

Fuzzy Sets and Fuzzy Techniques

Lecture 4 – Membership functions and uncertainty

Joakim Lindblad
joakim@cb.uu.se

Centre for Image Analysis
Uppsala University

2010-02-20

Joakim Lindblad, 2010-02-20 (1/28)

Topics of today

- Constructing fuzzy sets I. Ch. 10
- Characterizing fuzzy sets I. Ch. 9.1–9.4
 - Nonspecificity
 - Fuzziness

Joakim Lindblad, 2010-02-20 (2/28)

Constructing fuzzy sets I

Methods of construction

- Direct methods and indirect methods
- One expert and multiple experts

Joakim Lindblad, 2010-02-20 (3/28)

Direct methods with one expert

- Define the complete membership function based on a justifiable mathematical formula
 - Often based on mapping of directly measurable features of the elements of X
- Exemplifying it for some selected elements of X and interpolate (/extrapolate) MF in some way.
 - Expert of some kind

Joakim Lindblad, 2010-02-20 (4/28)

Direct methods with one expert

Directly defining the membership function

Examples:

- Fuzzy set of straight lines, S
 - Least squares straight-line fitting, with residual error $\epsilon(x)$ for given line x , where ϵ_t is the largest acceptable error.

$$S(x) = \begin{cases} 1 - \epsilon(x)/\epsilon_t & \text{when } \epsilon(x) < \epsilon_t \\ 0 & \text{otherwise} \end{cases}$$

- Direct mapping of grey values of an image (fuzzy thresholding)

Joakim Lindblad, 2010-02-20 (5/28)

Direct methods with one expert

Defining (learning) the membership function through examples

Example:

Expert judging, often with the aid of **ideal prototypes**.

- "What is the degree of membership of x in A ?"
- "What is the degree of compatibility of x with L_A ?"
where L_A is a linguistic term we wish to represent.

Reverse formulation:

- "Which elements x have the degree $A(x)$ of membership in A ?"
- "Which elements x are compatible with L_A to degree $A(x)$?"

Joakim Lindblad, 2010-02-20 (6/28)

Direct methods with one expert

Interpolating the membership function

Interpolation of the MF requires, in general, some distance measure in X , e.g. if $X \subseteq \mathbb{R}^n$, or if we have some mapping $\Phi : X \rightarrow \mathbb{R}^n$ (numerical features of X) which induces a distance in X .

Section 10.7: Curve fitting techniques.

- Lagrange interpolation
- Least squares curve fitting
- Neural networks
- Spline fitting
- ...

Joakim Lindblad, 2010-02-20 (7/28)

Direct methods with multiple experts

The opinions of several experts need to be aggregated.

Example: Average (Probabilistic interpretation)

$$A(x) = \frac{1}{n} \sum_{i=1}^n a_i(x)$$

Joakim Lindblad, 2010-02-20 (8/28)

Direct methods with multiple experts

The opinions of several experts need to be aggregated.

Example: Weighted average,
 c_i being the degree of competence of expert i

$$A(x) = \sum_{i=1}^n c_i a_i(x),$$

where

$$\sum_{i=1}^n c_i = 1$$

Direct methods with multiple experts

Instead of aggregating the multiple suggested membership values to one, we may choose to use an interval-valued fuzzy set, or a fuzzy set of type 2.

Indirect methods with one expert

It may be easier/more objective to ask simpler questions to the experts, than the membership directly.

Example: Pairwise comparisons

- Problem: Determine membership $a_i = A(x_i)$
- Extracted information: Pairwise relative belongingness, matrix $\mathbf{P}$ with $p_{ij} \approx \frac{a_i}{a_j}$

Indirect methods with one expert

Example: Pairwise comparisons

If $\mathbf{P}$ is perfect, i.e. $p_{ij} = \frac{a_i}{a_j}$, then

$$\sum_{j=1}^n p_{ij} a_j = \sum_{j=1}^n a_i = n a_i$$

or in matrix form,

$$\mathbf{P}\mathbf{a} = n\mathbf{a}$$

n is an **eigenvalue** of $\mathbf{P}$ and $\mathbf{a}$ is the corresponding **eigenvector**.

Indirect methods with one expert

Example: Pairwise comparisons

In reality, $\mathbf{P}$ is not fully consistent, but is instead slightly perturbed.

When the values p_{ij} changes slightly, the eigenvalues of $\mathbf{P}$ changes in a similar way.

Finding the largest eigenvalue and its associated eigenvector provides us with a good solution. The closer $\lambda_{\max}$ is to n , the more accurate is the estimate of $\mathbf{a}$.

Note: This does not give us absolute memberships, the solution is often normalized as a post-processing step.

Indirect methods with multiple experts

Section 10.6. Iterated box classification scheme.

Characterizing fuzzy sets I

Uncertainty measures

Uncertainty measures

- Nonspecificity of crisp sets
- Nonspecificity of fuzzy sets
- Fuzziness of fuzzy sets

Information and uncertainty

Uncertainty is connected with some information deficiency.

Information may be

- incomplete
- imprecise
- fragmentary
- unreliable
- vague
- contradictory
- ...

Information and uncertainty

We are concerned with information in terms of reduction of uncertainty – **uncertainty based information**.

Alternative: Descriptive information or algorithmic information

The length of the shortest possible program by which the object is described/can be computed.

Information and uncertainty

Information theory
(Shannon)

- Probabilistic uncertainty

Generalized information theory

- Nonspecificity (or imprecision)
related to size (cardinality) of relevant sets
- Fuzziness (or vagueness)
imprecise boundaries of relevant sets
- Strife (or discord)
conflicts among various sets or alternatives

Nonspecificity of crisp sets

Hartley function

Hartley [1928] showed that a function

$$U(A) = c \log_b |A| ,$$

where $|A|$ is the cardinality of A , and $b > 1$ and $c > 0$ are constants, is the only sensible way to measure the amount of uncertainty associated with a finite set of possible alternatives.

$b = 2$ and $c = 1 \rightarrow$ uncertainty measure in **bits**

$$U(A) = \log_2 |A|$$

Relates to the nonspecificity inherent in each set.
Larger sets correspond to less specific predictions.

Nonspecificity of crisp sets

Assume that a set A is reduced to a subset B by some action. The amount of (uncertainty based) information produced by the action $I(A, B)$ is equal to the reduction in uncertainty

$$I(A, B) = U(A) - U(B) = \log_2 \frac{|A|}{|B|}$$

When the action eliminates all but one alternative, $|B| = 1$, then $I(A, B) = U(A) = \log_2 |A|$. That is, $U(A)$ is the amount of information needed to characterize one element of the set A . (Number of bits required to indicate one element.)

Nonspecificity of crisp sets

A meaningful counterpart to the Hartley function for infinite sets is the following

$$U(A) = \log[1 + \mu(A)],$$

where $\mu(A)$ is the *measure* of A defined by the Lebesgue integral of the characteristic function of A .

This uncertainty is not directly comparable with values obtained for finite sets. The choice of logarithm for infinite sets is often taken to be the natural logarithm.

Nonspecificity of fuzzy sets

 U -uncertainty

Generalized Hartley function

$$U(A) = \frac{1}{h(A)} \int_0^{h(A)} \log_2 |\alpha A| d\alpha$$

Weighted average of the Hartley function for all distinct α -cuts of the normalized counterpart of A .

Fuzzy sets that are equal when normalized have the same nonspecificity.

Nonspecificity of fuzzy sets

 U -uncertainty

Generalized Hartley function for infinite sets

$$U(A) = \frac{1}{h(A)} \int_0^{h(A)} \log[1 + \mu(\alpha A)] d\alpha,$$

where $\mu(A)$ is the measure of A .

Fuzziness of fuzzy sets

A measure of fuzziness is a function

$$f : \mathcal{F}(X) \rightarrow \mathbb{R}^+$$

For each fuzzy set A , $f(A)$ expresses the degree to which the boundary of A is not sharp.

The following three requirements are essential

- 1 $f(A) = 0$ iff A is a crisp set
- 2 $f(A)$ attains its maximum iff $A(x) = 0.5$ for all $x \in X$
- 3 $f(A) \leq f(B)$ when set A is “undoubtedly” sharper than set B
 - a) $A(x) \leq B(x)$ when $B(x) \leq 0.5$
 - b) $A(x) \geq B(x)$ when $B(x) \geq 0.5$

Fuzziness of fuzzy sets

One way to measure fuzziness of a set A is to measure the **distance** between A and the nearest crisp set. Remaining is to choose the distance measure.

Another way is to view the fuzziness of a set as the **lack of distinction between the set and its complement**. The less a set differs from its complement, the fuzzier it is. Also this path (which is the one we will take) requires a distance measure.

This last view, combined with a non-standard complement, requires the exchange of 0.5 in points 2 and 3 on the previous slide with the equilibrium of the fuzzy complement.

Fuzziness of fuzzy sets

A simple and intuitive distance measure is the Hamming distance.

$$d(A, B) = \sum |A(x) - B(x)|$$

The measure of fuzziness then becomes

$$\begin{aligned} f(A) &= d(X, \bar{X}) - d(A, \bar{A}) \\ &= \sum (1 - |A(x) - (1 - A(x))|) \\ &= \sum (1 - |2A(x) - 1|) \end{aligned}$$

Generalizes easily to infinite bounded subsets by replacing the sum with an integral.

Information gain?

Fuzziness and nonspecificity are distinct types of uncertainty and totally independent of each other.

They are also totally different in their connections to information. When nonspecificity is reduced, we view this as a gain in information, regardless of any associated change in fuzziness. The opposite, however, is not true.

A reduction of fuzziness is reasonable to consider as a gain of information only if the nonspecificity also decreases or remains the same.