Very on a discrete membership function*
Assume a discrete universe $\mathcal{U} = \langle 0, 20, 40, 60, 80\rangle$ of ages. In Matlab we can assign

```
u = [0 20 40 60 80]
```

and

```
young = [1 .6 .1 0 0]
```

The discrete membership function for the set 'very young' is `young.^2`*,*

```
1 0.36 0.01 0 0
```

The notation '`.^`*' is Matlab notation for the power operator. The set 'very very young' is, by repeated application,* `young.^4`*,*

```
1 0.13 0 0 0
```

The derived sets inherit the universe of the primary set.

A *primary term* is a term that must be defined a priori, for example *Young* and *Old* in Fig. 4, whereas the sets *Very young* and *Not very young* are modified sets. The primary terms can be modified by negation ('not') or hedges ('very', 'more or less'), and the resulting sets can be connected using connectives ('and', 'or', 'implies', 'equals'). Long sentences can be built using this vocabulary, and the result is still a membership function.

4.2 Modus Ponens Inference

Modus ponens generalised to fuzzy logic is the core of fuzzy reasoning. Consider the argument

$$\begin{array}{l} A' \\ A \Rightarrow B \\ \hline B' \end{array} \qquad (14)$$

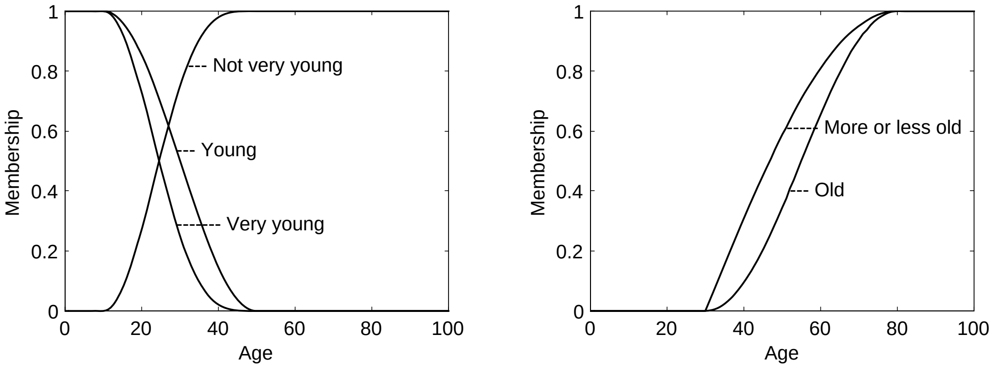

Figure 4. The membership functions for *very young* and *not very young* are derived from *young* (left), and the membership function for *more or less old* from *old* (right). (figage.m)

It is based on modus ponens, but the premise A' is slightly different from A and thus the conclusion B' is slightly different from B. For instance, given the rule 'if x is High, then y is Low'; if x in fact is 'very high', we would like to conclude that y is 'very low'.

Let A and A' be fuzzy sets defined on $\mathcal{X}$, and B a fuzzy set defined on $\mathcal{Y}$. Then the fuzzy set B', induced by 'x is A' from the fuzzy rule

$$\text{if } x \text{ is } A \text{ then } y \text{ is } B,$$

is given by

$$B' = A' \circ (A \Rightarrow B) .$$

The operation $\circ$ is composition, the inner $\vee - \wedge$ product.

Example 17 *Generalised modus ponens*

Given the rule 'if altitude is High, then oxygen is Low'. Let the fuzzy set High be defined on a set of altitude ranges from 0 to 4000 metres (about 12 000 feet),

$$High = \{\langle 0, 0\rangle, \langle 1000, 0.25\rangle, \langle 2000, 0.5\rangle, \langle 3000, 0.75\rangle, \langle 4000, 1\rangle\}$$

and Low be defined on a set of percentages of normal oxygen content,

$$Low = \{\langle 0, 1\rangle, \langle 25, 0.75\rangle, \langle 50, 0.5\rangle, \langle 75, 0.25\rangle, \langle 100, 1\rangle\}$$

As a shorthand notation we write the rule as a logical proposition High $\Rightarrow$ Low, where it is understood that the proposition concerns altitude and oxygen. We construct the relation $\mathbf{R}$ *connecting High and Low using sharp implication, or* $x \leq y$,

$\mathbf{R} =$

	1	.75	.5	.25	0
0	1	1	1	1	1
.25	1	1	1	1	0
.5	1	1	1	0	0
.75	1	1	0	0	0
1	1	0	0	0	0

The matrix is displayed with boxes and axis annotations to make the construction of the table clear: each element r_{xy} is the evaluation of $\mu_{High}(x) \leq \mu_{Low}(y)$. The numbers on the vertical axis correspond to μ_{High} and the numbers on the horisontal axis correspond to μ_{Low}. Assuming altitude is High, we find by modus ponens

$$\begin{aligned}
\boldsymbol{\mu}^t &= \boldsymbol{\mu}^t_{High} \circ \mathbf{R} \\
&= \begin{pmatrix} 0 & .25 & .5 & .75 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & .75 & .5 & .25 & 0 \end{pmatrix}
\end{aligned}$$

The result is identical to Low, thus modus ponens returns a result as expected in that case.

Assume instead altitude is Very High,

$$\boldsymbol{\mu}^t_{VeryHigh} = \begin{pmatrix} 0 & .06 & .25 & .56 & 1 \end{pmatrix},$$

the square of $\boldsymbol{\mu}^t_{High}$. Modus ponens yields

$$\begin{aligned}
\boldsymbol{\mu}^t &= \boldsymbol{\mu}^t_{VeryHigh} \circ \mathbf{R} \\
&= \begin{pmatrix} 0 & .06 & .25 & .56 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & .56 & .25 & .06 & 0 \end{pmatrix}
\end{aligned}$$

The result is identical to the square of μ_{Low}. Written as an argument, we have

Very High
High ⇒ Low
———————
Very Low

This is not always the case, though. With a slightly different definition of Low, say,

$$Low = \{\langle 0, 1\rangle, \langle 25, 0.8\rangle, \langle 50, 0.5\rangle, \langle 75, 0.3\rangle, \langle 100, 0\rangle\}$$

the resulting vector $\boldsymbol{\mu}^t$ would be the same as previously, but only approximately equal to Very Low.

5 Summary

Fuzzy reasoning is based on fuzzy logic, which is again based on fuzzy set theory. The idea of a fuzzy set is basic and simple: an object is allowed to have a gradual membership of a set. This simple idea pervades all derived mathematical aspects of set theory. In fuzzy logic an assertion is allowed to be more or less true. A truth value in fuzzy logic is a real number in the interval $[0, 1]$, rather than just the set of two truth values $\{0, 1\}$ of classical logic. Classical logic can be fuzzified in many ways, but the central problem is to find a suitable definition for the connective ’implication’. Fuzzy reasoning is based on the modus ponens rule of inference, which again rests on the definition of ’implication’.

Not all laws in classical logic can be valid in fuzzy logic, therefore the design of a fuzzy system is a trade-off between mathematical rigour and engineering requirements.

A backward approach is recommended:

1. Start by deciding which laws are required to hold;
2. define ’and’, ’or’, ’not’, ’implication’, and ’equivalence’;
3. check by means of their truth tables whether the laws in step 1 hold; and
4. if not, go to 2.

An assumption made in the tutorial is that the laws can be checked by just checking all combinations of three truth-values $\{0, 0.5, 1\}$. This is true for all expressions involving variables connected with $\wedge$, $\vee$, and $\neg$, but one must be cautious with expressions involving implication, as these generally cannot be reduced to those three connectives. Furthermore, it is assumed in the tutorial that truth-values are taken from a finite, discrete set of truth-values. This is not the case in general.

References

[1] O. I. Franksen. Group representation of finite polyvalent logic. In A. Niemi, editor, *Proceedings 7th Triennial World Congress*, Helsinki, 1978. International federation of automatic control, IFAC, Pergamon.

[2] S. Fukami, M. Mizumoto, and K. Tanaka. Some considerations of fuzzy conditional inference. *Fuzzy Sets and Systems*, 4:243–273, 1980.

[3] Mai Gerhke, Carol L. Walker, and Elbert A. Walker. Normal forms and truth tables for fuzzy logics. *Fuzzy Sets And Systens*, 138:25–51, 2003.

[4] J.-S. R. Jang, C.-T. Sun, and E. Mizutani. *Neuro-Fuzzy and Soft Computing*. Number isbn 0-13-261066-3 in Matlab Curriculum Series. Prentice Hall, Upper Saddle River, NJ, USA, 1997.

[5] J. Jantzen. Array approach to fuzzy logic. *Fuzzy Sets and Systems*, 70:359–370, 1995.

[6] J. B. Kiszka, M. E. Kochanska, and D. S. Sliwinska. The influence of some fuzzy implication operators on the accuracy of a fuzzy model. *Fuzzy Sets and Systems*, 15:(Part1) 111–128; (Part 2) 223 – 240, 1985.

[7] C. C. Lee. Fuzzy logic in control systems: Fuzzy logic controller. *IEEE Trans. Systems, Man & Cybernetics*, 20(2):404–435, 1990.

[8] E. H. Mamdani. Application of fuzzy logic to approximate reasoning using linguistic

synthesis. *IEEE Transactions on Computers*, C-26(12):1182–1191, December 1977.

[9] M. Mizumoto, S. Fukami, and K. Tanaka. Some methods of fuzzy reasoning. In Gupta, Ragade, and Yager, editors, *Advances in Fuzzy Set Theory Applications*. North-Holland, New York, 1979.

[10] Robert R. Stoll. *Set Theory And Logic*. Dover Publications, New York, dover edition, 1979. (org 1963).

[11] F. Wenstøp. Quantitative analysis with linguistic values. *Fuzzy Sets and Systems*, 4(2):99–115, 1980.

[12] L. A. Zadeh. Fuzzy sets. *Inf. and control*, 8:338–353, 1965.

[13] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Systems, Man & Cybernetics*, 1:28–44, 1973.

[14] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, 8:43–80, 1975.

[15] L. A. Zadeh. Fuzzy logic. *IEEE Computer*, 21(4):83–93, April 1988.

[16] H.-J. Zimmermann. *Fuzzy set theory - and its applications*. Kluwer, Boston, second edition, 1993. (1. ed. 1991).