Wiener’s Conjecture About Transformation Groups Helps Predict Which Fuzzy Techniques Work Better

Francisco Zapata\textsuperscript{1}, Olga Kosheleva\textsuperscript{2}, and Vladik Kreinovich\textsuperscript{3}

\textsuperscript{1}Research Institute for Manufacturing & Engineering Systems
Departments of \textsuperscript{2}Teacher Education and \textsuperscript{3}Computer Science
University of Texas at El Paso, El Paso, Texas 79968, USA
fazg74@gmail.com, olgak@utep.edu, vladik@utep.edu
1. Formulation of the Problem

- Often, application succeeds only when we select specific fuzzy techniques (t-norm, membership f-n, etc.).
- In different applications, different techniques are the best.
- How to find the best technique?
- Exhaustive search of all techniques is not an option: there are too many of them.
- We need to come up with a narrow class of promising techniques, so that trying them all is realistic.
- We show that transformation groups – motivated by N. Wiener’s conjecture – lead to such a narrowing.
- This conjecture was, in its turn, motivated by observations about human vision.
2. Wiener’s Conjecture: Reminder

• The closer we are to an object, the better we can determine its shape.

• Experiments show that there are distinct phases in this determination.

• When the object is very far, all we see is a formless blurb.

• In other words, objects obtained from other by arbitrary smooth transformations cannot be distinguished.

• When the object gets closer, we can detect whether it is smooth or has sharp angles.

• We may see a circle as an ellipse, a square as a rhombus (diamond).

• At this stage, images obtained by a projective transformation are indistinguishable.
3. Wiener’s Conjecture (cont-d)

- When the object gets closer, we can detect which lines are parallel but we may not yet detect the angles.

- For example, we are not sure whether what we see is a rectangle or a parallelogram.

- This stage corresponds to affine transformation.

- Then, we have a stage of similarity transformations – when we detect the shape but cannot yet detect its size.

- Finally, when the object is close enough, we can detect both its shape and its size.

- Each stage can be thus described by an appropriate transformation group (see a formal description below).
4. Wiener’s Conjecture: Result

- Humans result from billions of years of evolution. So, Wiener conjectured that:
  - if there was a group intermediate between, e.g., all projective and all continuous transformations,
  - our vision mechanism would have used it.

- Thus, according to the 1940s Wiener’s conjecture, such intermediate groups are not possible.

- In the 1960s, Wiener’s conjecture was proven.

- In the 1-D case, projective transformations are simply fractionally linear, and affine are simply linear.

- Thus, any group containing all 1-D linear transformation is:
  - either the group of all fractionally-linear transf.
  - or the group of all transformations.
5. How Wiener’s Conjecture Helps: General Idea

- Fuzzy degrees are not uniquely determined.
- Different elicitation techniques lead, in general, to different values.
- Sometimes, different scales are related by a linear transformation, sometimes by a non-linear one.
- In practice, we want a description with finitely many parameters.
- Thus, we want a finite-dimensional transformation group.
- Due to the above result, all such transformations are fractionally linear.
- We show that this can explain why some t-norms, membership functions, etc., are empirically more successful.
6. Different Assignment Procedures Are In Use

- Intelligent systems use several different procedures for assigning numeric values that describe uncertainty.

- The same expert’s degree of uncertainty that he expresses, e.g., by the expression “for sure”, can lead:
  - to 0.9 if we apply one procedure, and
  - to 0.8 if another procedure is used.

- 1 foot and 12 inches describe the same length, but in different scales.

- We can say that 0.9 and 0.8 represent the same degree of certainty in two different scales.

- Some scales are different even in the fact that they use an interval different from $[0, 1]$ to represent uncertainty.

- For example, the famous MYCIN system uses the interval $[-1, 1]$. 
7. Transformations Between Reasonable Scales

• Let $F$ denote the class of reasonable transformations of degrees of uncertainty. If:
  – a function $x \rightarrow f(x)$ is a reasonable transformation from a scale $A$ to some scale $B$, and
  – a function $y \rightarrow g(y)$ is a reasonable transformation from $B$ into some other scale $C$,
  – then the transformation $x \rightarrow g(f(x))$ from $A$ to $C$ is also reasonable.

• In other words, the class $F$ of all reasonable transformations must be closed under composition. Also:
  – if $x \rightarrow f(x)$ is a reasonable transformation from a scale $A$ to scale $B$,
  – then the inverse function is a reasonable transformation from $B$ to $A$.

• Thus, $F$ must be a transformation group.
8. Examples of Reasonable Transformations

• A natural method to assign a truth value \( t(S) \) to a statement \( S \) is to ask several experts and take

\[
t(S) = \frac{N(S)}{N}.
\]

• The more expert we ask, the more reliable is this estimate.

• However, in the presence of Nobelists, experts may say nothing or follow the majority.

• After we add \( M \) experts who do not answer anything and \( M' \) who follow the majority, we get

\[
t' = \frac{N(S) + M'}{N + M + M'} = \frac{N \cdot t + M'}{N + M + M'} = a \cdot t + b.
\]

• The transformation from an old scale \( t(S) \) to a new scale \( t' \) is a linear function.
9. Definition and Main Result

- By a *rescaling* we mean a strictly increasing continuous function \( f \) that is defined on an interval \([a, b] \subseteq \mathbb{R}\).

- Suppose a set \( F \) of rescalings is a connected Lie group which contains, for all \( N, M, M' \geq 0 \), a transformation
  \[
  t \mapsto \frac{N \cdot t + M'}{N + M + M'}.
  \]

- Elements of this set \( F \) will be called *reasonable transformations*.

- *Result:* Every reasonable transformation \( f(x) \) is fractionally linear:
  \[
  f(x) = \frac{a \cdot x + b}{c \cdot x + d}.
  \]
10. Normalizations

- To compare degrees in different scales, we need to “normalize” them.
- Often, there exists an alternative $a$ for which we are absolute sure that it is not possible: $\mu(a) = 0$.
- It is natural to require that this value 0 should remain the same after the “normalization” transformation.
- By a normalization we mean a reasonable transformation $f(x)$, for which $f(0) = 0$.
- **Result:** Every normalization has the form $f(x) = \frac{k \cdot x}{1 + d \cdot x}$.
- **Comment.** This class includes the most widely used linear normalization $\mu'(x) = \frac{\mu(x)}{\max_y \mu(y)}$. 
11. Selecting Membership Functions

- Suppose that we have a fuzzy notion like “small”.
- For $x = 0$, we are sure that it is small.
- Until we reach large values, the bigger $x$, the less we are certain that $x$ is small.
- There are thus two ways to represent our uncertainty:
  - we can use the value of a membership function $\mu(x)$;
  - we can also use the value $x$ itself – since the larger $x$, the larger our uncertainty.
- The transformation between these scales must be reasonable.
- So, a membership function must be piecewise fractionally linear.
- Triangular and trapezoid functions – most efficient – are indeed examples of such functions.
12. “And”-Operations

- When we communicate, we often make implicit assumptions.
- For example, when we ask a doctor to estimate the efficiency of a certain treatment \( t \):
  - the doctor may interpret it as estimating the proportion of patients who gets well,
  - or as proportion of patients who got well because of \( t \) – and not by itself.
- In other words, we estimate either \( d(W) \) or \( d(W & T) \), where \( T \) means that the treatment worked.
- It makes sense to require that the transformation \( d(W) \rightarrow d(W & d(T)) \) is reasonable.
- In fuzzy logic, we estimate \( d(W & T) \) as \( f_{\&}(d(W), d(T)) \).
- So, we require that \( a \rightarrow f_{\&}(a, b) \) is reasonable for all \( b \).

• **Reminder:** we require that $a \rightarrow f_\&(a, b)$ is reasonable for all $b$.

• **Result:** all such t-norms are either $f_\&(a, b) = \min(a, b)$ or
  \[
  f_\&(a, b) = \frac{a \cdot b}{k + (1 - k) \cdot (a + b - a \cdot b)}.
  \]

• For t-conorms (“’or’”-operations), we similarly get $f_\lor(a, b) = \max(a, b)$ or
  \[
  f_\lor(a, b) = \frac{a + b + (k - 1) \cdot a \cdot b}{1 + k \cdot a \cdot b}.
  \]

• Most widely used $\min$, $\max$, $a \cdot b$, and $a + b - a \cdot b$ are indeed examples of such operations.
14. Negation Operations etc.

- A negation operation can be defined as a function $f_\neg(x)$ which extends the usual negation from $\{0, 1\}$ to $[0, 1]$: 
  $$f_\neg(0) = 1 \text{ and } f_\neg(1) = 0.$$ 

- We can express our uncertainty in a statement $A$:
  - either by our degree of belief $d(A)$ in $A$,
  - or by our degree of belief $d(\neg A) = f_\neg(d(A))$ in $\neg A$.

- The transformation $f_\neg(x) : d(A) \rightarrow d(\neg A)$ is reasonable, so $f_\neg(x) = \frac{1-x}{1+k \cdot x}$.

- For $k = 0$, we get the original negation $f_\neg(x) = 1 - x$.

- For $k \neq 0$, we get Sugeno operations which are known to be a good fit for human reasoning.

- Similarly, we explain which defuzzification to use, why sigmoid activation functions are efficient, etc.
15. Acknowledgment

This work was supported in part by the National Science Foundation grants:

- HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and
- DUE-0926721.