Uncertainty and reasoning

Frank Dignum
Department of Information and Computing Sciences
Utrecht University
The Netherlands
Outline

- Fuzzy logic
- Inferences with fuzzy logic
Uncertainty

Let action $A_t = \text{leave for airport} \ t \ \text{minutes before flight}$
Will $A_t$ get me there on time?

**Problems:**

1. partial observability (road state, other drivers' plans, etc.)
2. noisy sensors (traffic reports)
3. uncertainty in action outcomes (flat tire, etc.)
4. immense complexity of modeling and predicting traffic

Hence a purely logical approach either

1. risks falsehood: "$A_{125}$ will get me there on time", or
2. leads to conclusions that are too weak for decision making:

"$A_{125}$ will get me there on time if there's no accident on the bridge and it doesn't rain and my tires remain intact etc etc."

($A_{1440}$ might reasonably be said to get me there on time but I'd have to stay overnight in the airport ...)

Methods for handling uncertainty

- Default or nonmonotonic logic:
  - Assume my car does not have a flat tire
  - Assume $A_{125}$ works unless contradicted by evidence

- Issues: What assumptions are reasonable? How to handle contradiction?

- Rules with fudge factors:
  - $A_{125} \rightarrow_{0.3} \text{get there on time}$
  - $\text{Sprinkler} \rightarrow_{0.99} \text{WetGrass}$
  - $\text{WetGrass} \rightarrow_{0.7} \text{Rain}$

- Issues: Problems with combination, e.g., $\text{Sprinkler}$ implies $\text{Rain}$??

- Fuzzy Logic
  - The road is “busy”
  - Being at the airport 120 minutes before departure is “more than in time”
  - $\text{IF } \text{road(busy)} \text{ and } A_{125} \text{THEN } \text{at_airport(just-in-time)}$

- Probability
  - Model agent's degree of belief
  - Given the available evidence,
  - $A_{25} \text{ will get me there on time with probability 0.04}$
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 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.
Definition

- 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.
Fuzzy, or multi-valued logic, was introduced in the 1930s by Jan Lukasiewicz, a Polish philosopher. While classical logic operates with only two values 1 (true) and 0 (false), Lukasiewicz introduced logic that extended the range of truth values to all real numbers in the interval between 0 and 1.

For example, the possibility that a man 181 cm tall is really tall might be set to a value of 0.86. It is likely that the man is tall. This work led to an inexact reasoning technique often called possibility theory.

In 1965 Lotfi Zadeh, published his famous paper “Fuzzy sets”. Zadeh extended the work on possibility theory into a formal system of mathematical logic, and introduced a new concept for applying natural language terms. This new logic for representing and manipulating fuzzy terms was called fuzzy logic.
Applications

- Automatic parking cars
- Video cameras (noise reduction, steady image)
- Classifying situations in games
- ...

...
Fuzzy logic game
Fuzzy logic is a branch of fuzzy set theory, which deals (as logical systems do) with the representation and inference from knowledge. Fuzzy logic, unlike other logical systems, deals with imprecise or uncertain knowledge.
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.

![Diagram](attachment:image.png)

(a) Boolean Logic.  
(b) Multi-valued Logic.
## Fuzzy Sets

<table>
<thead>
<tr>
<th>Name</th>
<th>Height, cm</th>
<th>Degree of Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Crisp</td>
</tr>
<tr>
<td>Chris</td>
<td>208</td>
<td>1</td>
</tr>
<tr>
<td>Frank</td>
<td>205</td>
<td>1</td>
</tr>
<tr>
<td>John</td>
<td>198</td>
<td>1</td>
</tr>
<tr>
<td>Tom</td>
<td>181</td>
<td>1</td>
</tr>
<tr>
<td>David</td>
<td>179</td>
<td>0</td>
</tr>
<tr>
<td>Mike</td>
<td>172</td>
<td>0</td>
</tr>
<tr>
<td>Bob</td>
<td>167</td>
<td>0</td>
</tr>
<tr>
<td>Steven</td>
<td>158</td>
<td>0</td>
</tr>
<tr>
<td>Bill</td>
<td>155</td>
<td>0</td>
</tr>
<tr>
<td>Peter</td>
<td>152</td>
<td>0</td>
</tr>
</tbody>
</table>
Crisp Vs Fuzzy Sets

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.
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 \( A \):

\[
f_A(x) : X \rightarrow \{0, 1\}, \quad \text{where} \quad f_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}
\]

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 \( A \).
A Fuzzy Set has Fuzzy Boundaries

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

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

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 \( A \).
Fuzzy Set Representation

- 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.

![Diagram of Fuzzy Set Representation](image)
Linguistic Variables and Hedges

- 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.
Linguistic Variables and Hedges

Degree of Membership

Height, cm

Very Short

Average

Short

Tall

Very Tall
### Linguistic Variables and Hedges

<table>
<thead>
<tr>
<th>Hedge</th>
<th>Mathematical Expression</th>
<th>Graphical Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A little</td>
<td>$[\mu_A(x)]^{1.3}$</td>
<td><img src="image" alt="Graphical Representation" /></td>
</tr>
<tr>
<td>Slightly</td>
<td>$[\mu_A(x)]^{1.7}$</td>
<td><img src="image" alt="Graphical Representation" /></td>
</tr>
<tr>
<td>Very</td>
<td>$[\mu_A(x)]^{2}$</td>
<td><img src="image" alt="Graphical Representation" /></td>
</tr>
<tr>
<td>Extremely</td>
<td>$[\mu_A(x)]^{3}$</td>
<td><img src="image" alt="Graphical Representation" /></td>
</tr>
</tbody>
</table>
### Linguistic Variables and Hedges

<table>
<thead>
<tr>
<th>Hedge</th>
<th>Mathematical Expression</th>
<th>Graphical Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very very</td>
<td>$[\mu_A(x)]^4$</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>More or less</td>
<td>$\sqrt{\mu_A(x)}$</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>Somewhat</td>
<td>$\sqrt{\mu_A(x)}$</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>Indeed</td>
<td>$2 [\mu_A(x)]^2$</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td></td>
<td>if $0 \leq \mu_A \leq 0.5$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$1 - 2 [1 - \mu_A(x)]^2$</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td></td>
<td>if $0.5 &lt; \mu_A \leq 1$</td>
<td></td>
</tr>
</tbody>
</table>
Operations of Fuzzy Sets

- **Intersection**
- **Union**
- **Complement**
- **Containment**

[Diagrams showing the operations of fuzzy sets]
Complement

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

The complement of a set is an opposite of this set. For example, if we have the set of tall men, its complement is the set of NOT tall men. When we remove the tall men set from the universe of discourse, we obtain the complement.

If \( A \) is the fuzzy set, its complement \( \neg A \) can be found as follows:

\[
\mu_{\neg A}(x) = 1 - \mu_A(x)
\]
Containment

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

- Similar to a Chinese box, a set can contain other sets. The smaller set is called the *subset*. For example, the set of tall men contains all tall men; very tall men is a subset of tall men. However, the tall men set is just a subset of the set of men. In crisp sets, all elements of a subset entirely belong to a larger set. In fuzzy sets, however, each element can belong less to the subset than to the larger set. Elements of the fuzzy subset have smaller memberships in it than in the larger set.
Intersection

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

In classical set theory, an intersection between two sets contains the elements shared by these sets. For example, the intersection of the set of tall men and the set of fat men is the area where these sets overlap. In fuzzy sets, an element may partly belong to both sets with different memberships.

A fuzzy intersection is the **lower membership** in both sets of each element. The fuzzy intersection of two fuzzy sets $A$ and $B$ on universe of discourse $X$:

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

where $x \in X$.  

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

The union of two crisp sets consists of every element that falls into either set. For example, the union of tall men and fat men contains all men who are tall OR fat.

In fuzzy sets, the union is the reverse of the intersection. That is, the union is the largest membership value of the element in either set. The fuzzy operation for forming the union of two fuzzy sets A and B on universe X can be given as:

$$\mu A \cup B(x) = \max [\mu A(x), \mu B(x)] = \mu A(x) \cup \mu B(x),$$

where $$x \in X$$
Operations of Fuzzy Sets

- Complement

- Containment

- Intersection

- Union
Properties of Fuzzy Sets

- Equality of two fuzzy sets
- Inclusion of one set into another fuzzy set
- Cardinality of a fuzzy set
- An empty fuzzy set
- \( \alpha \)-cuts (alpha-cuts)
Equality

- Fuzzy set $A$ is considered equal to a fuzzy set $B$, IF AND ONLY IF (iff):
  \[ \mu_A(x) = \mu_B(x), \quad \forall x \in X \]

\[ A = 0.3/1 + 0.5/2 + 1/3 \]
\[ B = 0.3/1 + 0.5/2 + 1/3 \]

therefore $A = B$
Inclusion of one fuzzy set into another fuzzy set. Fuzzy set $A \subseteq X$ is included in (is a subset of) another fuzzy set, $B \subseteq X$:

$$\mu_A(x) \leq \mu_B(x), \forall x \in X$$

Consider $X = \{1, 2, 3\}$ and sets $A$ and $B$

$A = 0.3/1 + 0.5/2 + 1/3$;
$B = 0.5/1 + 0.55/2 + 1/3$

then $A$ is a subset of $B$, or $A \subseteq B$
Cardinality

- Cardinality of a non-fuzzy set, Z, is the number of elements in Z. BUT the cardinality of a fuzzy set A, the so-called SIGMA COUNT, is expressed as a SUM of the values of the membership function of A, $\mu_A(x)$:

$$\text{card}_A = \mu_A(x_1) + \mu_A(x_2) + \ldots + \mu_A(x_n) = \sum \mu_A(x_i), \quad \text{for } i=1..n$$

Consider $X = \{1, 2, 3\}$ and sets $A$ and $B$

$A = 0.3/1 + 0.5/2 + 1/3$;
$B = 0.5/1 + 0.55/2 + 1/3$

$\text{card}_A = 1.8$
$\text{card}_B = 2.05$
Empty Fuzzy Set

- A fuzzy set $A$ is empty, IF AND ONLY IF:
  \[ \mu_A(x) = 0, \ \forall x \in X \]

Consider $X = \{1, 2, 3\}$ and set $A$

\[ A = 0/1 + 0/2 + 0/3 \]

then $A$ is *empty*
An $\alpha$-cut or $\alpha$-level set of a fuzzy set $A \subseteq X$ is an ORDINARY SET $A_\alpha \subseteq X$, such that:

$$A_\alpha = \{ \mu_A(x) \geq \alpha, \ \forall x \in X \}.$$

Consider $X = \{1, 2, 3\}$ and set $A$

$$A = 0.3/1 + 0.5/2 + 1/3$$

then $A_{0.5} = \{2, 3\}$,

$A_{0.1} = \{1, 2, 3\}$,

$A_1 = \{3\}$
Assume \( A \) is a fuzzy subset of \( X \):

- the **support** of \( A \) is the crisp subset of \( X \) consisting of all elements with membership grade:
  \[
  supp(A) = \{ x \mid \mu_A(x) > 0 \text{ and } x \in X \}
  \]

- the **core** of \( A \) is the crisp subset of \( X \) consisting of all elements with membership grade:
  \[
  core(A) = \{ x \mid \mu_A(x) = 1 \text{ and } x \in X \}
  \]
Consider two fuzzy subsets of the set $X$,

$X = \{a, b, c, d, e\}$

referred to as $A$ and $B$

$A = \{1/a, 0.3/b, 0.2/c, 0.8/d, 0/e\}$

and

$B = \{0.6/a, 0.9/b, 0.1/c, 0.3/d, 0.2/e\}$
Fuzzy Sets Examples

- **Support:**
  \[ supp(A) = \{a, b, c, d\} \]
  \[ supp(B) = \{a, b, c, d, e\} \]

- **Core:**
  \[ core(A) = \{a\} \]
  \[ core(B) = \{o\} \]

- **Cardinality:**
  \[ card(A) = 1 + 0.3 + 0.2 + 0.8 + 0 = 2.3 \]
  \[ card(B) = 0.6 + 0.9 + 0.1 + 0.3 + 0.2 = 2.1 \]

- **\(a\)-cut:**
  \[ A_{0.2} = \{a, b, c, d\} \]
  \[ A_{0.3} = \{a, b, d\} \]
Fuzzy Sets Examples

- **Complement:**
  \[ A = \{1/a, 0.3/b, 0.2/c, 0.8/d, 0/e\} \]
  \[ \neg A = \{0/a, 0.7/b, 0.8/c, 0.2/d, 1/e\} \]

- **Union:**
  \[ A \cup B = \{1/a, 0.9/b, 0.2/c, 0.8/d, 0.2/e\} \]

- **Intersection:**
  \[ A \cap B = \{0.6/a, 0.3/b, 0.1/c, 0.3/d, 0/e\} \]
A fuzzy rule can be defined as a conditional statement in the form:

\[
\text{IF } x \text{ is } A \\
\text{THEN } y \text{ is } B
\]

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.
Classical Vs Fuzzy Rules

- A classical IF-THEN rule uses binary logic,

  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.

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

- The Mamdani-style fuzzy inference process is performed in four steps:

1. Fuzzification of the input variables
2. Rule evaluation (inference)
3. Aggregation of the rule outputs (composition)
4. Defuzzification.
Mamdani Fuzzy Inference

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

Rule: 1
IF x is A3 OR y is B1 THEN z is C1

Rule: 2
IF x is A2 AND y is B2 THEN z is C2

Rule: 3
IF x is A1 THEN z is C3

Rule: 1
IF project_funding is adequate OR project_staffing is small THEN risk is low

Rule: 2
IF project_funding is marginal AND project_staffing is large THEN risk is normal

Rule: 3
IF project_funding is inadequate THEN risk is high
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.

For $x_1$:
- $\mu(x = A_1) = 0.5$
- $\mu(x = A_2) = 0.2$

For $y_1$:
- $\mu(y = B_1) = 0.1$
- $\mu(y = B_2) = 0.7$
Step 2: Rule Evaluation

- The second step is to take the fuzzified inputs, $\mu_{(x=A_1)} = 0.5$, $\mu_{(x=A_2)} = 0.2$, $\mu_{(y=B_1)} = 0.1$, and $\mu_{(y=B_2)} = 0.7$, 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.
Step 2: Rule Evaluation

**Rule 1:** IF $x$ is $A_3$ (0.0) OR $y$ is $B_1$ (0.1) THEN $z$ is $C_1$ (0.1)

**Rule 2:** IF $x$ is $A_2$ (0.2) AND $y$ is $B_2$ (0.7) THEN $z$ is $C_2$ (0.2)

**Rule 3:** IF $x$ is $A_1$ (0.5) THEN $z$ is $C_3$ (0.5)
Step 2: Rule Evaluation

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

- There are two main methods for doing so:
  - Clipping
  - Scaling
Step 2: Rule Evaluation

- 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 (alpha-cut).

- 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.
Step 2: Rule Evaluation

- 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.
Step 2: Rule Evaluation

Degree of Membership

clipping

scaling
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.
Step 3: Aggregation of the rule outputs

- $z$ is $C1 (0.1)$
- $z$ is $C2 (0.2)$
- $z$ is $C3 (0.5)$
- $\sum$
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.
Step 4: Defuzzification

- 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:

\[
COG = \frac{\int_a^b \mu_A(x) x \, dx}{\int_a^b \mu_A(x) \, dx}
\]
Step 4: Defuzzification

- 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.
Step 4: Defuzzification

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

- Fuzzy sets can be used to model vague terms
- Important where crisp distinctions would lead to arbitrary cut offs
- Inferencing with fuzzy sets possible, but not as “clean” as with first order logic