COMP474/6741

Fuzzy logic and fuzzy inference

Negnevitsky: Chap4

Today

- Introduction

- Fuzzy sets
- How to represent a fuzzy set in a computer?
- Operations on fuzzy sets
- Fuzzy inference
    - Mamdani inference Model
- Appendix : example 2

Introduction
what is fuzzy thinking?

- Experts rely on **common sense** when they solve problems.
- **How can we represent expert knowledge that uses vague and ambiguous terms in a computer?**
- Fuzzy logic is not logic that is fuzzy, but logic that is used to describe fuzziness. Fuzzy logic is the theory of fuzzy sets, sets that calibrate vagueness.
- Fuzzy logic is based on the idea that all things admit of degrees. Temperature, height, speed, distance, beauty – all come on a sliding scale. The motor is running *really hot*. Tom is a *very tall* guy.

Introduction cont.

- Boolean logic uses sharp distinctions. It forces us to draw lines between members of a class and non-members. For instance, we may say, Tom is tall because his height is 181 cm. If we drew a line at 180 cm, we would find that David, who is 179 cm, is small. Is David really a small man or we have just drawn an arbitrary line in the sand?
- Fuzzy logic reflects how people think. It attempts to model our sense of words, our decision making and our common sense. As a result, it is leading to new, more human, intelligent systems.

Introduction cont.

- ***Why fuzzy?***

  As Zadeh said, the term is concrete, immediate and descriptive; we all know what it means. However, many people in the West were repelled by the word *fuzzy*, because it is usually used in a negative sense.

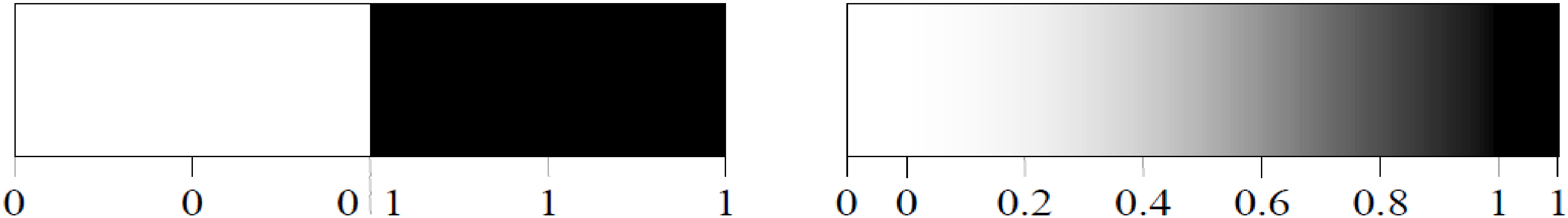
- ***Why logic?***

  Fuzziness rests on fuzzy set theory, and fuzzy logic is just a small part of that theory.

Fuzzy logic is a set of mathematical principles for knowledge representation based on degrees of membership.

Unlike two-valued Boolean logic, fuzzy logic is **multi-valued**. It deals with **degrees of membership** and **degrees of truth**. Fuzzy logic uses the continuum of logical values between 0 (completely false) and 1 (completely true). Instead of just black and white, it employs the spectrum of colours, accepting that things can be partly true and partly false at the same time.

Range of logical values in Boolean and fuzzy logic



(*a*) Boolean Logic. (*b*) Multi-valued Logic.

Today

- Introduction
- Fuzzy sets

- How to represent a fuzzy set in a computer?
- Operations on fuzzy sets
- Fuzzy inference
  - Mamdani inference Model
- Appendix : example 2

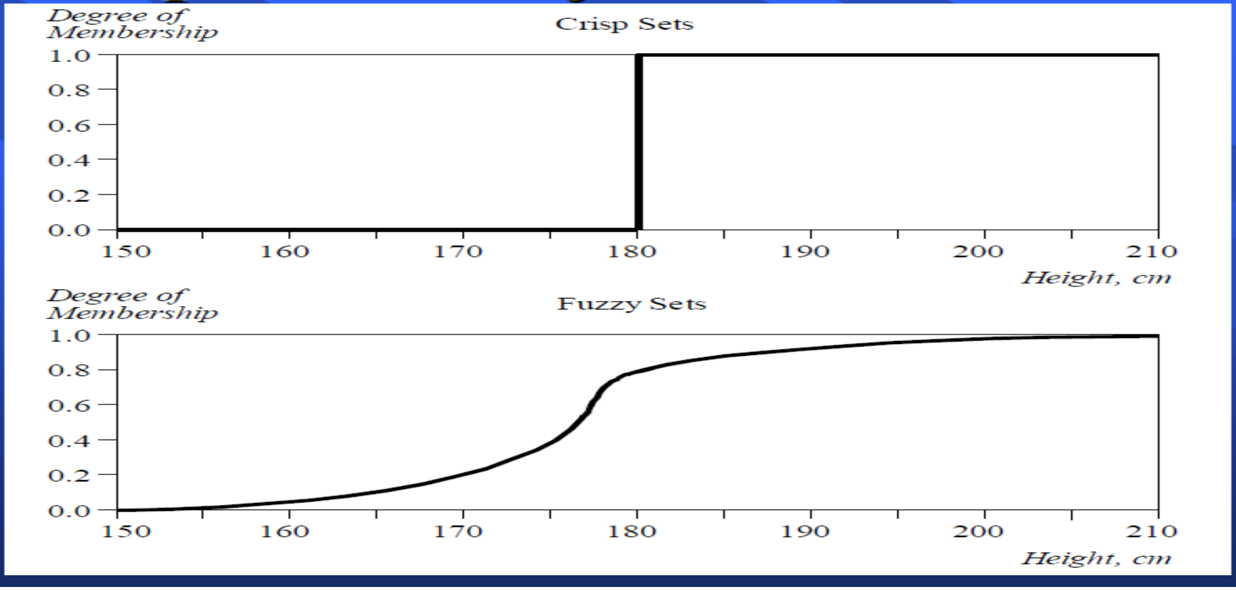
Fuzzy sets

- The concept of a **set** is fundamental to mathematics.
- However, our own language is also the supreme expression of sets. For example, *car* indicates the *set of cars*. When we say *a car*, we mean one out of the set of cars.

- The classical example in fuzzy sets is *tall men*. The elements of the fuzzy set "tall men" are all men, but their degrees of membership depend on their height.

Name	Height, cm	Degree of Membership	
		Crisp	*Fuzzy*
Chris	208	1	1.00
Mark	205	1	1.00
John	198	1	0.98
Tom	181	1	0.82
David	179	0	0.78
Mike	172	0	0.24
Bob	167	0	0.15
Steven	158	0	0.06
Bill	155	0	0.01
Peter	152	0	0.00

Crisp and fuzzy sets of "*tall men*"



- The *x*-axis represents the **universe of discourse** – the range of all possible values applicable to a chosen variable. In our case, the variable is the man height. According to this representation, the universe of men's heights consists of all tall men.
- The *y*-axis represents the **membership value of the fuzzy set**. In our case, the fuzzy set of "*tall men*" maps height values into corresponding membership values.

A fuzzy set is a set with fuzzy boundaries.

- Let X be the universe of discourse and its elements be denoted as x. In the classical set theory, **crisp set A of X is defined as function $f_A(x)$ called the characteristic function of A**

$f_A(x)$: $X \rightarrow \{0, 1\}$, where

$$f_A(x) = \begin{cases} 1, \text{if } x \in A \\ 0, \text{if } x \notin A \end{cases}$$

This set maps universe X to a set of two elements. For any element x of universe X, characteristic function $f_A(x)$ is equal to 1 if x is an element of set A, and is equal to 0 if x is not an element of A.

- In the fuzzy theory, fuzzy set A of universe X is defined by function $\mu_A(x)$ called the *membership function* of set A

$\mu_A(x)$: $X \rightarrow [0, 1]$, where
$\mu_A(x) = 1$ if x is totally in A;
$\mu_A(x) = 0$ if x is not in A;
$0 < \mu_A(x) < 1$ if x is partly in A.

This set allows a continuum of possible choices. For any element x of universe X, membership function $\mu_A(x)$ equals the degree to which x is an element of set A. This degree, a value between 0 and 1, represents the **degree of membership**, also called **membership value**, of element x in set A.

Today

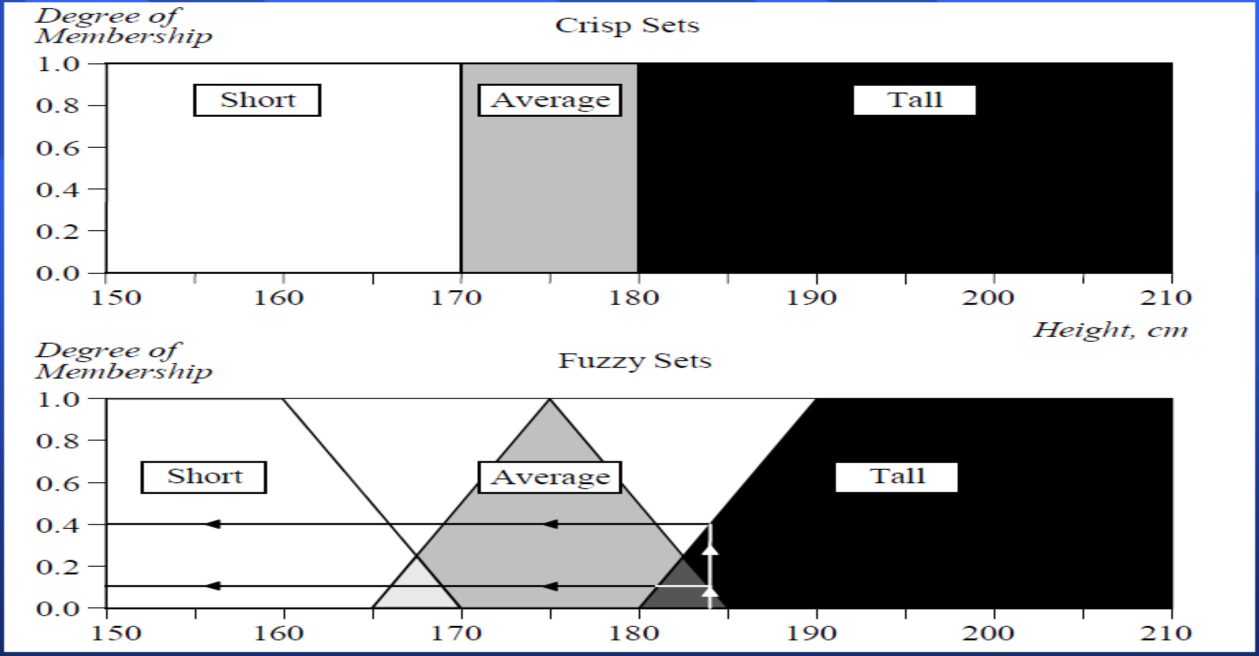
- Introduction
- Fuzzy sets
- How to represent a fuzzy set in a computer?

- Operations on fuzzy sets
- Fuzzy inference
  - Mamdani inference Model
- Appendix : example 2

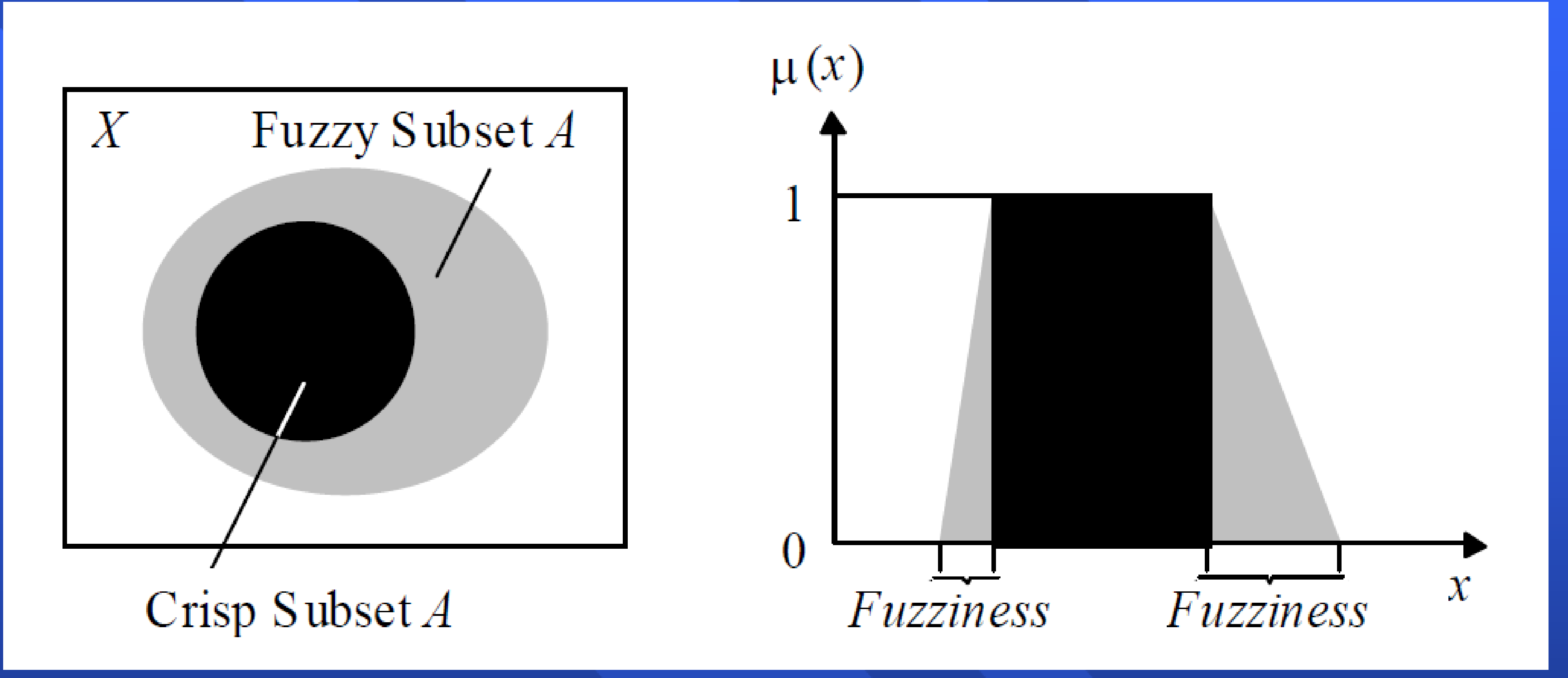
How to represent a fuzzy set in a computer?

- First, we determine the membership functions. In our "*tall men*" example, we can obtain fuzzy sets of *tall*, *short* and *average* men.
- The universe of discourse – the men's heights – consists of three sets: *short*, *average* and *tall men*. As you will see, a man who is 184 cm tall is a member of the *average men* set with a degree of membership of 0.1, and at the same time, he is also a member of the *tall men* set with a degree of 0.4.

Crisp and fuzzy sets of short, average and tall men



Representation of crisp and fuzzy subsets



Typical functions that can be used to represent a fuzzy set are sigmoid, gaussian and pi. However, these functions increase the time of computation. Therefore, in practice, most applications use **linear fit functions**.

Linguistic variables and hedges

- At the root of fuzzy set theory lies the idea of linguistic variables.
- **A linguistic variable is a fuzzy variable.** For example, the statement "John is tall" implies that the linguistic variable *John* takes the linguistic value *tall*.

In fuzzy expert systems, linguistic variables are used in fuzzy rules. For example:

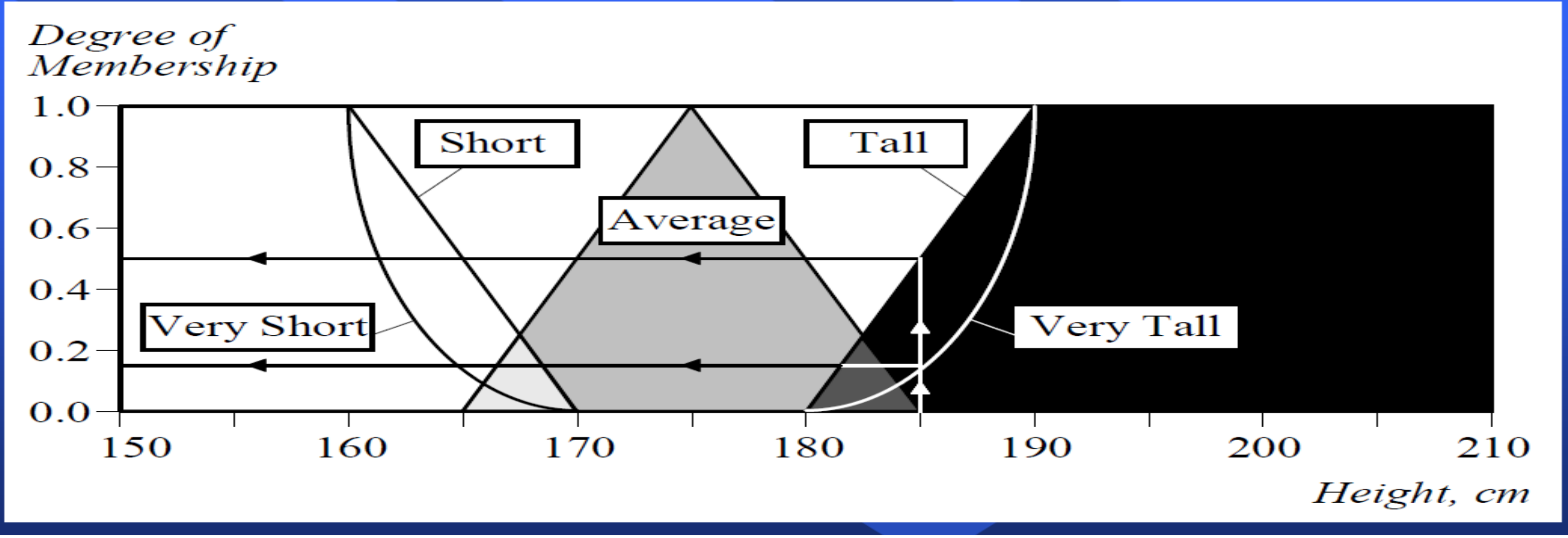
IF **wind is strong**
THEN **sailing is good**

IF **project_duration is long**
THEN **completion_risk is high**

IF **speed is slow**
THEN **stopping_distance is short**

- The range of possible values of a linguistic variable represents the universe of discourse of that variable. For example, the universe of discourse of the linguistic variable *speed* might have the range between 0 and 220 km/h and may include such fuzzy subsets as *very slow*, *slow*, *medium*, *fast*, and *very fast*.
- A linguistic variable carries with it the concept of fuzzy set qualifiers, called *hedges*.
- **Hedges are terms that modify the shape of fuzzy sets. They include adverbs such as *very*, *somewhat*, *quite*, *more or less* and *slightly*.**

Fuzzy sets with the hedge *very*



Today

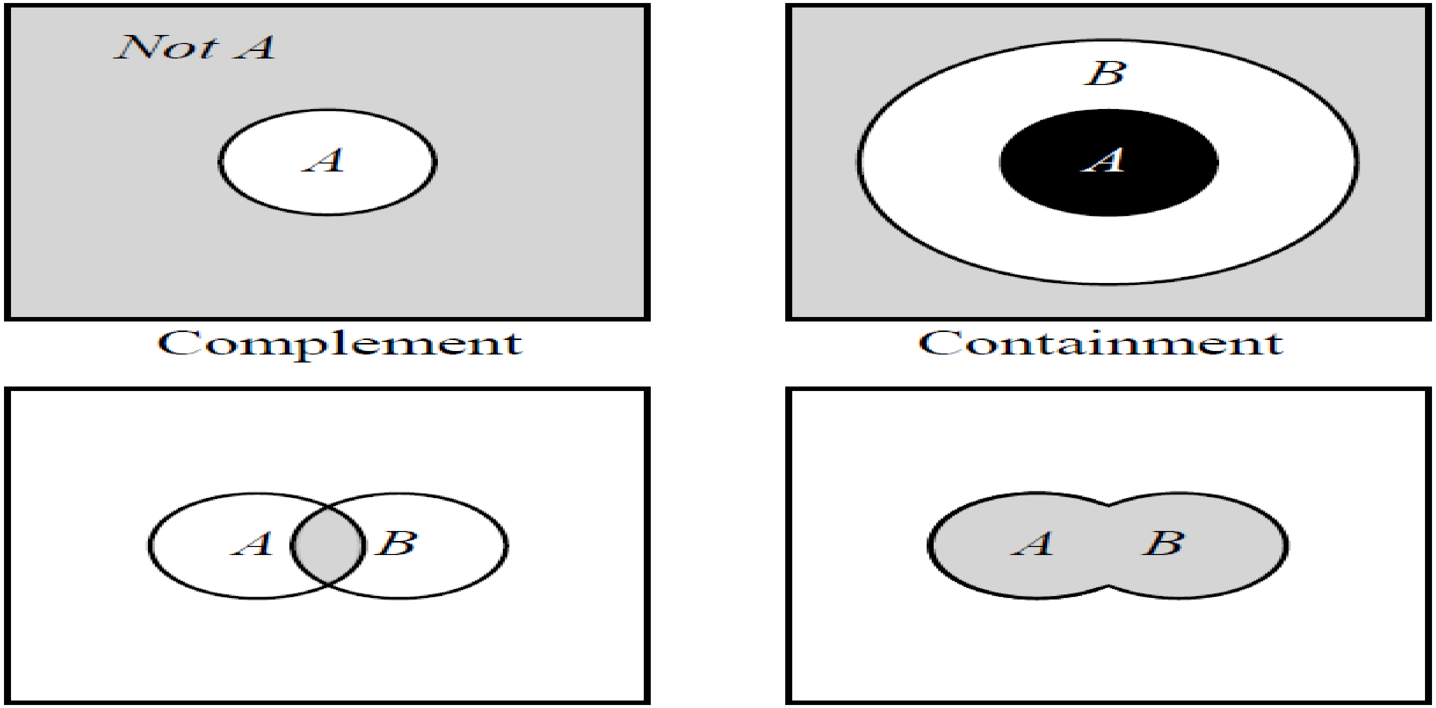
- Introduction
- Fuzzy sets
- How to represent a fuzzy set in a computer?
- Operations on fuzzy sets

- Fuzzy inference
  - Mamdani inference Model
- Appendix : example 2

Operations of fuzzy sets

The classical set theory developed in the late 19th century by Georg Cantor describes how crisp sets can interact. These interactions are called **operations**.

Cantor's sets



Complement

Containment

Intersection

Union

Complement

Crisp Sets: Who does not belong to the set?
Fuzzy Sets: How much do elements not belong to the set?

Containment

Crisp Sets: Which sets belong to which other sets?
Fuzzy Sets: Which sets belong to other sets?

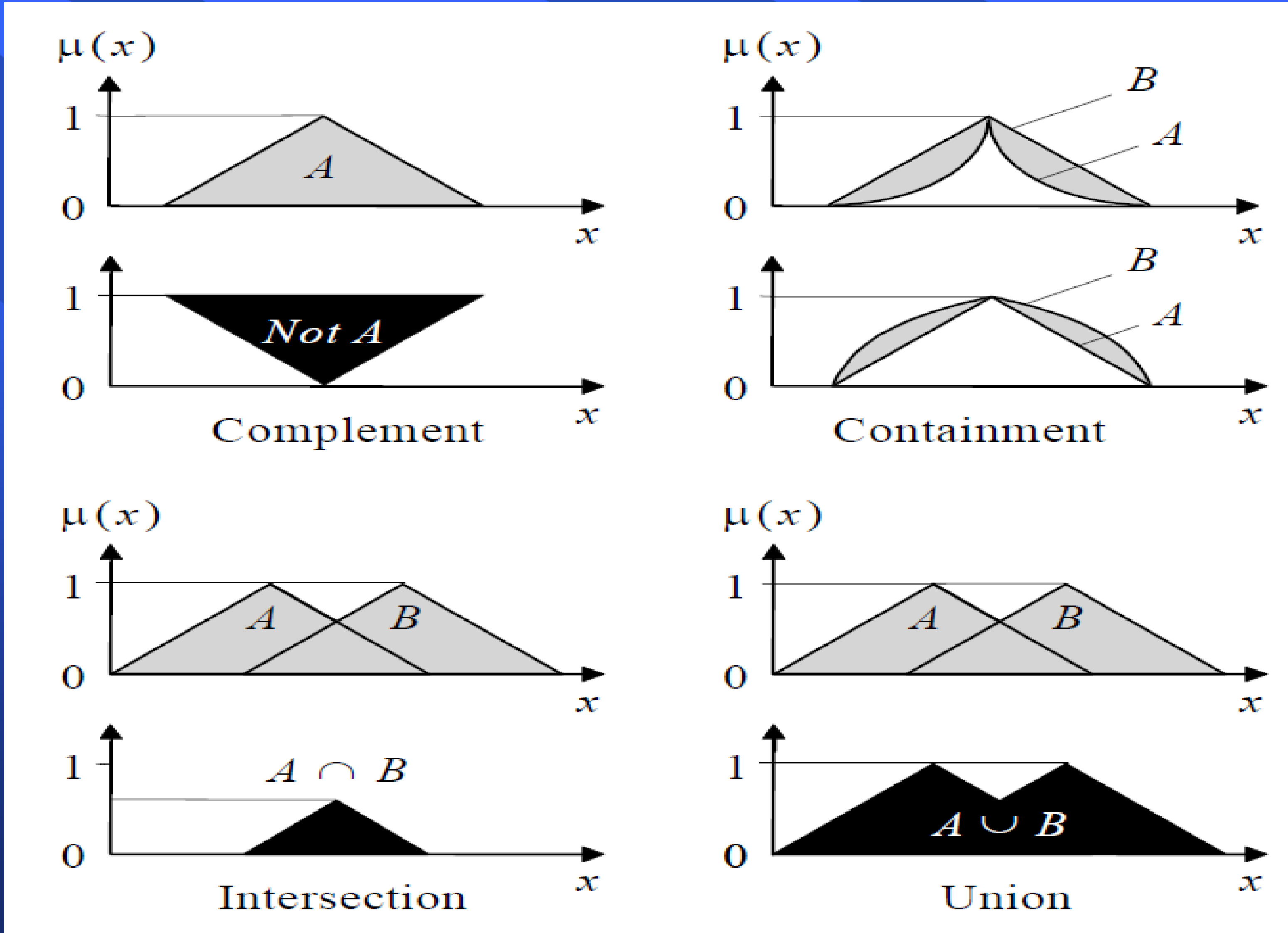
Intersection

Crisp Sets: Which element belongs to both sets?
Fuzzy Sets: How much of the element is in both sets?

Union

Crisp Sets: Which element belongs to either set?
Fuzzy Sets: How much of the element is in either set?

Operations of fuzzy sets



What is a fuzzy rule?

A fuzzy rule can be defined as a conditional statement in the form:

IF *x* **is** *A*
THEN *y* **is** *B*

where *x* and *y* are linguistic variables; and *A* and *B* are linguistic values determined by fuzzy sets on the universe of discourses *X* and *Y*, respectively.

What is the difference between classical and fuzzy rules?

A classical IF-THEN rule uses binary logic, for example,

Rule: 1
IF speed is > 100
THEN stopping_distance is long

Rule: 2
IF speed is < 40
THEN stopping_distance is short

The variable *speed* can have any numerical value between 0 and 220 km/h, but the linguistic variable *stopping_distance* can take either value *long* or *short*. In other words, classical rules are expressed in the black-and-white language of Boolean logic.

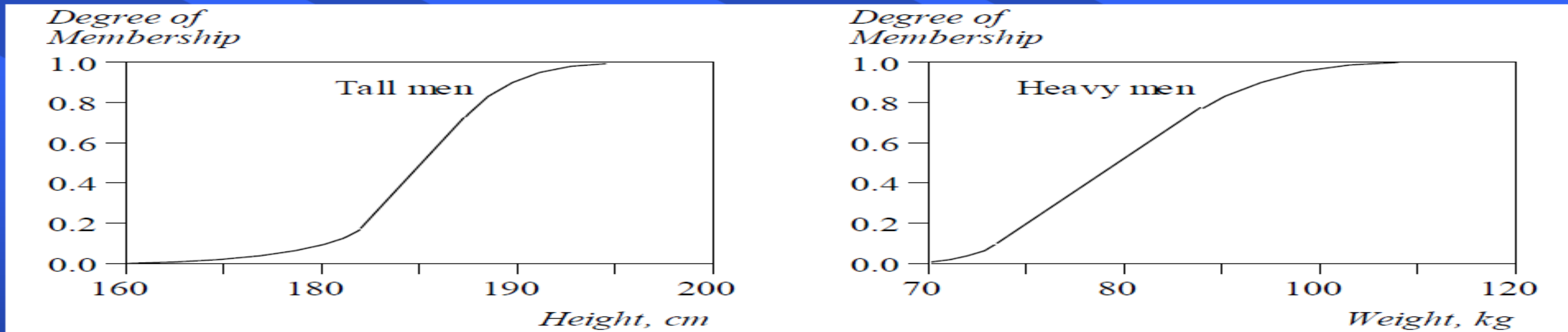
We can also represent the stopping distance rules in a fuzzy form:

Rule: 1
IF speed is fast
THEN stopping_distance is long

Rule: 2
IF speed is slow
THEN stopping_distance is short

In fuzzy rules, the linguistic variable *speed* also has the range (the universe of discourse) between 0 and 220 km/h, but this range includes fuzzy sets, such as *slow*, *medium* and *fast*. The universe of discourse of the linguistic variable *stopping_distance* can be between 0 and 300 m and may include such fuzzy sets as *short*, *medium* and *long*.

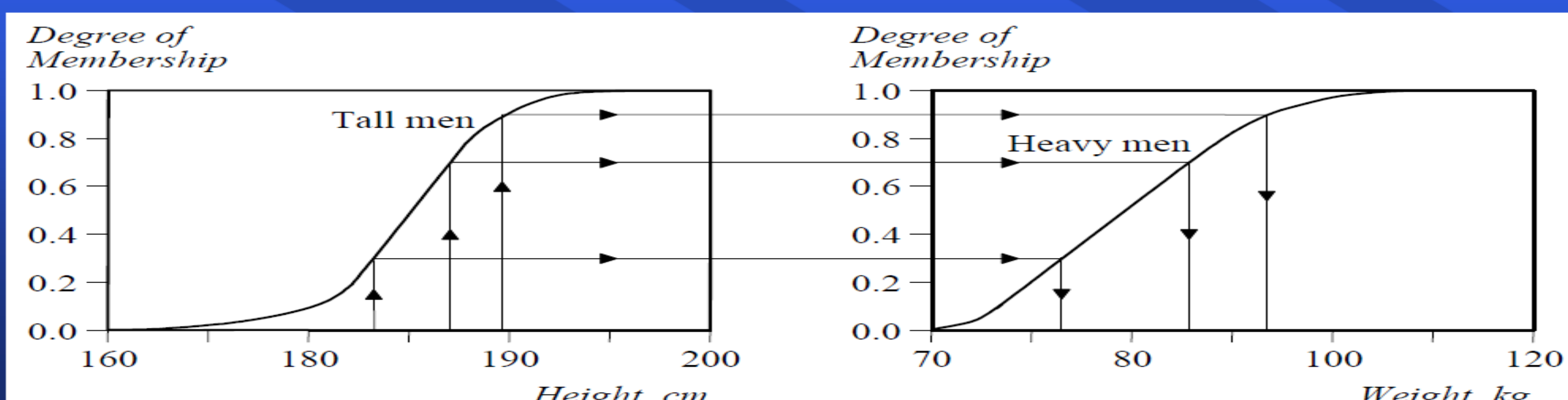
Fuzzy sets of *tall* and *heavy* men



These fuzzy sets provide the basis for a weight estimation model. The model is based on a relationship between a man's height and his weight:

IF height is *tall*
THEN weight is *heavy*

The value of the output or a truth membership grade of the rule consequent can be estimated directly from a corresponding truth membership grade in the antecedent. This form of fuzzy inference uses a method called **monotonic selection**.



A fuzzy rule can have multiple antecedents, for example:

IF project_duration is long
AND project_staffing is large
AND project_funding is inadequate
THEN risk is high

IF service is excellent
OR food is delicious
THEN tip is generous

The consequent of a fuzzy rule can also include multiple parts, for instance:

IF temperature is hot
THEN hot_water is reduced;
cold_water is increased

Today

- Introduction
- Fuzzy sets
- How to represent a fuzzy set in a computer?
- Operations on fuzzy sets
- Fuzzy inference
  - Mamdani inference Model
- Appendix : example 2

Fuzzy expert systems: Fuzzy inference

Two models are the most used

- **Mamdani inference model:** The consequent of a rule is a fuzzy fact, the inference consists in to apply the possibility value of the antecedent to the consequent
- **Sugeno inference model:** The consequent of a rule is a linear function of the possibility values of the antecedent, the value of this function is the result of the inference

Fuzzy inference

The most commonly used fuzzy inference technique is the so-called Mamdani method. In 1975, Professor **Ebrahim Mamdani** of London University built one of the first fuzzy systems to control a steam engine and boiler combination. He applied a set of fuzzy rules supplied by experienced human operators.

Mamdani inference model

fuzzification

- **Antecedent evaluation:** Computation for each rule of the possibility value of the crisp values using the antecedent facts and their combination
- **Consequent evaluation:** Each consequent is weighted using the possibility value of its antecedent
- **Conclusion combination:** The conclusions are combined in a fuzzy set that represents the joint conclusion
- **Nitidification:** The crisp value corresponding to the conclusion is computed from the fuzzy (for example computing the center of mass of the fuzzy set)

defuzzification

$$CDG(f(x)) = \frac{\int_a^b f(x) \cdot x dx}{\int_a^b f(x) dx}$$

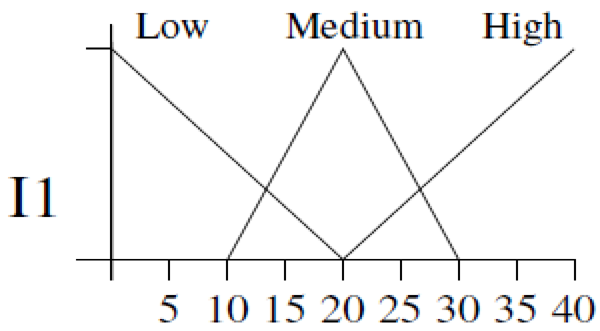
Fuzzy inference: example (1/8)

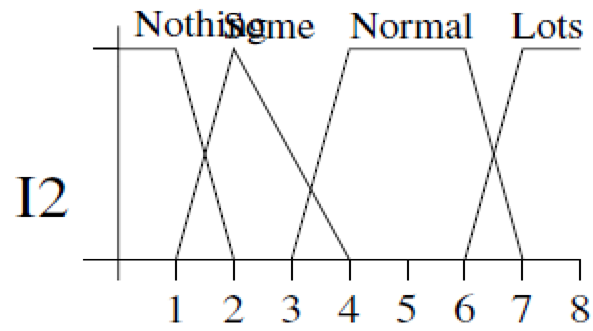
Assume that we have two input variables $I1$ and $I2$ and an output variable O that have the values defined by the following linguistic labels:

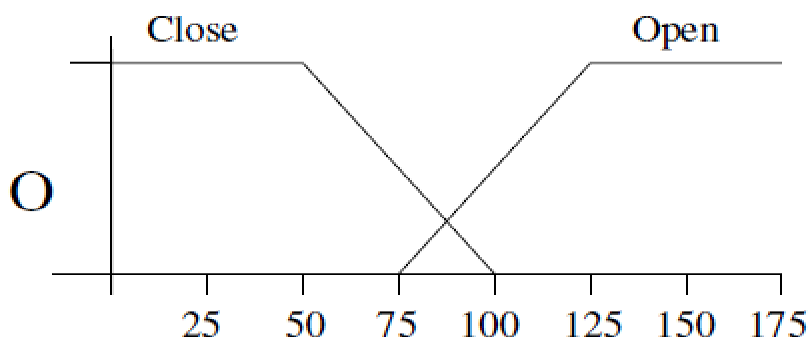
- $I1$ = {low,medium,high}
- $I2$ = {nothing,some,normal,lots}
- O = {close,open}

Fuzzy inference: example (2/8)

These labels are represented by the following fuzzy sets







Fuzzy inference: example (3/8)

Assume that we have the following rules:

R1. if ([I1=medium] or [I1=high]) y [I2=lots] then [O=close]

R2. if [I1=high] and [I2=normal] then [O=close]

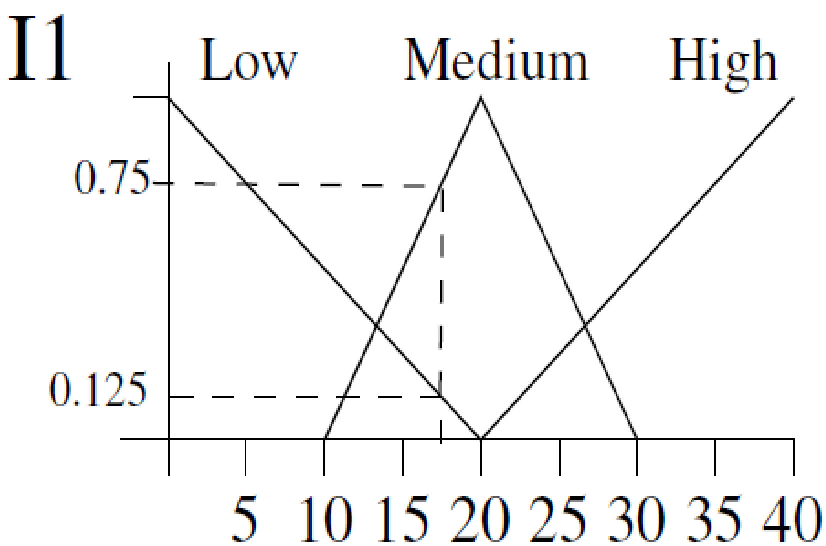
R3. if [I1=low] and not([I2=lots]) then [O=open]

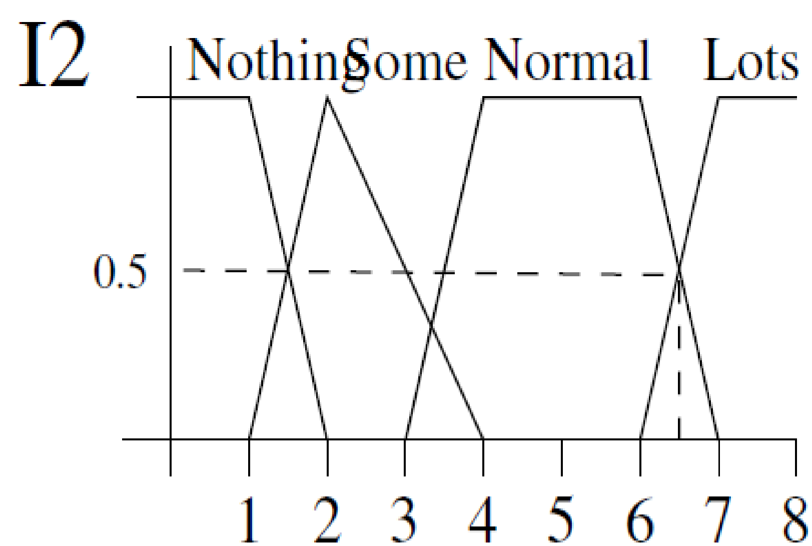
R4. if ([I1=low] or [I1=medium]) and [I2=some] then [O=open]

The combination functions are the minimum for the conjunction, the maximum for the disjunction and $1 - x$ for the negation and the crisp values for variables $I1$ and $I2$ are 17.5 and 6.5 respectively

Fuzzy inference: example (4/8)

If we use the values on the fuzzy sets of variables $I1$ and $I2$ we obtain the following possibility values





Fuzzy inference: example (5/8)

- If we evaluate the **rule R1** we have:

  [I1=medium] = 0.75, [I1=high] = 0, [I2=lots] = 0.5 $\Rightarrow$ min(max(0.75,0),0.5) = 0.5
  So we have 0.5 $\cdot$ [O=close]

- If we evaluate the **rule R2** we have:

  [I1=high] = 0, [I2=normal] = 0.5 $\Rightarrow$ min(0,0.5) = 0
  So we have 0 $\cdot$ [O=close]

- If we evaluate the **rule R3** we have:

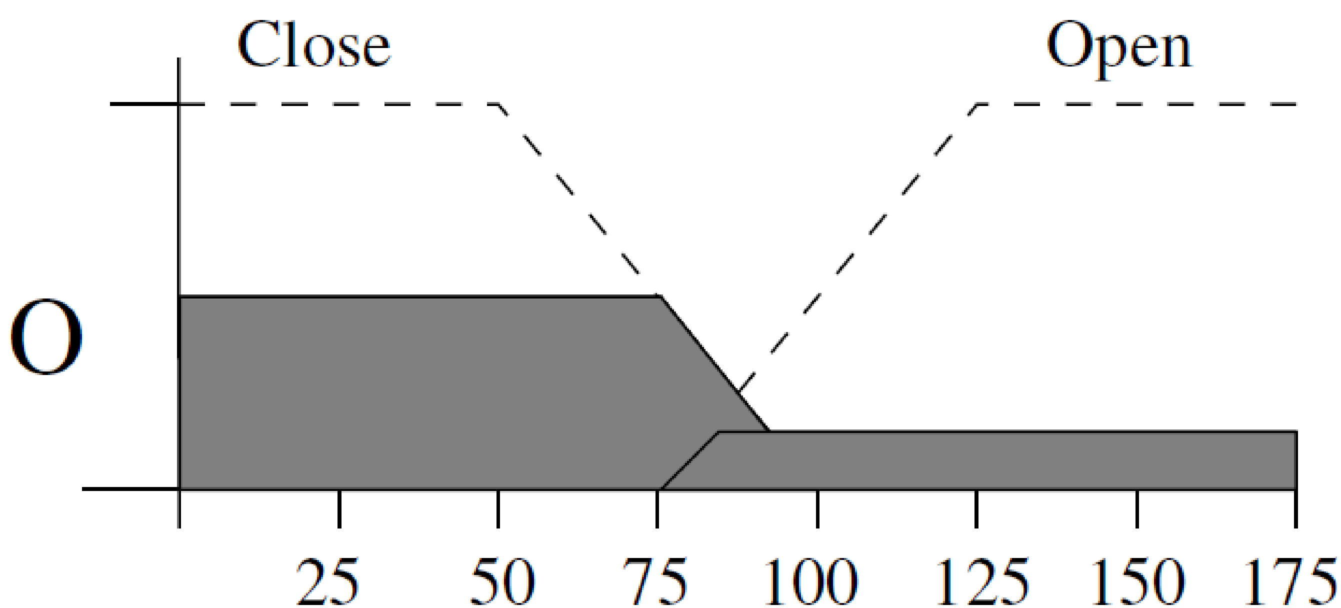
  [I1=low] = 0.125, [I1=lots] = 0.5, $\Rightarrow$ min(0.125,1-0.5) = 0.125
  So we have 0.125 $\cdot$ [O=open]

- If we evaluate the **rule R4** we have:

  [I1=low] = 0.125, [I1=medium] = 0.75, [I2=some] = 0 $\Rightarrow$ min(max(0.125,0.75),0) = 0
  So we have 0 $\cdot$ [O=open]

Fuzzy inference: example (6/7)

The fuzzy set resulting from the combination of the conclusion is:



Fuzzy inference: example (7/8)

Now we compute the center of mass of the fuzzy set, this can be described with the function:

$$f(x) = \begin{cases} 0.5 & x \in [0 - 75] \\ (100 - x)/50 & x \in (75 - 93.75] \\ 0.125 & x \in (93.75 - 100] \end{cases}$$

Fuzzy inference: example (8/8)

$$CDG(f(x)) = \frac{\int_a^b f(x) \cdot x dx}{\int_a^b f(x) dx}$$

The CoM can be computed as:

$$\frac{\int_0^{75} 0.5 \cdot x \cdot dx + \int_{75}^{93.75} (100 - x)/50 \cdot x \cdot dx + \int_{93.75}^{175} 0.125 \cdot x \cdot dx}{\int_0^{75} 0.5 \cdot dx + \int_{75}^{93.75} (100 - x)/50 \cdot dx + \int_{93.75}^{175} 0.125 \cdot dx} =$$

$$\frac{0.25 \cdot x^2|_0^{75} + (x^2 - x^3/150)|_{75}^{93.75} + 0.0625 \cdot x^2|_{93.75}^{175}}{0.5 \cdot x|_0^{75} + (2 \cdot x - x^2/100)|_{75}^{93.75} + 0.125 \cdot x|_{93.75}^{175}} =$$

$$\frac{(0.25 \cdot 75^2 - 0) + ((93.75^2 - 93.75^3/150) - (75^2 - 75^3/150)) + (0.0625 \cdot 175^2 - 0.0625 \cdot 93.75^2)}{(0.5 \cdot 75 - 0) + ((2 \cdot 93.75 - 93.75^2/100) - (2 \cdot 75 - 75^2/100)) + (0.125 \cdot 175 - 0.125 \cdot 93.75)}$$

$$\frac{3254.39}{53.53} = 60.79$$

This mean open at 60.79 percent

Today

- Introduction
- Fuzzy sets
- How to represent a fuzzy set in a computer?
- Operations on fuzzy sets
- Fuzzy inference
  - Mamdani inference Model
- Appendix : example 2

Appendix: example 2

Mamdani fuzzy inference

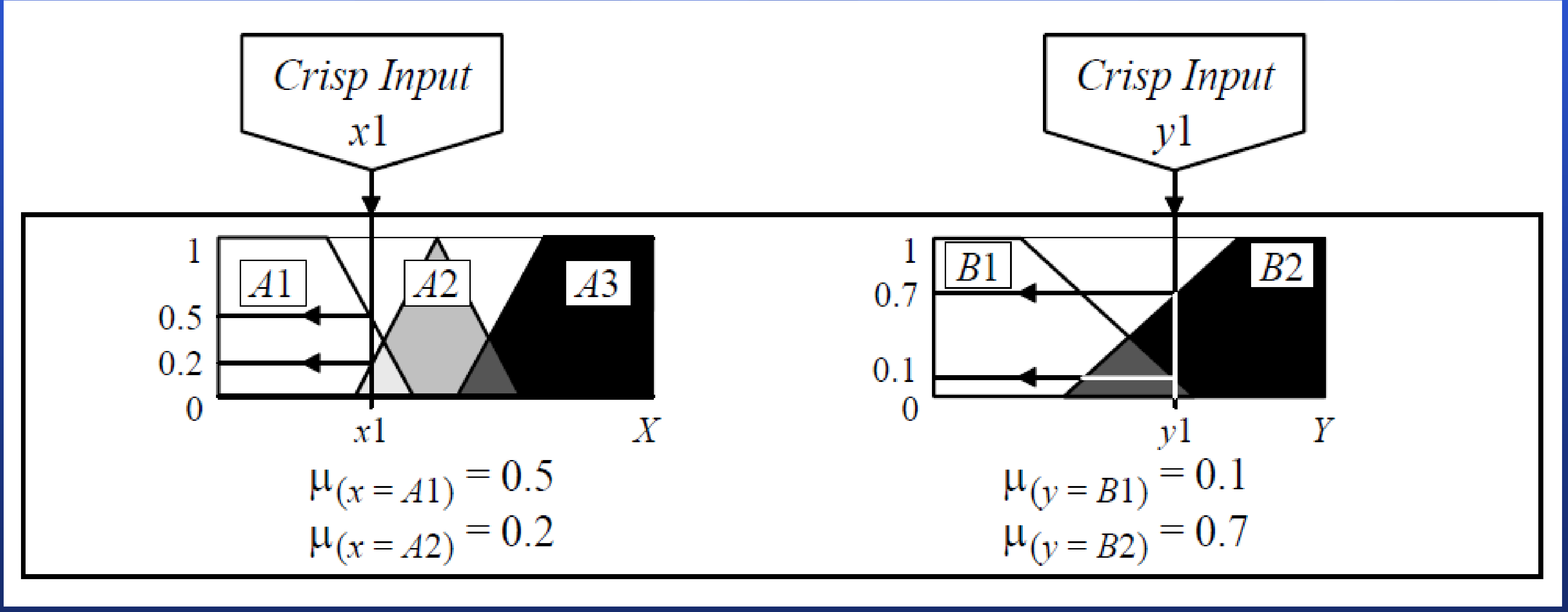
- The Mamdani-style fuzzy inference process is performed in four steps:
  - fuzzification of the input variables,
  - rule evaluation;
  - aggregation of the rule outputs, and finally
  - defuzzification.

We examine a simple two-input one-output problem that includes three rules:

Rule: 1	**Rule: 1**
IF *x* is *A3*	**IF** *project_funding* is *adequate*
OR *y* is *B1*	**OR** *project_staffing* is *small*
THEN *z* is *C1*	**THEN** *risk* is *low*
Rule: 2	**Rule: 2**
IF *x* is *A2*	**IF** *project_funding* is *marginal*
AND *y* is *B2*	**AND** *project_staffing* is *large*
THEN *z* is *C2*	**THEN** *risk* is *normal*
Rule: 3	**Rule: 3**
IF *x* is *A1*	**IF** *project_funding* is *inadequate*
THEN *z* is *C3*	**THEN** *risk* is *high*

Step 1: Fuzzification

The first step is to take the crisp inputs, *x*1 and *y*1 (*project funding* and *project staffing*), and determine the degree to which these inputs belong to each of the appropriate fuzzy sets.



Step 2: Rule Evaluation

The second step is to take the fuzzified inputs, $\mu_{(x=A1)} = 0.5$, $\mu_{(x=A2)} = 0.2$, $\mu_{(y=B1)} = 0.1$ and $\mu_{(y=B2)} = 0.7$, and apply them to the antecedents of the fuzzy rules. If a given fuzzy rule has multiple antecedents, the fuzzy operator (AND or OR) is used to obtain a single number that represents the result of the antecedent evaluation. This number (the truth value) is then applied to the consequent membership function.

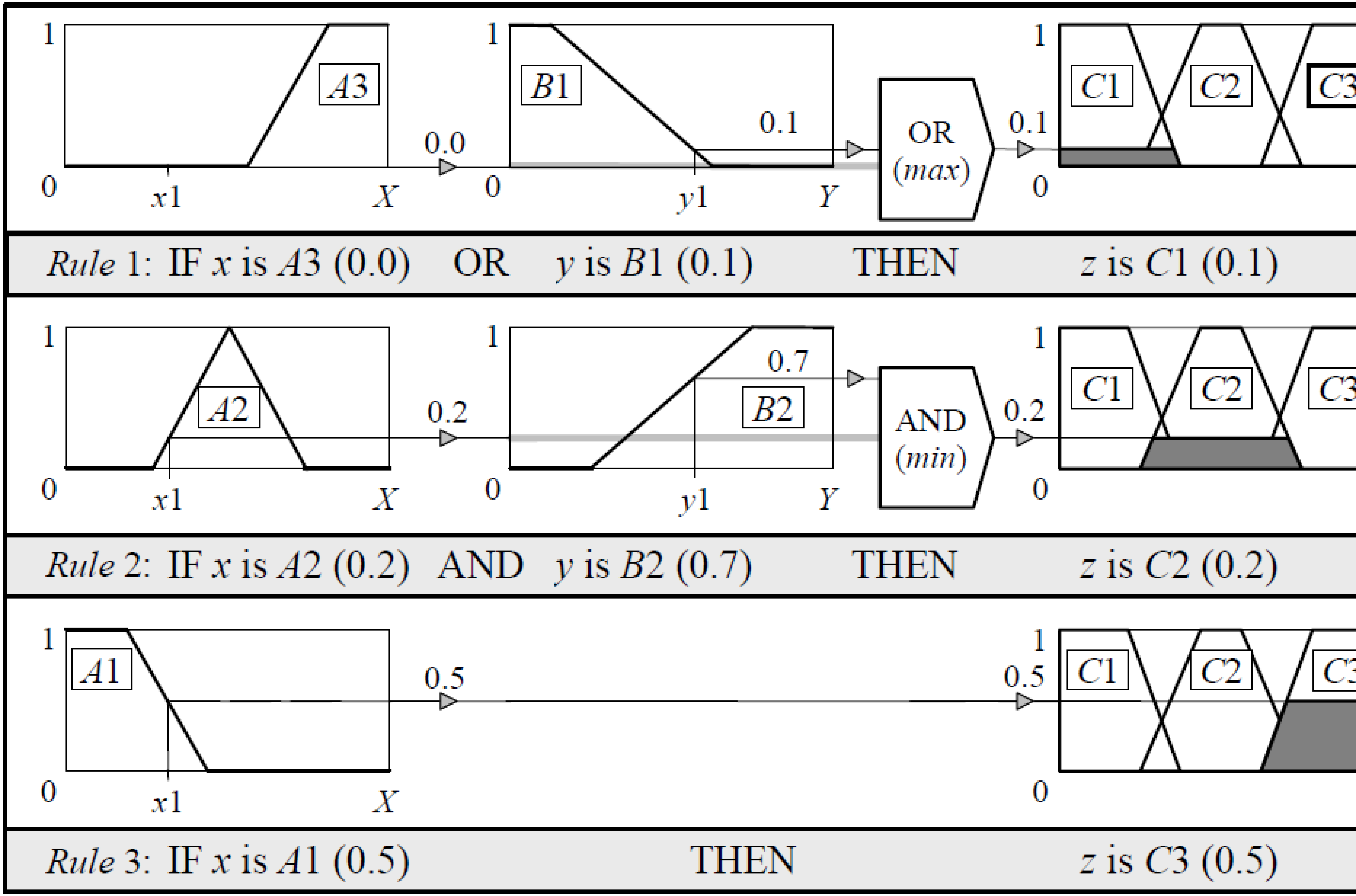
To evaluate the disjunction of the rule antecedents, we use the **OR fuzzy operation**. Typically, fuzzy expert systems make use of the classical fuzzy operation **union**:

$$\mu_{A \cup B}(x) = max\ [\mu_A(x), \mu_B(x)]$$

Similarly, in order to evaluate the conjunction of the rule antecedents, we apply the **AND fuzzy operation intersection**:

$$\mu_{A \cap B}(x) = min\ [\mu_A(x), \mu_B(x)]$$

Mamdani-style rule evaluation

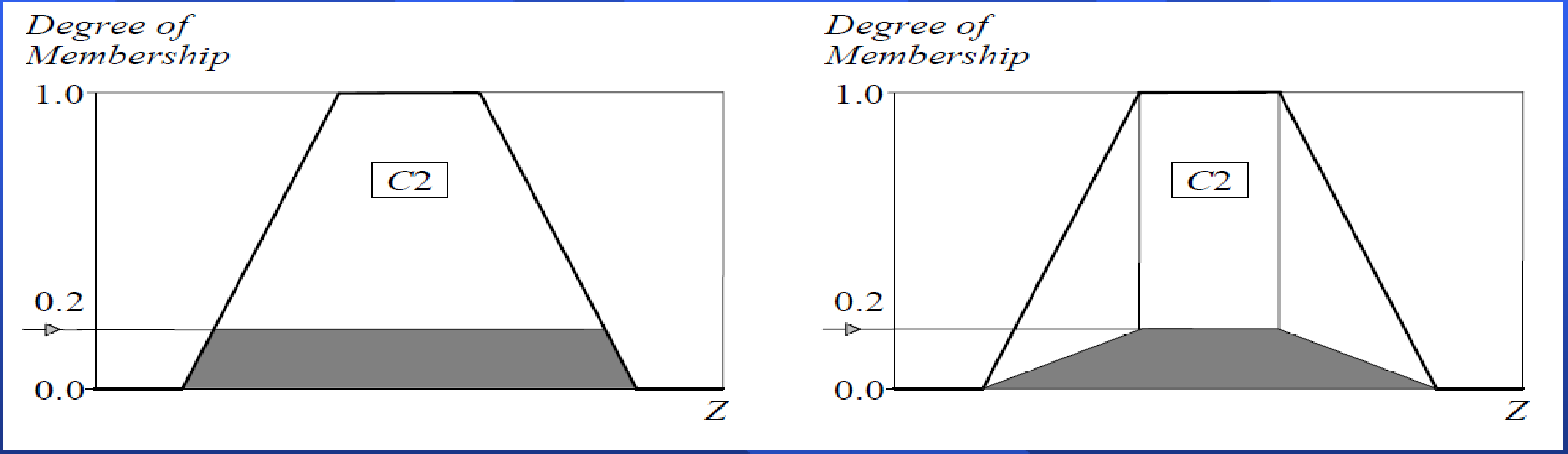


Now the result of the antecedent evaluation can be applied to the membership function of the consequent.

- The most common method of correlating the rule consequent with the truth value of the rule antecedent is to cut the consequent membership function at the level of the antecedent truth. This method is called **clipping**. Since the top of the membership function is sliced, the clipped fuzzy set loses some information. However, clipping is still often preferred because it involves less complex and faster mathematics, and generates an aggregated output surface that is easier to defuzzify.

- While clipping is a frequently used method, **scaling** offers a better approach for preserving the original shape of the fuzzy set. The original membership function of the rule consequent is adjusted by multiplying all its membership degrees by the truth value of the rule antecedent. This method, which generally loses less information, can be very useful in fuzzy expert systems.

Clipped and scaled membership functions

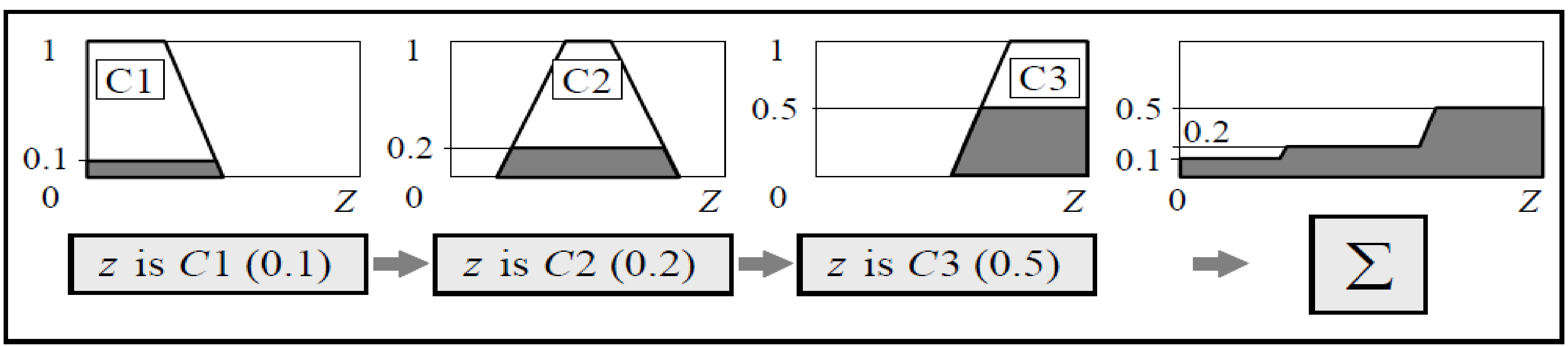


Step 3: Aggregation of the rule outputs

Aggregation is the process of unification of the outputs of all rules. We take the membership functions of all rule consequents previously clipped or scaled and combine them into a single fuzzy set.

The input of the aggregation process is the list of clipped or scaled consequent membership functions, and the output is one fuzzy set for each output variable.

Aggregation of the rule outputs



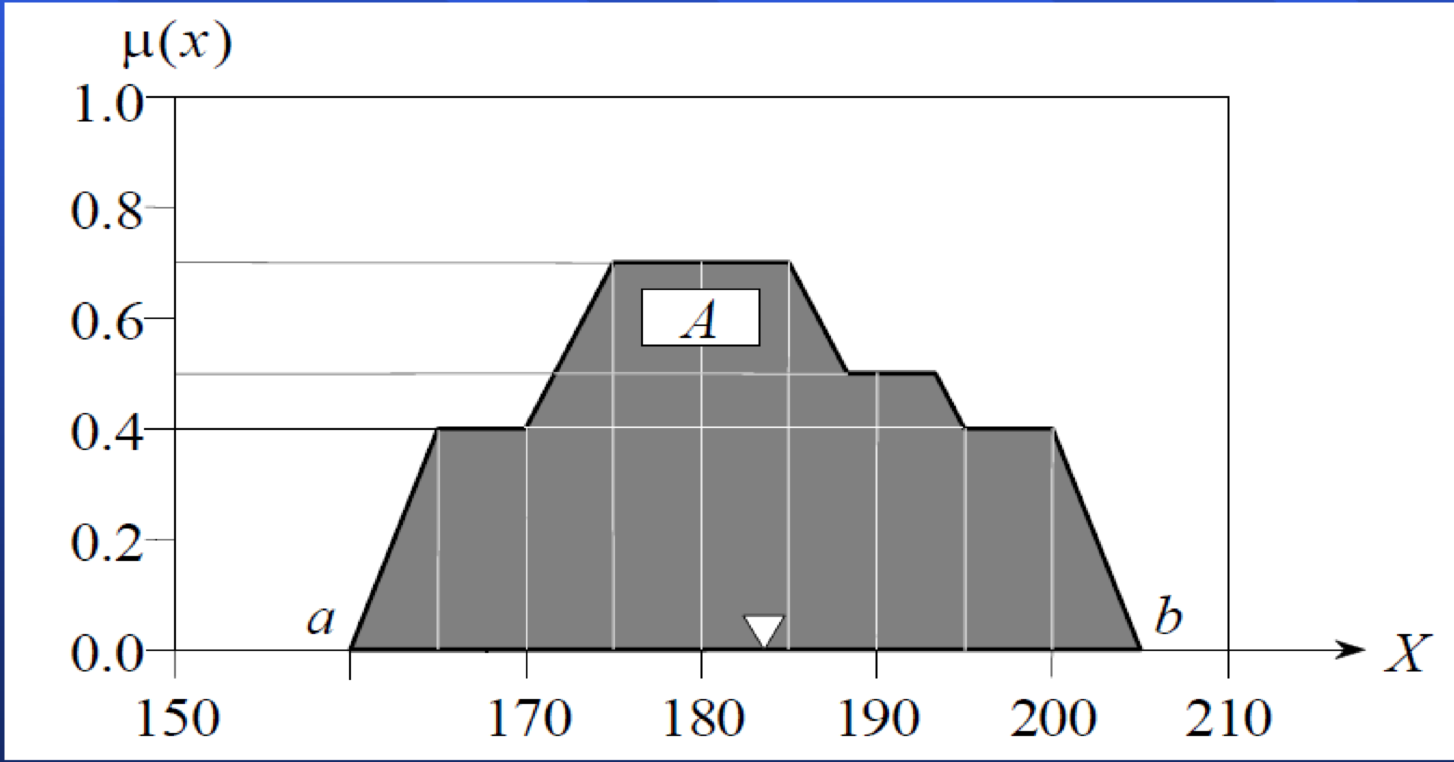
Step 4: Defuzzification

The last step in the fuzzy inference process is defuzzification. Fuzziness helps us to evaluate the rules, but the final output of a fuzzy system has to be a crisp number. The input for the defuzzification process is the aggregate output fuzzy set and the output is a single number.

- There are several defuzzification methods, but probably the most popular one is the **centroid technique**. It finds the point where a vertical line would slice the aggregate set into two equal masses. Mathematically this **centre of gravity (COG)** can be expressed as:

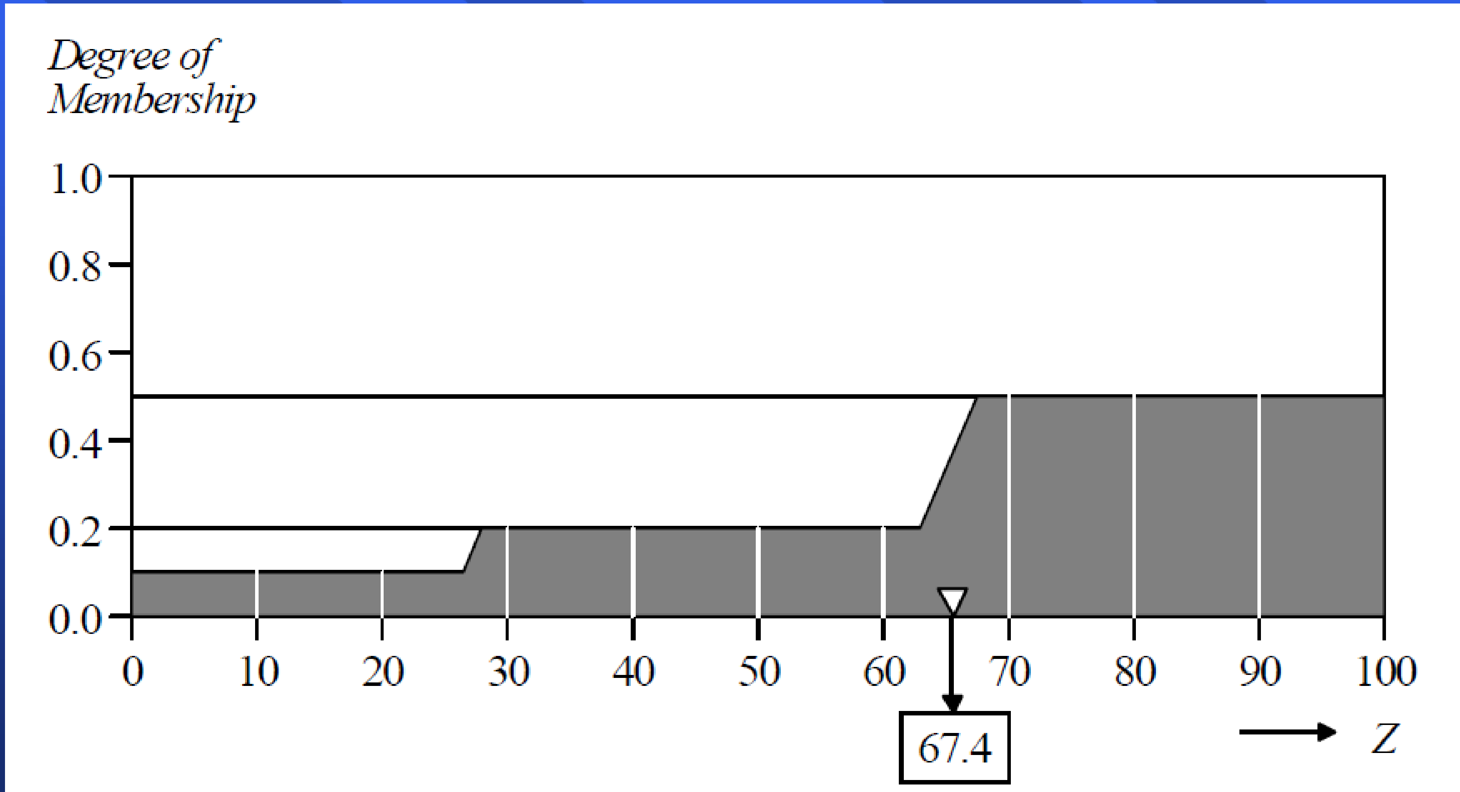
$$COG = \frac{\int_a^b \mu_A(x)\, x\, dx}{\int_a^b \mu_A(x)\, dx}$$

- Centroid defuzzification method finds a point representing the centre of gravity of the fuzzy set, *A*, on the interval, *ab*.
- A reasonable estimate can be obtained by calculating it over a sample of points.



Centre of gravity (COG):

$$COG = \frac{(0+10+20)\times 0.1 + (30+40+50+60)\times 0.2 + (70+80+90+100)\times 0.5}{0.1+0.1+0.1+0.2+0.2+0.2+0.2+0.5+0.5+0.5+0.5} = 67.4$$



Which means the risk involved in our project is 67.4 per cent