Artificial Neural Networks
Lecture Notes

Stephen Lucci, PhD

Part 9

About this file:

- This is the printer-friendly version of the file "*lecture09.htm*". In case the page is not properly displayed, use IE 5 or higher.
- Since this is an offline page, each file "lecture*nn*.htm" (for instance "lecture09.htm") is accompanied with a "img*nn*" folder (for instance "img09") containing the images which make part of the notes. So, to see the images, Each html file must be kept in the same directory (folder) as its corresponding "img*nn*" folder.
- If you have trouble reading the contents of this file, or in case of transcription errors, email gi0062@bcmail.brooklyn.cuny.edu
- Acknowledgments:
Background image is from http://www.anatomy.usyd.edu.au/online/neuroanatomy/tutorial1/tutorial1.html (edited) at the University of Sydney Neuroanatomy web page. Mathematics symbols images are from metamath.org's GIF images for Math Symbols web page. Other image credits are given where noted, the remainder are native to this file.

Contents

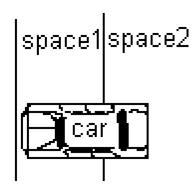
Fuzzy Logic

- Neural networks can learn. However, they are rather opaque.
- "The network acts like a black box by computing a statistically sound approximation to a function known only from a training set.", Rojas.

- Fuzzy logic has an explanatory capability. However, it is unable to learn.

Fuzzy Sets Vs. Crisp Sets

- "Raise your hand if you're male" ... hands down. Now,
- "Raise your hand if you're female" ... hands down.
- These are *crisp sets*.
- "Raise your hand if you're satisfied with your jobs" ... hands down. Now,
- Some hands probably went up both times, and
- Hands may have been only somewhat raised in each case.
- We would say that job satisfaction is a *fuzzy concept*, in that most people are not entirely satisfied or dissatisfied with their jobs.
- Parking spaces in a lot would be another example: Most of the time one has a situation like the one shown on the right.



- *Fuzzy logic* was developed by Lotfi Zadeh during the mid 1960's.
- Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite set
- The subset A of X consisting of x_1 alone, can be described by the n-dimensional membership vector

  $$Z(A) = (1, 0, 0, \ldots 0).$$

- The subset B of X with elements x_1 and x_n is described by the vector

  $$Z(A) = (1, 0, 0, \ldots, 1).$$

- Any other crisp subset of X can be represented in the same way by an n-dimensional binary vector.
- Consider the fuzzy set C

  $$Z(C) = (0.5, 0, 0, \ldots, 0).$$

In classical set theory, this would be impossible. For classically, either $x_1 \in C$ or it doesn't. The world is black and white in classical set theory.

- With *Fuzzy* set theory, this type of description is permitted.
- The element x_1 belongs to the set C only to some extent.
- The degree of membership is expressed by a real number in the interval [0,1].
- An example:
  "x_1 is a tall person"
  Is a 6-foot individual tall? ... yes.
- Other diffuse statements:
  x is old
  y is young
  z is mature

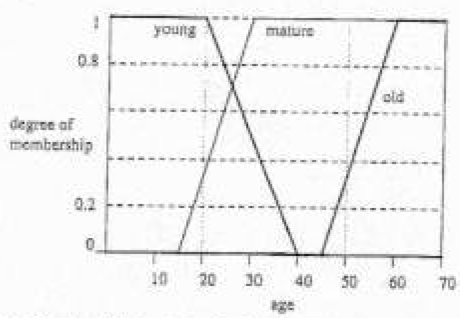


Fig. 11.1. Membership functions for the concepts young, mature and old

Three membership functions in the interval 0:70 years.

- The three functions define the degree of membership of any given age in the sets of young, mature and old ages.
- A 20-year old - degree of membership in the set
  of young people is 1.0
  of mature adults is 0.35 , and
  of old persons is 0.0.
- If someone is 50 years old, the degrees of membership are 0.0, 1.0, 0.3 in the respective sets.
- <u>Definition 11.1</u>:
  Let X be a classical universal set. A real function $\mu_A : X \longrightarrow [0,1]$ is called the membership function of A, and defines the fuzzy set A of X.
  This is the set of all pairs $(x, \mu_A(x))$ with $x \in X$.
- A fuzzy set is completely determined by its membership function.

- The *set of support* of a fuzzy set A, is the set of all elements x of X for which $(x, \mu_A(x)) \in A$, and $\mu_A(x) > 0$ holds.

- A fuzzy set A with finite set of support $\{a_1, a_2, \ldots, a_m\}$ can be described as:

$$A = \mu_1 / a_1 + \mu_2 / a_2 + \ldots + \mu_m / a_m$$

  where $\mu_i = \mu_A(a_i)$ for $i = 1, \ldots, m$.

- Let $X = \{x_1, x_2, x_3\}$.

  The classical (crisp) subsets $A = \{x_1, x_2\}$ and $B = x_2, x_3\}$ can be represented as

$$A = 1/x_1 + 1/x_2 + 0/x_3$$

$$B = 0/x_1 + 1/x_2 + 1/x_3$$

- The union of A and B is computed by taking for each element x_i the maximum of its membership in both sets. I.e.,

$$A \cup B = 1/x_1 + 1/x_2 + 1/x_3 .$$

- The fuzzy union of two fuzzy sets can be computed in the same way.
  Let

$$C = 0.5/x_1 + 0.6/x_2 + 0.3/x_3$$

$$D = 0.7/x_1 + 0.2/x_2 + 0.8/x_3$$

  Then

$$C \cup D = 0.7/x_1 + 0.6/x_2 + 0.8/x_3 .$$

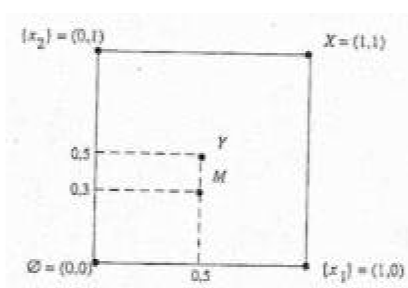
- The fuzzy intersection of two sets can be defined in a similar manner.

Geometric Interpretation of Fuzzy Sets

- Let $X = \{x_1, x_2\}$ be a universal set.

- Each point in the interior represents a subset of X

- The crisp subsets of X are located at the vertices of the unit square.



- The point (1,0) represents the set x_1.
  The point (0,1) represents the set x_2
  etc. ...

- The crisp subsets of X are located at the vertices of the unit square.

- The inner region of a unit hypercube in an n-dimensional space - the fuzzy region.

- The point M (in the figure above) corresponds to the fuzzy set

  $$M = 0.5/x_1 + 0.3/x_2.$$

- Center of square - most diffuse of all fuzzy sets.

- The degree of fuzziness of a fuzzy set can be measured by its entropy (here different from the entropy used in information theory or physics.)

- The **entropy** (index of fuzziness) corresponds inversely to the distance between the representation of the set and the center of the unit square.

- Entropy here is the index of fuzziness, crispness, certitude or ambiguity.

- The set Y in the previous figure has the maximum entropy.

- The vertices represent the crisp sets and have the lowest entropy, i.e. 0.

Entropy of a Fuzzy Set M

- Quotient of the distance of the nearest corner from M to the distance d_2 from the corner which is farthest away.

  $$E(M) = d_1/d_2.$$

  $0 \leqslant$ entropy of a set $\leqslant 1$.

- Maximum entropy - at the center of square.

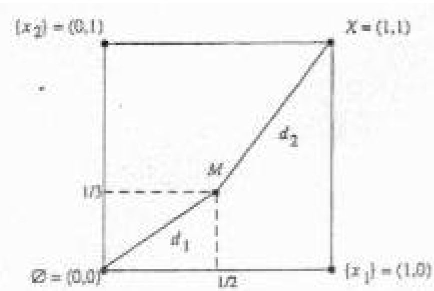


Fig. 11.3. Distance of the set M to the universal and to the void set

Visualization for Union and Intersection of Fuzzy Sets

- The membership function for the *union* of two sets A and B is

$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)), \forall x \in X.$

- This corresponds to the maximum of the corresponding coordinates in the geometric visualization.

- The membership function for the *intersection* of two sets A and B is given by

$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)), \forall x \in X.$

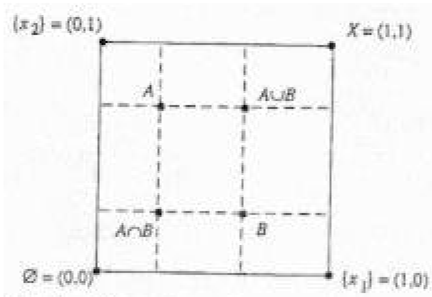


Fig. 11.4. Intersection and union of two fuzzy sets

- The union or intersection of two fuzzy sets is in general a fuzzy, not a crisp set.

- The *complement* A^C of a fuzzy set is given by

$\mu_{A^c}(x) = 1 - \mu_A(x), \forall x \in X.$

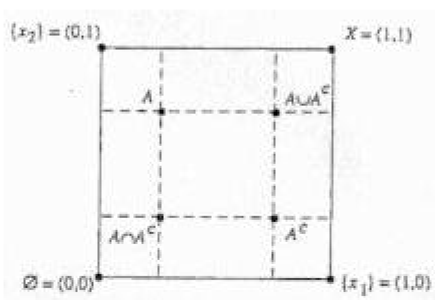


Fig. 11.5. Complement A^c of a fuzzy set

The line joining the representation of A and A^C goes through the center of the square.

- In general, for fuzzy sets we have

$A \cup A^C \neq X$, and

$A \cap A^C \neq X$, and

Fuzzy Sets and Logic:

On the Relationship between Set Theory and the Propositional Logic

- Let A, B be two crisp sets. We have,

  $\mu_A , \mu_B : X \longrightarrow \{0,1\}$.

- The membership function $\mu_{A \cup B}$ for $A \cup B$ is

  $\mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) , \quad \forall x \in X.$

  where $0 \equiv$ False , $1 \equiv$ True.

- Similarly, for the membership function for $A \cap B$, we have

  $\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) , \quad \forall x \in X.$

- For the complement A^c,

  $\mu_{A^c}(x) = \neg \mu_A(x)$.

- *De Morgan's Laws*

  $(A \cup B)^c \equiv A^c \cap B^c$ corresponds to $\neg(A \vee B) \equiv \neg A \wedge \neg B$

  $(A \cap B)^c \equiv A^c \cup B^c$ corresponds to $\neg(A \wedge B) \equiv \neg A \vee \neg B$

- *Fuzzy Counterparts to Logic Operators*
  We may identify:

  the OR operation ($\bar{\vee}$) with *maximum*,
  the AND ($\bar{\wedge}$) with *minimum*, and
  Complementation ($\bar{\neg}$) with $x \mapsto 1 - x$.

- Then we may write:

  $\mu_{A \cup B}(x) = \mu_A(x) \,\bar{\vee}\, \mu_B(x) , \forall x \in X,$
  $\mu_{A \cap B}(x) = \mu_A(x) \,\bar{\wedge}\, \mu_B(x) , \forall x \in X,$
  $\mu_{A^c}(x) = \bar{\neg} \mu_A(x) , \forall x \in X.$

- The properties of the isomorphism present in the classical theories are preserved.

- Many classical logic rules are valid for fuzzy operators.
  E.g., *min* and *max* are commutative and associative.

- But note: the principle of no contradiction does *not* hold.
  E.g., proposition A with truth value 0.4:

  $A \bar{\wedge} \neg A = \min(0.4, (1 - 0.4)) \neq 0.$

- Similarly, the *Law of Excluded Middle* is not valid either.

Axiomatic Definitions of Fuzzy Operators

Fuzzy OR Operator

- Axiom $\mathcal{U}_1$: *Boundary Conditions*
  $0 \bar{\vee} 0 = 0$
  $1 \bar{\vee} 0 = 1$
  $0 \bar{\vee} 1 = 1$
  $1 \bar{\vee} 1 = 1.$
- Axiom $\mathcal{U}_2$: *Commutativity*
  $a \bar{\vee} b = b \bar{\vee} a.$
- Axiom $\mathcal{U}_3$: *Monotonicity*
  If $a \leqslant a'$ and $b \leqslant b'$,
  then $a \bar{\vee} b \leqslant a' \bar{\vee} b'$.
- Axiom $\mathcal{U}_4$: *Associativity*
  $a \bar{\vee} (b \bar{\vee} c) = (a \bar{\vee} b) \bar{\vee} c.$
- Axiom $\mathcal{U}_5$: *Idempotence*
  $a \bar{\vee} a = a.$

  - The *maximum* function satisfies $\mathcal{U}_1 \longrightarrow \mathcal{U}_4$

Fuzzy AND Operator

- The fuzzy AND is monotonic, commutative and associative.

  Boundary Conditions
  $0 \bar{\wedge} 0 = 0$
  $1 \bar{\wedge} 0 = 0$
  $0 \bar{\wedge} 1 = 0$
  $1 \bar{\wedge} 1 = 1.$

Fuzzy Negation

- Axiom N_1: *Boundary Conditions*
  $\bar{\neg} 0 = 1$
  $\bar{\neg} 1 = 0$
- Axiom N_2: *Monotonicity*

If $a \leqslant b$, then $\neg b \leqslant \neg a$.

- Axiom N_3: *Involution*
  $a = \neg\neg a$.

- Graphs of the functions max and min (i.e., a possible fuzzy AND and fuzzy OR combination,)

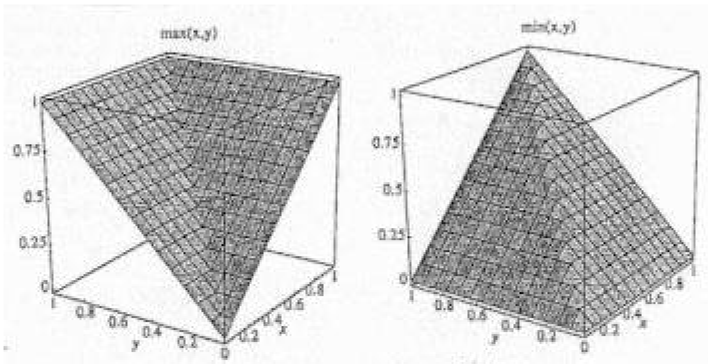


Fig. 11.7. The max and min functions

- They could be used as activation functions in neural networks.

- Using output functions derived from fuzzy logic can have the added benefit of providing a logical interpretation of the neural output.

Fuzzy Inferences

- Fuzzy logic operators can be used as the basis for inference systems.

- Fuzzy inference rules have the same structure as classical ones.
  E.g., rule R1: if $(A \bar{\wedge} B)$ Then C
  $(A \bar{\vee} B)$ Then D.

- Note $\bar{\wedge}$ ~ min and $\bar{\vee}$ ~ max, though other choices are possible.

- Let the truth values of A and B be 0.4 and 0.7. Then

$$A \bar{\wedge} B = \min(0.4, 0.7) = 0.4$$
$$A \bar{\vee} B = \max(0.4, 0.7) = 0.7$$

- Fuzzy inference mechanism - Rules R1 and R2 can only be partially applied (i.e., rule R1 is applied 40% and rule R2 70%.)

- The result of the inference is a combination of the propositions C and D.