Computing with
Words and Granules

Ronald R. Yager
Machine Intelligence Institute
Iona College

yager@panix.com
Human Sourced Information is Soft Information

Based on Perceptions
Human Focused
Linguistically Expressed
Imprecision
Uncertain

Soft Information is Granular in Nature
Granular Probabilities

- Close to half
- About 0.2
- Large

0 0.2 0.5 1
Granular Computing

Is a Collection of Set Based Technologies that Allow for the Formal Representation and Manipulation of Human Focused Soft Linguistic Concepts
Granular Computing Technologies

Fuzzy Set Theory
Dempster-Shafer Belief Structures
Rough Sets
Probabilistic Reasoning
Possibility Probability Granules
SITUATION ASSESSMENT REQUIRED FOR DECISION MAKING AND OTHER TASKS OFTEN INVOLVES THE FUSION OF INFORMATION FROM MULTIPLE SOURCES
On the Linguistic Presentation of Information for Human Comprehension
Man-Machine Communications

• Information fusion systems manipulate mathematical representations of information

• Result of fusion in formal mathematical structure (fuzzy set/ probability distribution)

• Human Clients of fusion systems prefer information expressed in natural language like statements
RETRANSLATION

Process of converting the formal mathematical representations resulting from information fusion into natural language like statements understandable to the human client.
Point of Departure

• Result of some fusion process
  \( V \text{ is } A \) A fuzzy subset of \( X \)

• \( L \) Natural language vocabulary

• \( F \): Representations of NL terms in \( L \) as fuzzy subsets of \( X \)

Task
Replace \( V \text{ is } A \) by \( V \text{ is } L \)
\( L \) is element from \( L \)
Implementation of Retranslation

• Substitute the proposition $V \text{ is } F$ for $V \text{ is } A$ where $F$ is some element from $F$.

• Express the output as $V \text{ is } L$ where $L$ is the natural language term associated with $F$.

• The key issue is the substitution of $V \text{ is } F$ for $V \text{ is } A$. 
• Decide upon the best retranslation using satisfaction to some criteria

• Criteria can be user selected or default

• Types of Criteria:
  Faithfulness of $F$ to $A$.
  Features of the replacement set $F$
  Some external goal or purpose
Truthfulness of a Retranslation

• **Don't want result that is not true!**

• Normally most important criteria

• If fusion result is
  
  John's age is \( \{x | x = 1 \text{ to } 19\} \)

  Don't to retranslate this into
  
  John is **old**.

• **True:** "John is less than 20"

• **False:** John is over 50

• **Not Justified:** John is 13
The entailment principle provides a tool for addressing this issue.

The Entailment Principle

From the proposition \( V \text{ is } A \) we can truthfully infer \( V \text{ is } F \) where \( A \subseteq F \).
Containment of Fuzzy Sets

- **Zadeh Definition**: \( A \subseteq F \) if \( A(x) \leq F(x) \) for all \( x \in X \)

- **Crisp definition**: Either \( A \) is contained in \( F \) or \( A \) not contained in \( F \)

- **Crisp interpretation of the entailment principle**.
Softer definition

• Deg(A ⊆ F) the degree to which A is a subset of F.
  \[ \text{Deg}(A \subseteq F) \in [0, 1]. \]

• Using softer definition in entailment principle we can get a degree to which the statement \( V \ is \ A \) entails the statement \( V \ is \ F \).

Associates with each \( F \in \mathcal{F} \) a degree to which it is valid to infer \( V \ is \ F \) from \( V \ is \ A \)
General Definition of Containment

• $\text{Deg}(A \subseteq F) = \text{Min}_x[I(A(x), F(x))]$

  $I$ is called a ply operator

• $I: [0, 1] \times [0, 1] \rightarrow [0, 1]$ such that
  1. $I(a_1, b) \geq I(a_2, b)$ if $a_2 \geq a_1$
  2. $I(a, b_1) \geq I(a, b_2)$ if $b_1 \geq b_2$
  3. $I(0, b) = 1$
  4. $I(a, b) = 1$ if $b \geq a$
Example:

\[ I(a, b) = \min[1, 1 + (b - a)] \]

Allow a range \( \Delta \) in which we neglect any difference.

\[ I(a, b) = \min[1, 1 + \Delta + (b - a)] \]
Closeness of A and F

• Useful criteria to determine appropriateness of replacing $V \text{ is } A$ with $V \text{ is } F$ is that $F$ is close to $A$.

• Not a required but may be desired by client

• Validity doesn't imply closeness
• Need relationship *Close* between two fuzzy subsets

• The fundamental unit in discussing the closeness between fuzzy sets is elementary difference

\[ \Delta_j = |A(x_j) - F(x_j)| \]

• From this we obtain

\[ Close(A, F) = 1 - \left( \frac{1}{n} \sum_j \Delta_j^\alpha \right)^{1/\alpha} \]
In selecting the retranslation we may want to consider some features associated solely with the retranslated value $V$ is $F$. 

Fuzziness

• Related to type of boundary distinguishing membership from non-membership in F.

• Saying that John's age is between 22 and 28 is very crisp

• Saying that he is about 25 is more fuzzy

• Intuitively fuzziness is related to the lack of distinction between a set and its negation.
Properties of Measure of Fuzziness

• $\text{FUZZ}(F) \in [0, 1]$,

• $\text{FUZZ}(F)$ attains its minimum value, if $F$ is a crisp set, $F(x) \in \{0, 1\}$

• $\text{FUZZ}(F)$ attains maximum value for the fuzzy subset $F(x) = 1/2$ for all $x$

• $\text{FUZZ}(F) \geq \text{FUZZ}(F^*)$ if $F^*$ is sharpened version of $F$
  
  $F^*(x) \geq F(x)$ when $F(x) \geq 1/2$

  $F^*(x) \leq F(x)$ if $F(x) \leq 1/2$
• Measure of fuzziness based on the lack of distinction between $F$ and $\bar{F}$

• The closer $F$ and $\bar{F}$ the more fuzzy.

• $\text{FUZZ}(F) = 1 - \text{Ave}(|F(x_j) - \bar{F}(x_j)|$

• Definition satisfies above conditions .

This type criteria just depends upon $F$

Doesn't involve the relationship between $F$ and $A$.

Just about features of $F$. 
Specificity

• Related to amount of information in statement.

• Statement Rachel is 32 is more informative than statement "she is in her early thirties".

• Specificity measures degree to which subset contains one and only one element.
Fuzziness and Specificity

• **Fuzziness** related to type of boundary

• **Specificity** related to the cardinality of the set.

Statement can be crisp & non–specific,

John's over 10 years old

non–specific but crisp

"close to 20" is specific but fuzzy.
Properties of a Measure of Specificity

• $Sp(F) \in [0, 1]$

• $Sp(F) = 1$ if there exists one $x^*$ with $F(x^*) = 1$ and $F(x) = 0$ for all other $x$

• $Sp(F) = 0$ if $F(x) = 0$ for all $x$

• $Sp(F_1) \leq Sp(F_2)$ if $F_1$ & $F_2$ are normal and $F_1(x) \geq F_2(x)$ for all $x$

• $Sp(F_1) \geq Sp(F_2)$ if $F_1$ & $F_2$ are non-null & crisp with $\text{Card}(A_1) \leq \text{Card}(A_2)$
Specificity can be measured as difference between the largest membership grade and the average of membership grade

$$Sp(F) = \max_x[F(x)] - \avex[F(x)]$$
Providing Retranslations that Give Desired Perceptions

In some instances we have an agenda when expressing the fused value

Convey a particular perception

Let P be a fuzzy subset corresponding to desired perception

\[ \text{Perp}(P/F) = 1 - \frac{\text{Card}(F \cap \overline{P})}{\text{Card}(F)} \]

Perception of P using term corresponding to F
Selecting Best Retranslation

• \( V \) is \( A \) Fusion output,

• \( F \) Fuzzy subsets corresponding to available linguistic terms.

• \( C = \{C_1, ..., C_q\} \) collection of criteria of interest to the user

• For \( F \in F \) we have \( C_j(F) \), the degree to which \( F \) satisfies \( C_j \)
Multi-criteria decision Problem

- For each $F$ calculate overall satisfaction to the criteria,
  \[ Val(F) = Agg(C_1(F), ..., C_q(F)). \]

- Select the $F$ with largest satisfaction.

- Use the OWA Operator to model multi-criteria decision function
• Associate with $C$ a measure: $\mu: 2^C \rightarrow [0, 1]$

• $\mu(G)$ is degree of acceptability of a term that satisfies the subset $G$ of criteria.

• $\mu(\emptyset) = 0, \mu(C) = 1$ & $\mu(G_2) \geq \mu(G_1)$ if $G_2 \supset G_1$.

• $H_j$: subset of $j$ criteria where $F$ scores best.

• $m_j = \mu(H_j) - \mu(H_{j-1})$

• $\text{Val}(F) = \sum_j m_j C_{\text{index}(j)}(F)$

  $C_{\text{index}(j)}(F)$ is $j$th best score
Disinformation Dissemination

• At times we have an agenda when expressing the fused value

• Convey a particular perception

• Denability/Accountability

• Imbed true value in larger granular that gives desired perception without being false
The Entailment Principle

- Basic Inference rule in fuzzy logic

- The principle states:
  From knowledge that $V$ is $A$ we can infer $V$ is $F$ for any $A \subseteq F$

  ($A \subseteq F$ if $A(x) \leq F(x)$ for all $x$)

- Example  Knowing that John's age is between $25$ and $35$ allows us to infer that John's age is between $10$ and $50$

- Related to logic inference $a \Rightarrow a \text{ or } b$
In yellow circle must be in blue circle
REDUCTION

In orange saying it is in blue
REDUCTION

• From less precise to more precise

• From knowing that John's is between 25 and 35 years old we infer that John's age is 30

• NOT SOUND  Can lead to wrong inferences

• Used in human reasoning: Default reasoning

• Pragmatic & Often useful
  Mary lives in China ⇒ Mary is Chinese
Learning From Imprecise Granular Observations Using Trapezoidal Fuzzy Set Representations
• Huge amounts of linguistic information available

• Generally granular and imprecise

• Objective to use this information for learning

• Must represent information in terms of some formal granular object

• Representation step affords user freedom in selection of the representing granule
Major Considerations in Granule Selection

• **Cointention**
  
  Ability of representing object to convey the meaning of the information it is being used to represent

• **Functionality**
  
  Ease with which the representing object can be manipulated in the context of the application
Trapezoidal Fuzzy Set Representation
Rich Representational Capability

- $a = b$ and $c = d \implies$ Interval Type

- $b = c \implies$ Triangular Type

- $a = b = c = d \implies$ Singleton point case

- $a = b \implies$ Book end type
• Membership grade easily obtained from the four parameters, a, b, c, d

\[
\begin{align*}
F(x) &= 0 & \text{if } x \leq a \\
F(x) &= \frac{x - a}{b - a} & \text{if } a < x \leq b \\
F(x) &= 1 & \text{if } b \leq x \leq c \\
F(x) &= \frac{d - x}{d - c} & \text{if } c < x \leq d \\
F(x) &= 0 & \text{if } x > d
\end{align*}
\]

**Piecewise Linear Membership Function**
Level Sets

- $\alpha$-level set of $F$ is a crisp set $F_{\alpha} = \{ x / F(x) \geq \alpha \}$

- **Representation Theorem**: Any fuzzy subset can be expressed in terms of its level sets
  \[
  F = \bigcup_{\alpha \in [0, 1]} \alpha \otimes F_{\alpha}
  \]

- **Special Level Sets**
  - **Core**: $F_1 = \{ x / F(x) = 1 \}$
  - **Support**: $F_0^+ = \{ x / F(x) > 0 \}$
  - **Mid**: $F_m = F_{\alpha} = \{ x / F(x) \geq 0.5 \}$
Level Sets and Trapezoidal Fuzzy Subsets

- Level sets are interval subsets \( F_\alpha = [L_\alpha, U_\alpha] \)

- \( \alpha_1 > \alpha_2 \) then \( F_{\alpha_1} \subseteq F_{\alpha_2} \)
  \[ L_{\alpha_1} \geq L_{\alpha_2} \text{ and } U_{\alpha_1} \leq U_{\alpha_2} \]

- Linearity of trapezoid useful
  
  \textit{Can get every level set from any two level sets}
Assume $F_{\alpha_1} = [L_{\alpha_1}, U_{\alpha_1}]$ and $F_{\alpha_2} = [L_{\alpha_2}, U_{\alpha_2}]$

For any $0 \leq \alpha \leq 1$

$$L_{\alpha} = \frac{1}{(\alpha_1 - \alpha_2)} \left[ L_{\alpha_1} (\alpha - \alpha_2) + L_{\alpha_2} (\alpha_1 - \alpha) \right]$$

$$U_{\alpha} = \frac{1}{(\alpha_1 - \alpha_2)} \left[ U_{\alpha_1} (\alpha - \alpha_2) + U_{\alpha_2} (\alpha_1 - \alpha) \right]$$

*Trapezoid completely specified by two level sets*
Trapezoidal Preserving Operations

• If A and B are trapezoids and G is an arithmetic operation via extension principle F = G(A, B)

• G is a trapezoidal preserving operation if F is also a trapezoid

• F = w_1 A + w_2 B is trapezoidal preserving
Level Sets and Weighted Sums

• If $F = w_1 A + w_2 B$ and $w_1$ and $w_2$ non-negative, the operation simply performed on level sets

• $A_\alpha = [L_A, U_A]$ and $B_\alpha = [L_B, U_B]$

• $F_\alpha = [L_F, U_F]$

• $F_\alpha = [w_1 L_A + w_2 L_B, w_1 U_A + w_2 U_B]$

• To get $F$ all we need is two $F_\alpha$

Trapezoidal representation very efficient for processes involving these types of operation
Learning from Observations

- \( V \) is variable whose domain is \([a, b]\)

- Let \( E \) be the current estimate of the value of \( V \)

- Let \( D \) be new observation of the value of \( V \)

- New estimate \( F \) of the value of \( V \)

\[
F = E + \lambda (D - E) = \lambda D + \bar{\lambda} E
\]

where \( \lambda \in [0, 1] \) is learning rate
Learning from Granular Observations

• Represent observations and estimates using trapezoidal representation

• Take advantage of efficiency of trapezoids in this linear environment
Granular Calculation

• \( F = \lambda D + \bar{\lambda} E \)

• \( E_\alpha = [L_{E_\alpha}, U_{E_\alpha}] \)
  \( D_\alpha = [L_{D_\alpha}, U_{D_\alpha}] \)
  \( F_\alpha = [L_{F_\alpha}, U_{F_\alpha}] \)

• \( F_\alpha = [\lambda L_{D_\alpha} + \bar{\lambda} L_{E_\alpha}, \lambda U_{D_\alpha} + \bar{\lambda} U_{E_\alpha}] \)

• \( F \) obtained from any two level sets
  \( F_1 = [\lambda L_{D_1} + \bar{\lambda} L_{E_1}, \lambda U_{D_1} + \bar{\lambda} U_{E_1}] \)
  \( F_{0.5} = [\lambda L_{D_{0.5}} + \bar{\lambda} L_{E_{0.5}}, \lambda U_{D_{0.5}} + \bar{\lambda} U_{E_{0.5}}] \)
Uncertainty in Granular Estimates

Three examples of estimates of $V$:

A: $A_1 = [5, 5]$ and $A_{0.5} = [5, 5]$

B: $B_1 = [4, 8]$ and $B_{0.5} = [3, 10]$

C: $C_1 = [0, 10]$ and $C_{0.5} = [0, 10]$

A provides more information about $V$, it says it is precisely 5.

B provides less information than A but it is better then that provided by C
Measuring Information in Estimates

- Specificity measures information in fuzzy subset

- Assume V is variable with domain \([a, b]\)

- Let \(A = [c, d]\) be an interval subset
  \[
  SP(A) = 1 - \frac{\text{Length}(A)}{b - a} = 1 - \frac{d - c}{b - a}
  \]
  Inversely related to the size of the interval.

- Let \(F\) be a trapezoidal fuzzy subset on \([a, b]\)
  \[
  SP(F) = 1 - \frac{\text{Length}(F_{0.5})}{b - a} = SP(F_{0.5})
  \]
  Bigger \(F_{0.5}\) the less information
Effect of Learning on Specificity

• New Estimate
  \[ F_\alpha = [\lambda L_D_\alpha + \bar{\lambda} L_E_\alpha, \lambda U_D_\alpha + \bar{\lambda} U_E_\alpha] \]

• New Estimate Level Set Length
  \[ \text{Length}(F_\alpha) = \lambda \text{Length}(D_\alpha) + \bar{\lambda} \text{Length}(E_\alpha) \]

• Specificity of New Estimate
  \[ \text{Sp}(F) = \lambda \text{Sp}(D) + \bar{\lambda} \text{Sp}(E) \]

Specificity of new estimate is weighted average of specificity of the observation and the specificity of the current estimate.
Very uncertain/imprecise observations, those with small specificity, will tend to decrease the specificity of our estimate.

WE MUST FIX THIS!
Modified Learning Rule

- \( F = E + \lambda \sigma (D - E) \)

- \( \sigma \in [0, 1] \) term relating the specificity of observation \( Sp(D) \) with specificity of current estimate \( Sp(E) \)

- Smaller \( Sp(D) \) relative to \( Sp(E) \) smaller \( \sigma \)

- Smaller \( \sigma \) less the effect of observation
  \[
  F = \lambda \sigma D + (1 - \lambda \sigma) E
  \]
  \[
  Sp(F) = \lambda \sigma Sp(D) + (1 - \lambda \sigma) Sp(E)
  \]
Possible Forms for $\sigma$

- $\sigma = 1$ if $\text{SP}(D) \geq \text{SP}(E)$

$$\sigma = \frac{\text{Sp}(D)}{\text{Sp}(E)}$$ if $\text{SP}(D) < \text{SP}(E)$

- $\sigma = 1$ if $\text{SP}(D) \geq \text{SP}(E)$

$$\sigma = \left(\frac{\text{Sp}(D)}{\text{Sp}(E)}\right)^r$$ if $\text{SP}(D) < \text{SP}(E)$
• Obtain $\sigma$ using a fuzzy systems model

If $SP(D)$ is $A_1$ and $SP(E)$ is $B_1$ then $\sigma = g_1$

If $SP(D)$ is $A_2$ and $SP(E)$ is $B_2$ then $\sigma = g_2$

...  

If $SP(D)$ is $A_q$ and $SP(E)$ is $B_q$ then $\sigma = g_q$