Introduction to Fuzzy Logic

Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
vladik@utep.edu
http://www.cs.utep.edu/vladik
1. Need for Expert Knowledge

- In some cases, we have a precise knowledge – e.g., an autopilot perfectly pilots a plane.

- In many other cases, we have to rely on expert knowledge.

- So far, computer-based systems have not (yet) replaced skilled medical doctors or even skilled drivers.

- In the ideal world, everyone should go to the best doctor.

- However, in real life, this is not possible.

- It is therefore desirable to have a computer-based tool that contains the knowledge of the best doctors.

- Such tool will help all other doctors make good decisions.
2. Need to Describe Imprecise (Fuzzy) Knowledge

- Many doctors (and experts in general) are absolutely willing to share their knowledge.
- Challenge: this knowledge is often described in terms of imprecise (“fuzzy”) words from natural language.
- A medical doctor can say “if a patient has high fever” or “if a skin mole has an irregular shape”.
- A driver cannot say with what force to hit the brakes if the car 10 m in front slows down from 100 to 90 km/h.
- He/she will say “brake a little bit”.
- We thus need to translate these fuzzy words into computer-understandable language. about the object we can use.
- This is what Zadeh’s fuzzy logic is about.
3. **Toy Example: a Thermostat**

- To illustrate the main idea of fuzzy logic, let us consider a simplified thermostat with a dial.
- Turning the dial to the *left* makes it *cooler*.
- Turning it to the *right* makes it *warmer*.
- We want to reach a comfort temperature $T_0$.
- In other words, we want the difference $x = T - T_0$ to be 0.
- We need to describe, for each $x$, to which angle $u$ we turn the dial: $u = f(x)$. 
4. Thermostat: Rules

- For such an easy system, we do not need any expert to formulate reasonable rules.
- We can immediately describe several reasonable control rules.
- If the room is comfortable, no control is needed.
- So, if the difference $x = T - T_0$ is negligible, then the control $u$ should also be negligible.
- If the room is slightly overheated, cool it a little bit.
- So, if $x$ is positive and small, $u$ must be negative and small.
- If the temperature is a little lower than we would like it to be, then we need to heat the room a little bit.
- In other terms, if $x$ is small negative, then $u$ must be small positive.
5. Thermostat: Rules (cont-d)

• We can formulate many similar natural rules.

• For simplicity, we will restrict ourselves to the above three:
  – if \( x \) is negligible, then \( u \) must be negligible;
  – if \( x \) is small positive, then \( u \) must be small negative;
  – if \( x \) is small negative, then \( u \) must be small positive.
6. Rules: General Case

- Let us denote “negligible” by $N$, “small positive” by $SP$, and “small negative” by $SN$.

- Then, the rules take the following form:

  \[ N(x) \Rightarrow N(u); \quad SP(x) \Rightarrow SN(u); \quad SN(x) \Rightarrow SP(u). \]

- In general, the expert’s knowledge about the dependence of $y$ on $x_1, \ldots, x_n$ can be expressed by rules:

  If $x_1$ is $A_{r_1}$, \ldots, and $x_n$ is $A_{r_n}$, then $y$ is $B_r$.

- Here, $A_{r_i}$ and $B_r$ are words from natural language like “small”, “medium”, “large”, “approximately 1”.

- These rules have the form

  \[ A_{r_1}(x_1) \& \ldots \& A_{r_n}(x_n) \Rightarrow B_r(y). \]
7. Step-by-Step Translation of These Rules

- Our goal is to represent rule bases in precise terms.
- A rule base has a clear structure.
- A rule base consists of rules.
- Each rule, in its turn, is obtained:
  - from properties (expressed by words from natural language)
  - by using logical connectives.
- In view of this structure, it is reasonable to represent the rule base:
  - by first representing the basic elements of the rule base, and then
  - by extending this representation to the rule base as a whole.
8. Step-by-Step Translation (cont-d)

- So, first, we represent the properties $A_{ri}(x_i)$ and $B_r(y)$.
- Second, we represent the logical connectives.
- Third, we use logical connectives to represent each rule.
- Fourth, we combine the representations of different rules into a representation of a rule base.
- As a result of these four steps, we get an advising (expert) system.
- For example, if we apply these four steps to the medical knowledge, we ideally, get a system that
  - given the patient’s symptoms,
  - provides the diagnostic and medical advice.
9. Step-by-Step Translation (cont-d)

- For example, it can say that most probably, the patient has a flu, but it is also possible that he has bronchitis.
- Such an advice, coming from an expert system, is usually used by a specialist to make a decision.
- However, there are situations like automatic control where there is no time to involve a human operator.
- For such control situations, we need an additional, fifth follow-up step: making a decision.
- Let us describe all five steps.
10. First Step: Representing Natural-language Properties

- For properties $A$ like “small”, for some values $x$, we are not 100% sure whether this value is small or not.
- A natural idea is to ask the expert to mark, on a scale from 0 to 1, to what extend the given value $x$ is small.
- We can use another scale – e.g., 0 to 10 – and then divide by 10.
- As a result, for several values $x_i$, we get a degree $A(x_i)$ to which $x_i$ satisfies the property $A$.
- Some experts are not comfortable marking this value.
- Then, we poll the experts and take $A(x_i) = m/n$ if $m$ out of $n$ consider $x_i$ to be, e.g., small.
11. Need for Interpolation

- To get the values \( A(x) \) for all other \( x \), we use interpolation.

- The resulting function is called a *membership function* or a *fuzzy set*.

- The simplest is linear interpolation.

- Let us consider the word “negligible”.

- The only case when we are 100% sure that \( x \) is negligible is when \( x = 0 \).

- So, we have \( N(0) = 1 \).

- Usually, we also know the value \( \Delta > 0 \) after which the difference in temperatures is no longer negligible.

- For example, for a thermostat that controls the room’s temperature, we can take \( \Delta = 10^\circ F \).
12. Interpolation (cont-d)

- This means that $N(x) = 0$ for $x \geq \Delta$ and for $x < -\Delta$.
- We know the value of the function $N(x)$ for $x \leq -\Delta$, for $x = 0$, and for $x \geq \Delta$.
- For $x \in (-\Delta, 0)$, we get the expression
  
  $$N(x) = 1 + \frac{x}{\Delta}.$$ 

- For $x \in (0, \Delta)$, we get the expression $N(x) = 1 - \frac{x}{\Delta}$.
- The graph of $N(x)$ is a triangle.
- Such functions are called *triangular*. 
13. Need to Combine Fuzzy Degrees

- Conclusions based on expert knowledge often take into account *several* expert statements.

- Our degree of confidence in such a conclusion is thus equal to our degree of confidence that:
  - the first of used statements is true *and*
  - the second used statement is true, etc.

- In other words,
  - in addition to the expert’s degrees of confidence in their statements $S_1, \ldots, S_n$,
  - we also need to estimate the degrees of confidence in “and”-combinations $S_i \& S_j$, $S_i \& S_j \& S_k$, etc.

- In the ideal world, we can ask the experts to estimate the degree of confidence in each such combination.

- However, this is not realistically possible.
14. Need to Combine (cont-d)

- Problem: for $n$ original statements, there are $2^n - 1$ such combinations.
- Indeed, combinations are in 1-1 correspondence with non-empty subsets of the set of $n$ statements.
- Already for reasonable $n = 30$, we get an astronomical number $2^{30} \approx 10^9$ combinations.
- There is no way that we can ask a billion questions to the experts.
- We cannot elicit the expert’s degree of confidence in “and”-combinations directly from the experts.
- So, we need to estimate these degrees based on the experts’ degrees of confidence in each statement $S_i$. 
15. Need to Combine (cont-d)

• In other words, we need to be able:
  – to combine the degrees of confidence \( a \) and \( b \) of statements \( A \) and \( B \)
  – into an estimate for degree of confidence in the “and”-combination \( A \& B \).

• The algorithm for such combination is called an “and”-operation or, for historical reasons, a \( t \)-norm.

• The result of applying this combination algorithm to numbers \( a \) and \( b \) will be denoted \( f_\&(a, b) \).
16. How to Combine Fuzzy Degrees?

- Which operation $f_\& (a, b)$ should we choose?
- First, since $A \& B$ means the same as $B \& A$, it is reasonable to require that the resulting estimates coincide:
  \[ f_\&(a, b) = f_\&(b, a). \]
- Since $A \&(B \& C)$ means the same as $(A \& B) \& C$, we must have
  \[ f_\&(a, f_\&(b, c)) = f_\&(f_\&(a, b), c). \]
- Since $A \& B$ is a stronger statement than each of $A$ and $B$ (it implies both $A$ and $B$),
  \[ f_\&(a, b) \leq a \text{ and } f_\&(a, b) \leq b. \]
17. How to Combine Fuzzy Degrees (cont-d)

• Finally:
  – if our degree of confidence in one or both of the statements $A$ and $B$ increases,
  – the resulting degree of confidence in $A \& B$ should also increase – or at least remain the same:
    
    if $a \leq a'$ and $b \leq b'$, then $f_{\&}(a, b) \leq f_{\&}(a', b')$.

• There are many such operations: $\min(a, b)$, $a \cdot b$, etc.

• We need to select the one that best represents human reasoning in a given knowledge domain.

• We can also require that since $A \& A$ means the same as $A$, it is reasonable to require that $f_{\&}(a, a) = a$.

• Then, the only possible t-norm is $\min(a, b)$.

• Similarly, we can define “or”-operations (aka t-conorms).
18. Historical Comment

- Historically the first determination of “and”-operation was done at Stanford.

- Researchers designed an expert system MYCIN for diagnosing rare blood diseases.

- The results were good, so they thought that they have discovered general laws of human reasoning.

- However, when they applied the same laws to geophysics, they failed.
19. Historical Comment (cont-d)

- Indeed, reasoning is different in these two areas.
- In medicine, we need to be absolutely sure before we recommend, e.g., surgery – else we hurt the patient.
- In geophysics, if there is a good chance to find oil, we start digging.
- If we wait until we are absolutely sure, competitors will be there first.
- So, now we know that in different domains, different t-norms are needed.
20. Back to Fuzzy Rules and Toy Example

- Let us go back to our rules
  
  \[ \text{N}(x) \Rightarrow \text{N}(u); \quad \text{SP}(x) \Rightarrow \text{SN}(u); \quad \text{SN}(x) \Rightarrow \text{SP}(u). \]

- For each input \( x \), the control \( u \) is reasonable \( R(x, u) \) if one of the rules if applicable:
  
  – either the 1st rule is applicable, so \( \text{N}(x) \) and \( \text{N}(u) \),
  – or the 2nd rule is applicable, so \( \text{SP}(x) \) and \( \text{SN}(u) \),
  – or the 3rd rule is applicable, so \( \text{SN}(x) \) and \( \text{SP}(u) \):

  \[ R(x, u) \iff (\text{N}(x) \& \text{N}(u)) \lor (\text{SP}(x) \& \text{SN}(u)) \lor (\text{SN}(x) \& \text{SP}(u)). \]

- From experts, we get the values \( \text{N}(x) \), \( \text{N}(u) \), etc.

- Based on experts, we select “and”- and “or”-operations.

- Thus, for each \( x \) and \( u \), we get \( R(x, u) \) as

  \[ f_\lor(f_\&(\text{N}(x), \text{N}(u)), f_\&(\text{SP}(x), \text{SN}(u)), f_\&(\text{SN}(x), \text{SP}(u))). \]
21. Numerical Example

- Let us assume that all three membership functions are piece-wise linear.

- Specifically, let us assume that they are described by the following graph:
22. Numerical Example (cont-d)

• What is the degree of confidence $\mu_C(4, -2)$ that for $x = 4^\circ$ the control $u = -2^\circ$ is reasonable?

• According to our formulas, let us first compute the values of the membership functions.

• By linear interpolation, we can find the analytical formulas for these membership functions:

• The term “negligible” is described by the following formulas:
  
  • $\mu_N(x) = 1 + x/5$ for $-5 \leq x \leq 0$;
  • $\mu_N(x) = 1 - x/5$ for $0 \leq x \leq 5$;
  • $\mu_N(x) = 0$ for all other $x$. 

23. Numerical Example (cont-d)

- The term “small positive” is described by the following formulas:
  - $\mu_{SP}(x) = \frac{x}{5}$ for $0 \leq x \leq 5$;
  - $\mu_{SP}(x) = 2 - \frac{x}{5}$ for $5 \leq x \leq 10$;
  - $\mu_{SP}(x) = 0$ for all other $x$.

- The term “small negative” is described by the following formulas:
  - $\mu_{SN}(x) = 2 + \frac{x}{5}$ for $-10 \leq x \leq -5$;
  - $\mu_{SN}(x) = -\frac{x}{5}$ for $-5 \leq x \leq 0$;
  - $\mu_{SN}(x) = 0$ for all other $x$. 

24. Numerical Example (cont-d)

- If we use $f_\& (a, b) = \min(a, b)$ and $f_\lor (a, b) = \max(a, b)$, then we get $\mu_C(4, -2) = \max(d_1, d_2, d_3)$, where

$$d_1 = \min(\mu_N(4), \mu_N(-2)); \quad d_2 = \min(\mu_{SP}(4), \mu_{SN}(-2));$$

$$d_3 = \min(\mu_{SN}(4), \mu_{SP}(-2)).$$

- Here, $\mu_N(4) = 0.2$, $\mu_N(-2) = 0.6$, $\mu_{SP}(4) = 0.8$, $\mu_{SN}(-2) = 0.4$, and $\mu_{SN}(4) = \mu_{SP}(-2) = 0$.

- Hence, $d_1 = \min(0.2, 0.6) = 0.2$, $d_2 = \min(0.8, 0.4) = 0.4$, $d_3 = \min(0, 0) = 0$, and

$$\mu_C(4, -2) = \max(0.2, 0.4, 0) = 0.4.$$
25. **Defuzzification**

- For each $x$ and for each $u$, we get a degree $\mu(u) = R(x,u)$ to which $u$ is reasonable.
- In other words, we get a fuzzy set of possible controls.
- For automatic control, we need to “defuzzify” this into a single value $\bar{u}$.
- For each expert’s opinion $u$, we want to have $\bar{u} - u \approx 0$.
- The vector formed by the differences should be as close to 0 as possible, so we minimize $\sum (\bar{u} - u)^2$. 
26. Defuzzification (cont-d)

- For each $u$, the degree $\mu(u)$ is proportional to the number of experts who believe that $u$ is reasonable.
- So, the minimized sum becomes $\sum_u \mu(u) \cdot (\bar{u} - u)^2$, i.e., $\int (\bar{u} - u) \cdot \mu(u) \, du$.
- To minimize this expression, we differentiate it relative to $\bar{u}$ and equate the result to 0; thus
  $$\bar{u} = \frac{\int u \cdot \mu(u) \, du}{\int \mu(u) \, du}.$$  
- This is known as centroid defuzzification.
- The resulting fuzzy control has indeed been very successful in many applications, from rice cookers to trains.
27. Fuzzy Control: Summarizing

- We start with the if-then expert rules of the type “If \( x \) is small, and ..., then \( y \) is small”:

\[
A_{r1}(x_1) \& ... \& A_{rn}(x_n) \Rightarrow B_r(y).
\]

- Here \( x_1, \ldots, x_n \) are inputs, and \( A_{ri} \) and \( B_r \) are words that describe properties of inputs and output.

- For each words \( w \) used in these rules, we pick several values \( x^{(1)}, \ldots, x^{(k)} \).

- We determine the degrees of confidence \( \mu_w(x^{(1)}), \ldots, \mu_w(x^{(n)}) \) that these values satisfy the property \( w \).

- Then, we use some interpolation technique to determine the membership functions \( \mu_w(x) \) for all \( x \).

- We choose “and”- and “or”-operations \( f_\& (a, b) \) and \( f_\lor (a, b) \).
28. Fuzzy Control: Summarizing (cont-d)

- For each rule \( r \), and for each possible values of input and output, we compute the firing degrees
  \[
  d_r(x_1, \ldots, x_n, y) = f_\&(\mu_{r1}(x_1), \ldots, \mu_{rn}(x_n), \mu_r(y)).
  \]

- Then we compute the membership function for control
  \[
  \mu_C(x_1, \ldots, x_n, y) = f_\lor(d_1(x_1, \ldots, x_n, y), \ldots, d_R(x_1, \ldots, x_n, y)).
  \]

- For every input \( x_1, \ldots, x_n \), \( \mu_C(y) \) is the degree of confidence that \( y \) is a reasonable control.

- If needed, we can then get a single recommended control value \( \bar{y} \):
  \[
  \bar{y} = \frac{\int y \cdot \mu_C(y) \, dy}{\int \mu_C(y) \, dy}.
  \]
29. Fuzzy Computations: A Problem

\[
\begin{align*}
\mu_1(x_1) \\
\mu_2(x_2) \\
\vdots \\
\mu_n(x_n)
\end{align*}
\]

\[
f : \mu = f(\mu_1, \ldots, \mu_n)
\]

- **Given:** an algorithm \( y = f(x_1, \ldots, x_n) \) and \( n \) fuzzy numbers \( \mu_i(x_i) \).

- **Compute:** \( \mu(y) = \max_{x_1, \ldots, x_n: f(x_1, \ldots, x_n) = y} \min(\mu_1(x_1), \ldots, \mu_n(x_n)) \).

- **Motivation:** \( y \) is a possible value of \( Y \leftrightarrow \exists x_1, \ldots, x_n \) s.t. each \( x_i \) is a possible value of \( X_i \) and \( f(x_1, \ldots, x_n) = y \).

- **Details:** “and” is min, \( \exists \) (“or”) is max, hence

\[
\mu(y) = \max_{x_1, \ldots, x_n} \min(\mu_1(x_1), \ldots, \mu_n(x_n), t(f(x_1, \ldots, x_n) = y)).
\]
30. Fuzzy Computations: Reduction to Interval Computations

- **Problem (reminder):**
  - *Given:* an algorithm \( y = f(x_1, \ldots, x_n) \) and \( n \) fuzzy numbers \( X_i \) described by membership functions \( \mu_i(x_i) \).
  - *Compute:* \( Y = f(X_1, \ldots, X_n) \), where \( Y \) is defined by Zadeh’s extension principle:
    \[
    \mu(y) = \max_{x_1, \ldots, x_n : f(x_1, \ldots, x_n) = y} \min(\mu_1(x_1), \ldots, \mu_n(x_n)).
    \]

- **Idea:** represent each \( X_i \) by its \( \alpha \)-cuts
  \[
  X_i(\alpha) = \{ x_i : \mu_i(x_i) \geq \alpha \}.
  \]

- **Advantage:** for continuous \( f \), for every \( \alpha \), we have
  \[
  Y(\alpha) = f(X_1(\alpha), \ldots, X_n(\alpha)).
  \]

- **Resulting algorithm:** for \( \alpha = 0, 0.1, 0.2, \ldots, 1 \) apply interval computations techniques to compute \( Y(\alpha) \).