



# Summer School on Capability and Multidimensional Poverty

28 August-9 September, 2008

New Delhi, India

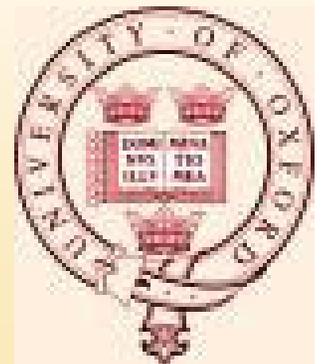


## HDCP-IRC

The Human Development, Capability  
and International Research Centre  
Istituto Universitario di Studi Superiori  
[www.iusspavia.it](http://www.iusspavia.it)

## OPHI

Oxford Poverty & Human  
Development Initiative  
University of Oxford  
[www.ophi.org.uk](http://www.ophi.org.uk)



# Fuzzy sets theory

Enrica Chiappero- Martinetti  
University of Pavia and HDC-IRC, IUSS

So far as laws of mathematics refer to reality, they are not certain.  
And so far as they are certain, they do not refer to reality (Albert Einstein 1922)

All traditional logic habitually assumes that precise symbols are being employed.  
It is therefore not applicable to this terrestrial life  
but only to an imagined celestial existence [...]  
logic takes us nearer to heaven than most other studies” (Russell 1923, p. 88-9).

# Some conceptual issues I

- Classical logic and the principle of bivalence do not accept vague predicates. Yet in all fields of science, as well as in ordinary language, many concepts and predicates escape the “law of the excluded middle” (either A or not A). This happens when:
  - a clear cut-off point between a given concept and its opposite does not exist
  - meanings can change according to the context or situation in which they are applied
  - Objects possess property (belong to a given set) to varying degree

# Some conceptual issues II

- Vagueness and Greek paradoxes (“how many grains of sand can you remove from a sand pile before it isn’t a pile anymore?”)
  - intrinsic vagueness (i.e. difficulty in “drawing a line”): it refers to the nature of a given concept or phenomenon;
  - vagueness in measurement: it relates to the way in which vagueness can be accounted for (i.e classical/bivalent logic vs many-valued logic),

# Some conceptual issues III

“Poor” is a vague predicate because:

- It involves borderline cases (a person is not clearly *poor* and not clearly *not poor*)
- It lacks sharp boundaries (along a hypothetical scale of well-being, an exact point at which a poor person ceases to be poor does not really exist)
- It is at risk to the Sorites paradox (a person with an income of \$1 billion can be considered as a rich person. If we take away \$1, he can still be rich. If we carry on, dollar by dollar, along a continuum line of people ranked by amounts of income that are always \$1 less than the previous one, we come to the conclusion that a person with no income is a rich person: but this is clearly false.)

# Fuzzy Sets Theory

- Generalization of crisp or classical set theory; applications range from consumer products (cameras, washing machines) to industrial process control, medical instrumentation, decision-making support systems, portfolio selection, etc.
- FST, Fuzzy logic, Fuzzy inference systems
- Economics: finance, oligopoly theory; fuzzy inequality and poverty measures (Basu 1987, Ok 1995, Shorrocks and Subramanian, 1994, Chakravarty 2006), CA literature

# FST II

- Some reasons for adding fuzzy methods to the social science toolbox
  - able to hand vagueness systematically
  - able to analyze multivariate relationship generalizing set-theoretic operations
  - even categorical concepts often turn out to be a matter of degree
  - combine set-wise thinking and continuous variables in a rigorous fashion

# FST III

Three possible level of analysis

1. Description of phenomena in their innate complexity and gradualness (pov & wb as a continuum between extreme deprivation and full achievement of functionings → membership functions
2. Relations across variables and domains → aggregation operators
3. Application of fuzzy logic rules (if....then) and fuzzy inference systems to infer a conclusion starting from given premises

# Membership function I

- first substitutes the characteristic function of a crisp set (that assigns a value of either 1 or 0 to each element in the universal set) with a generalized characteristic function (called membership function) which varies between 0 and 1. Larger values denote higher degrees of membership.
- if  $X$  denotes a universal set of objects, then the membership function  $\mu_A$  has the form

$$\mu_A(x) : X \rightarrow [0, 1]$$

$A$  is the fuzzy set;  $[0,1]$  is the interval of real numbers from 0 to 1:

$\mu_A(x) = 0$  if the element  $x \in X$  does not belong to  $A$

$\mu_A(x) = 1$  if  $x$  completely belongs to  $A$  and

$0 < \mu_A(x) < 1$  if  $x$  partially belongs to  $A$ .

## Membership functions II

Four steps:

- To specify the domain  $X$  (what is the universe of elements under consideration)
- to define an appropriate arrangement of modalities (or values) on the basis of the different degrees of hardship/well-being ;
- to identify the two extreme conditions such that  $\mu_A(x) = 1$  (full membership) and  $\mu_A(x) = 0$  (non-membership) ; (nb  $\mu_A(x) = 0.5$  is called neutral point)
- to specify the membership functions for all the other intermediate positions.

(measurement property)

## An example of gradually achieved functionings

---

“Being educated”	“Being nourished”
None (illiterate)	None (totally insufficient for survival; starvation)
Very low (ability to write and read only, but no formal education)	Very low (sporadic access to food and very serious under-nourishment)
Low (attendance at the lowest level of formal education, but with discontinuity and/or dropout)	Low (malnourishment; inadequate diet)
Sufficient (lower formal level of education achieved)	Sufficient (minimum calorie intake almost achieved, but diet not fully balanced)
Quite good (attendance in secondary school, but with discontinuity and/or dropout)	Quite good (minimum calorie intake achieved, but diet not fully balanced)
Good (secondary school achieved)	Good (minimum calorie intake achieved, and diet quite balanced)
Very Good (access to/achievement of the highest level of secondary school education)	Very good (nutritional level and balanced diet achieved)

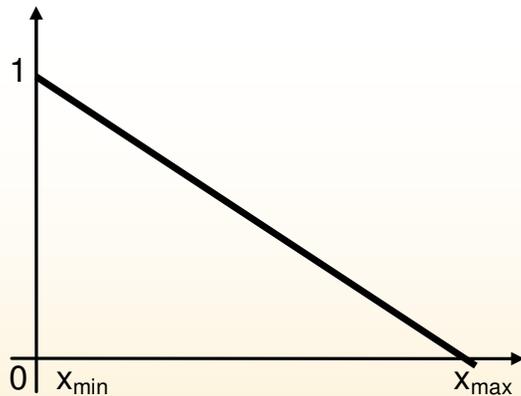
---

# Membership functions III

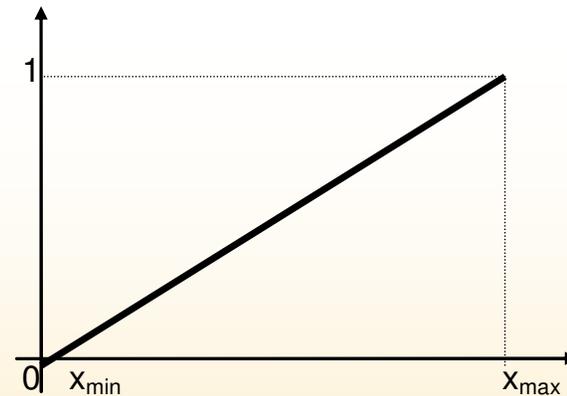
How the membership function  $\mu$  can be defined:

a) arbitrarily chosen by the investigator: some examples

a.1) Linear function



$$\mu(x) = \frac{x_{\max} - x}{x_{\max} - x_{\min}}$$



$$\mu(x) = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

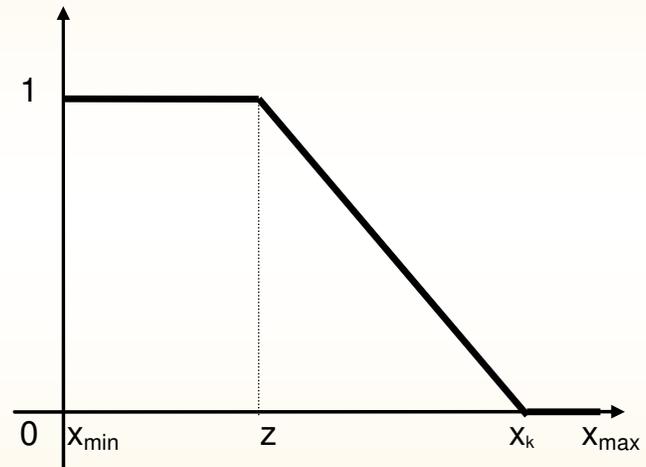
# Membership functions IV

## a.2) Trapezoidal function

$$\mu(x) = 1 \quad \text{if } x_{\min} \leq x \leq z$$

$$\mu(x) = 0 \quad \text{if } x_k \leq x \leq x_{\max}$$

$$\mu(x) = \frac{x_k - x}{x_k - z} \quad \text{if } z \leq x \leq x_k$$



# Membership functions V

## a.3) Sigmoid (or logistic) function

$$\mu(x) = 1/[1+e^{-a(x-b)}]$$

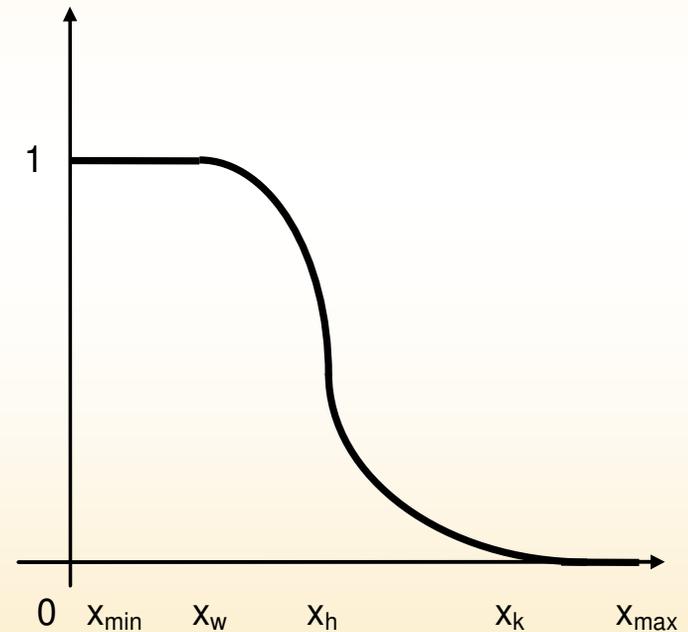
e.g.

$$x_k = \mu(x) = 0 \quad x_h = \mu(x) = 0.5 \quad x_w = \mu(x) = 1$$

$$\begin{aligned} \mu(x) &= 1 && \text{if } x_{\min} \leq x \leq x_w \\ \mu(x) &= 0 && \text{if } x_k \leq x \leq x_{\max} \end{aligned}$$

$$\mu(x) = 1 - \frac{1}{2} \left[ \frac{x_w - x}{x_w - x_h} \right]^2 \quad \text{if } x_w \leq x \leq x_h$$

$$\mu(x) = \frac{1}{2} \left[ \frac{x_k - x}{x_k - x_h} \right]^2 \quad \text{if } x_h \leq x \leq x_k$$



## Membership functions VI

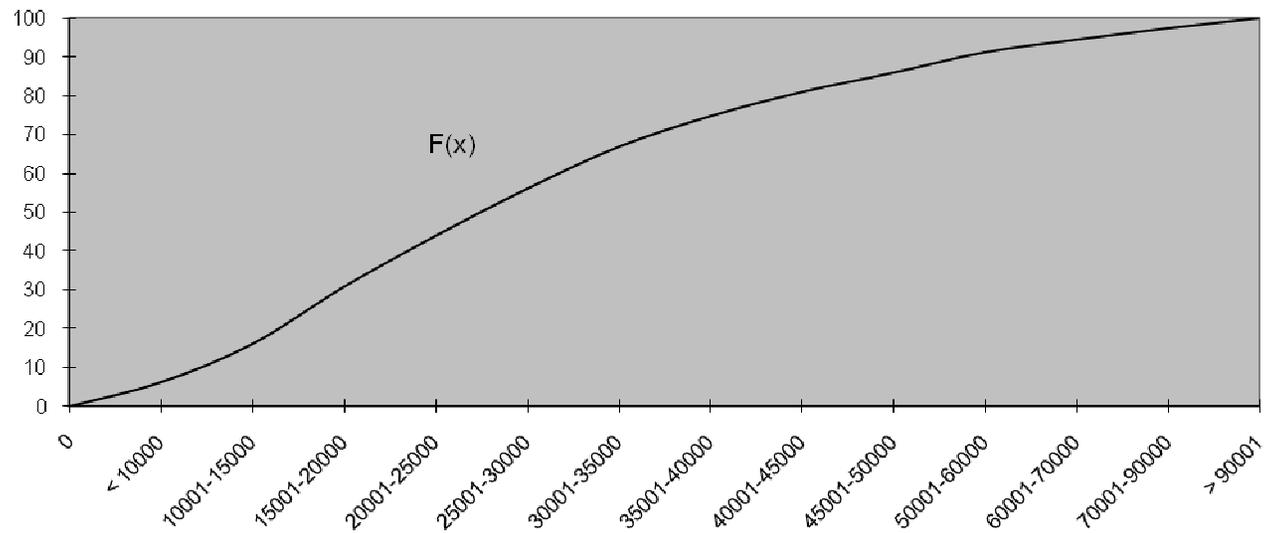
b) according to empirical evidence, interpolation of sample data (least-square method for fitting data and estimating the parameters): (e.g. Cheli and Lemmi 1995):

$$\mu(x^k) = \begin{cases} 0 & \text{if } k=1 \\ \mu(x_{k-1}) + \frac{F(x_k) - F(x_{k-1})}{1 - F(x_1)} & \text{if } k > 1 \end{cases}$$

where  $F(x)$  is the sampling distribution function of the variable  $x$  arranged in an increasing order according to  $k$

# Membership functions VII

Figure 5 - Equivalent income distribution function



# Membership functions VIII

Some possible interpretation of (approaches to) mf:

- Formalist interpretation: assigns mf in math terms by mapping the underlying var into the membership scale; var can come from different sources (subj perceptions of survey respondents, obj measured, external expert, indirect measurement models)
  - (nb: n.of plausible transformations is potentially limitless and sensitivity analysis is required )
- Probabilistic interpretation:  $\mu_A(x)$  is the (subjective or frequency-based) probability that x belongs to A
  - (nb. many fs theorist reject this interpretation

# Aggregation operators

## Aggregation operations:

- Across a set of elementary indicators for determining a given achieved functionings
- Across a set of functionings for determining an overall evaluation of well-being
- (nb we do not consider aggregation across individuals, see Chakravarty 2006)

FST includes and extends complement, union, intersection and inclusion operations of crisp set theory

# Aggregation operators II

## a) Standard fuzzy complement or negation

$$\sim A = 1 - \mu_A(x)$$

It contains all elements in  $X$  that are not in the set  $A$  and can be interpreted as the degree to which element  $x$  belongs to  $\sim A$  or, equivalently, does not belong to  $A$ . NB: negation determines the degree to which an element is complementary to the underlying fuzzy concept, but does not represent its opposite, as is the case in traditional logic.

## b) Inclusion

$A$  includes  $B$  ( $A \supset B$ ) if  $\mu_A(x) \geq \mu_B(x)$  for all  $x$

# Aggregation operators III

b.1) Standard fuzzy intersection (“and” operator)

$$\mu_{A \cap B} = \min [\mu_A, \mu_B]$$

Example:  $\mu_A=0.3$ ,  $\mu_B=0.7$   $\mu_{A \cap B}=0.3$

NB: implicitly rejects the hypothesis that a compensation or trade-off might be possible between A and B.

They can be usefully applied when there is a positive correlation between them, i.e. when the dimensions or attributes are complements (a higher level of education is associated with better achievement in terms of nutritional status).

# Aggregation operators IV

b.2) Standard fuzzy union (“or” operator)

$$\mu_{A \cup B} = \max[\mu_A, \mu_B]$$

Example:  $\mu_A=0.3$ ,  $\mu_B=0.7$   $\mu_{A \cup B} = 0.7$

NB: “or” operator correspond to a full compensation of lower degrees of membership by the maximum degree of membership (the higher membership score in the educational or nutritional functioning space would be a sufficient condition for achieving well-being).

# Aggregation operators $\vee$

Standard union and standard intersection:

- are simple to determine and intuitive;
- generalize the classical set theory when the range of values is restricted to 0 and 1
- can be adequate to deal with two dimensions for which a clear correlation can be identified.

In more complex, n-dimensional spaces, a potential limit of the standard operators (in particular, the standard intersection) is that extreme values can control the overall assessment.

# Aggregation operators VI

## Other common fuzzy sets operators

c.1) weak intersection (or algebraic product)

$$\mu_{A \cdot B} = [\mu_A \cdot \mu_B]$$

c.2) weak union (or algebraic sum)

$$\mu_{A+B} = [\mu_A + \mu_B - \mu_A \cdot \mu_B]$$

NB: they admit compensation between A and B and can be adequate when independent well-being dimensions are considered.

# Aggregation operators VII

d.1) bounded difference

$$\mu_{\underline{A \cap B}} = \max [0, \mu_A + \mu_B - 1]$$

d.2) bounded sum

$$\mu_{\underline{A \cup B}} = \min [1, \mu_A + \mu_B]$$

Appropriate in case of a negative correlation between indicators, but they decrease the possibility of “fuzzifying” the extreme values.

Bounded difference acts as a selective high filter (only high membership values are admitted when the summation of the two membership degrees exceeds 1)

Bounded sum, on the contrary, introduces a relatively low filter with aggregate membership degrees that quickly approach 1.

# Aggregation operators VIII

e) averaging operators: admit compensation between “and/or” goals

$$\mu A(x) = h(\mu A_1(x), \mu A_2(x), \dots, \mu A_n(x))$$

A parametric class of operators is the unweighted generalized means ; writing  $\mu A_1(x)$  simply as  $a_1$

$$h^\alpha = h(a_1, a_2, \dots, a_n) = [a_1^\alpha + a_2^\alpha + \dots + a_n^\alpha / n]^{1/\alpha}$$

$\alpha = 1$  for the arithmetic mean : (e.g.  $(a_1 + a_2) / 2$ )

$\alpha = 0$  for the geometric mean: (e.g.  $\sqrt{a_1 a_2}$ )

# Aggregation operators IX

f) class of wighted averaging operations:

$$h^\alpha = h(a_1, a_2, \dots, a_n; w_1, w_2, \dots, w_n) = [\sum w_i a_i^\alpha]^{1/\alpha}$$

with  $w_i \geq 0$  and  $\sum w_i = 1$

Problem: how to choose  $w_i$  (see OPHI workshop on weighting)

- neutral choice: equal weight to all constitutive elements (HDI)
- frequency-based weighting system. A few example:
  - Desai, Shah (1988): 1-H
  - Cerioli, Zani(1990):  $w_i = \ln 1/f_i$
  - Cheli, Lemmi (1995):  $w_i = \ln[1/n\sum_j \mu_{ij}]$

# Aggregation operators X

How to choose  $w_i$ :

- neutral choice: equal weight to all constitutive elements
- frequency-based weighting system
  - Desai, Shah (1988):  $1-H$
  - Cerioli, Zani(1990):  $w_i = \ln 1/f_i$
  - Cheli, Lemmi (1995):  $w_i = \ln[1/n \sum_j \mu_{ij}]$

# Fuzziness vs probability I

Probability and fuzziness are not competing, they are alternative tools for measuring two different types of uncertainty, and they can complement each other

FST captures a dimension of uncertainty that classical logic and crisp sets, are unable to grasp

# Fuzziness vs probability II

1. stochastic uncertainty or uncertainty as ambiguity: events or statements are well-defined; however, lack of information, time, degree of precision of measurement tools, our ability to use these tools in a proper way, make the choice between two or more alternatives unspecified → probability theory and statistics

# Fuzziness vs probability III

2. uncertainty as vagueness: difficulty in defining sharp boundaries and precise distinctions: fuzzy methodology provides a mathematical tool for dealing with this kind of uncertainty.

# Fuzziness vs probability IV

Probability is negatively related to the amount of information available (when information increases, probability disappears): with perfect and total information there is no uncertainty, and probability values will be equal to zero.

Furthermore, probability is "time-dependent": if a given event (for example, winning a lottery) has a probability of 0.02%, then it is sufficient to wait and see if the event will occur or not.

# Fuzziness vs probability V

On the contrary, fuzziness can be positively related to the amount of information (the more information available, the greater the vagueness may be) and it does not dissipate with time, since it is an intrinsic property of an event or a given object.

# Fuzziness vs probability VI

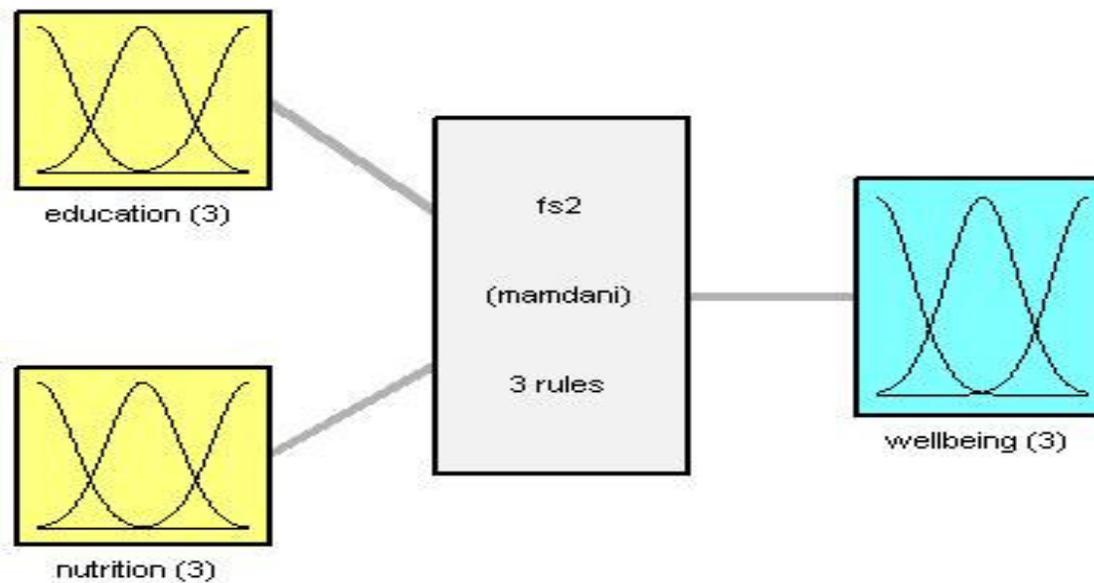
- Membership grades and probabilities take on similar values but they are not necessarily the same thing.

Ex:if a given individual  $i$  belongs to the fuzzy subset of poor people with a membership degree equal to 0.7, this does not mean that he or she has a 70% probability of being poor, but that his or her condition of being poor is vague or fuzzy in a measure of 70%.

# Fuzzy inference

- Fuzzy Inference System is a process based on fuzzy reasoning for mapping an input space to an output space: It involves mf, fuzzy operators and “if-then” rules
- Two main types of inference systems:
  - Mamdani-type: fuzzy sets are combined through aggregation operators and the resulting fs is defuzzified to yield the output of the system
  - Sugeno-type: similar to Mamdani FIS but works only with linear mf

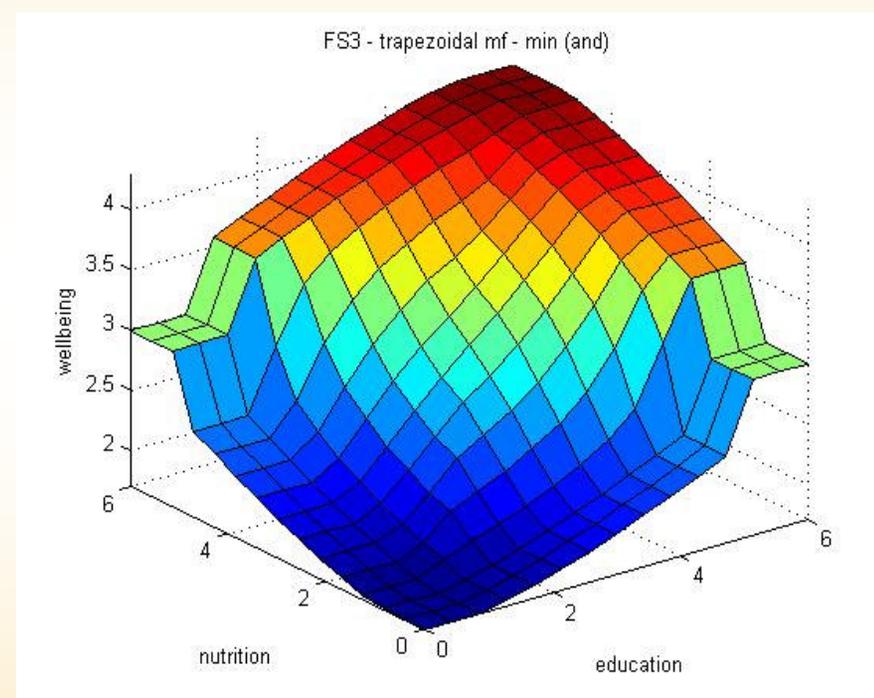
# How a FIS works



System fs2: 2 inputs, 1 outputs, 3 rules

# An example of well-being surfaces based on different mf and rules of aggregation

- trapezoidal mf
- 2 rules
- Standard intersection (and/min) operator



**Slide 36**

---

**ECM1**

nb thresholds in the top and bottom ranks determine a sort of "linear borders"

Enrica Chiappero Martinetti, 09/09/2005

## Some basic readings and textbooks

- Smithson M., Verkuilen J. (2006), Fuzzy set theory. Applications in social sciences, Sage Publications
- Ragin C., Pennings P. (eds.), (2005), Special issues on fuzzy sets and social research, Sociological Methods & Research, 33,4
- Cox E (1994) The fuzzy systems handbook. Academic Press, Chestelnut Hill, Ma
- Klir GJ, Yuan B (1995) Fuzzy sets and fuzzy logic: theory and applications. Prentice Hall, New Jersey
- Ragin CC (2000) Fuzzy set social science. The University of Chicago Press, Chicago