Fuzzy Methods for Constructing Multi-Criteria Decision Functions

Ronald R. Yager

Machine Intelligence Institute

Iona College

ryager@iona.edu

Mixing Words and Mathematics

Building Decision Functions Using Information Expressed in Natural Language

Beginning

Bellman, R. E., & Zadeh, L. A. (1970). Decision-making in a fuzzy environment. Management Science 17:4, 141-164

- Equivalence of Goals and Constraints (Criteria)
- Representation of Criteria as Fuzzy Sets Over the Set of Alternatives
- Linguistic Formulation of Relationship Between Criteria

Linguistic Expression of Multi-Criteria Decision Problem

Satisfy Criteria one and Criteria two and ....

- $D = C_1$ and C_2 and .... and C_n
- "and" as intersection of fuzzy sets
- $D = C_1 \cap C_2 \cap \ldots\ldots \cap C_n$
- $D(x) = \mathrm{Min}_j[C_j(x)]$
- Choose x* with biggest D(x)

Importance Weighting in Multi-Criteria Decision Problem

Yager, R. R. (1978). Fuzzy decision making using unequal objectives. Fuzzy Sets and Systems 1,87-95

- Associate with criteria C_j importance α_j
- $\alpha_j \in [0, 1]$ and $C_j(x) \in [0, 1]$
- $D(x) = \mathrm{Min}_j[(C_j(x))^{\alpha_j}]$
- $\mathrm{Min}[a, 1] = a$ & $(C_j(x))^0 = 1 \Rightarrow$ No effect of $\alpha_j = 0$

  Min[a, b] : Smaller argument more effect

Anxiety In Decision Making

- Alternatives: $X = \{x_1, x_2, x_3, \ldots\ldots., x_q\}$

- Decision function D

  $D(x_j)$ is satisfaction by x_j

- x^* best alternative

- Anxiety associated with selection

$$Anx(D) = D(x^*) - \frac{1}{q-1} \sum_{x_j \neq x^*} D(x_j)$$

Ordinal Scales

- $Z = \{z_0, z_1, z_3, \ldots\ldots\ldots, z_m\}$

  $z_i > z_k$ if $i > k$ (only ordering)

- Operations: Max and Min and Negation

  $Neg(z_j) = z_{m-j}$ (reversal of scale)

- Linguistic values generally only satisfy ordering

  Very High > High > Medium > Low > Very Low

- Often people only can provide information with this type of granulation

Ordinal Decision Making

Yager, R. R. (1981). A new methodology for ordinal multiple aspect decisions based on fuzzy sets. Decision Sciences 12, 589-600

- Criteria satisfactions and importances ordinal

- $\alpha_j \in Z$ and $C_j(x) \in Z$

- $D(x) = \mathrm{Min}_j[G_j(x)]$

  $G_j(x) = \mathrm{Max}(C_j(x), \mathrm{Neg}(\alpha_j))$

- $\alpha_j = z_0 \Rightarrow G_j(x) = z_m$ (No effect on D(x))

  $\alpha_j = z_m \Rightarrow G_j(x) = C_j(x)$

- Linguistic Expression: Satisfy Criteria one **and** Criteria two **and** ....

$$D = C_1 \text{ and } C_2 \text{ and } \ldots\ldots\ldots \text{ and } C_n$$

$$D = C_1 \cap C_2 \cap \ldots\ldots\ldots \cap C_n$$

$$D(x) = \mathrm{Min}_j[C_j(x)]$$

- Linguistic Expression: Satisfy Criteria one **or** Criteria two **or** ....

$$D = C_1 \text{ or } C_2 \text{ or } \ldots\ldots\ldots \text{ or } C_n$$

$$D = C_1 \cup C_2 \cup \ldots\ldots\ldots \cup C_n$$

$$D(x) = \mathrm{Max}_j[C_j(x)]$$

Building M-C Decision Functions

- **Linguistic Expression**

Satisfy Criteria one **and** Criteria two

or

Satisfy Criteria one **or** two **and** criteria 3

or

Satisfy criteria 4 **and** Criteria 3 **or** Criteria 2

- **Mathematical Formulation**

$$D = (C_1 \cap C_2) \cup ((C_1 \cup C_2) \cap C_3) \cup (C_4 \cap (C_3 \cup C_2))$$

Generalizing "and" Operators

t-norm operators generalize "and" (Min)

- $T: [0, 1] \times [0, 1] \rightarrow [0, 1]$
  1. $T(a, b) = T(b, a)$ Commutative
  2. $T(a, b) \geq T(c, d)$ if $a \geq c$ & $b \geq d$ Monotonic
  3. $T(a, T(b, c)) = T(T(a, b), c)$ Associative
  4. $T(a, 1) = a$ one as identity

- Many Examples of t-norms

  $T(a, b) = \text{Min}[a, b]$ $T(a, b) = a\,b$ (product)

  $T(a, b) = \text{Max}(a + b - 1, 0)$

  $$T(a, b) = \text{Max}(1 - ((1 - a)^{\lambda} + (1 - b)^{\lambda})^{\frac{1}{\lambda}}, 0)$$

  Family parameterized by λ

Generalizing “or” Operators

t-conorm operators generalize “or” (Max)

- S: $[0, 1] \times [0, 1] \rightarrow [0, 1]$
  1. $S(a, b) = S(b, a)$ Commutative
  2. $S(a, b) \geq S(c, d)$ if $a \geq c$ & $b \geq d$ Monotonic
  3. $S(a, S(b, c)) = S(S(a, b), c)$ Associative
  4. $S(a, 0) = a$ zero as identity

- Many Examples of t-norms

  $S(a, b) = Max[a, b]$ $S(a, b) = a + b - a\,b$

  $S(a, b) = Min(a + b, 1)$

  $S(a, b) = Min((a^{\lambda} + b^{\lambda})^{\frac{1}{\lambda}}, 1)$

  Family parameterized by λ

Alternative Forms of Basic M-C functions

- $D = C_1$ **and** C_2 **and** …….. **and** C_n
- $D(x) = T_j[C_j(x)]$
- $D(x) = \Pi_j C_j(x)$ (product)
- $D = C_1$ **or** C_2 **or** …….. **or** C_n
- $D(x) = S_j[C_j(x)]$
- $D(x) = \text{Min}(\Sigma_j C_j(x),\ 1]$ (Bounded sum)

- Use of families of t-norms enables a parameterized representation of multi-criteria decision functions

- This opens the possibility of learning the associated parameters from data

C_1	C_2	C_3	C_4	D
.3	.5	1	.7	.5

Generalized Importance Weighted "anding"

- $D = C_1$ **and** C_2 **and** …….. **and** C_n

- Associate with criteria C_j importance α_j

- $D(x) = \mathbf{T}_j[G_j(x)]$

  $G_j(x) = \mathbf{S}(C_j(x), 1 - \alpha_j)$

- $D(x) = \mathrm{Min}_j[(\mathrm{Max}(C_j(x), 1 - \alpha_j))$

  $D(x) = \prod(\mathrm{Max}(C_j(x), 1 - \alpha_j)$

Generalized Importance Weighted "oring"

- $D = C_1$ **or** C_2 **or** .... **or** C_n

- Associate with criteria C_j importance α_j

- $D(x) = \mathbf{S}_j[H_j(x)]$

  $H(x) = T(C_j(x), \alpha_j)$

- $D(x) = Max_j[Min(\alpha_j, C_j(x))]$

  $D(x) = Max_j[\alpha_j C_j(x)]$

  $D(x) = Min(\sum_j \alpha_j C_j(x), 1]$

Some Observations

- If any $C_j(x) = 0$ then

  $\mathbf{T}(C_1(x), C_1(x), \ldots\ldots, C_1(x)) = 0$

- Imperative of this class of decision functions is

  All *criteria must be satisfied*

- If any $C_j(x) = 1$ then

  $\mathbf{S}(C_1(x), C_1(x), \ldots\ldots, C_1(x)) = 1$

- Imperative of this class of decision functions is

  At least one *criteria must be satisfied*

$$D(x) = \frac{1}{n} \sum_{j=1}^{n} C_j(x)$$

Mean Operators

- $\mathbf{M}: R^n \to R$
  1. Commutative
  2. Monotonic

     $\mathbf{M}(a_1, a_2, \ldots\ldots, a_n) \geq \mathbf{M}(b_1, b_2, \ldots\ldots, b_n)$ if $a_j \geq b_j$
  3. Bounded

     $\mathrm{Min}_j[a_j] \leq \mathbf{M}(a_1, a_2, \ldots\ldots, a_n) \leq \mathrm{Max}_j[a_j]$

(Idempotent: $\mathbf{M}(a, a, \ldots\ldots, a) = a$

- Many Examples of Mean Operators

$\mathrm{Min}_j[a_j]$, $\mathrm{Max}_j[a_j]$, Median, Average

OWA Operators

Choquet Aggregation Operators

Ordered Weighted Averaging Operators

OWA Operators

Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Transactions on Systems, Man and Cybernetics 18, 183-190

OWA Aggregation Operators

- Mapping F: $R^n \rightarrow R$ with $F(a_1, \ldots\ldots, a_n) = \sum_{j=1}^{n} w_j \, b_j$
    - b_j is the j^{th} largest of the a_j
    - weights satisfy: **1.** $w_j \in [0, 1]$ and **2.** $\sum_{j=1}^{n} w_j = 1$

- Essential feature of the OWA operator is the reordering operation, **nonlinear operator**

- Weights not associated directly with an argument but with the ordered position of the arguments

- $W = [w_1 \ w_2 \ \quad w_n]$ called the **weighting vector**
- $B = [b_1 \ b_2 \ \quad b_n]$ is **ordered argument vector**
- $F(a_1, \ldots\ldots, a_n) = W \ B^T$
- If id(j) is index of jth largest of a_i then

$$F(a_1, \ldots\ldots, a_n) = \sum_{j=1}^{n} w_j \, a_{id(j)}$$

$$a_{id(j)} = b_j$$

Form of Aggregation is Dependent Upon the Weighting Vector Used

OWA Aggregation is Parameterized by W

Some Examples

- W^*: $w_1 = 1$ & $w_j = 0$ for $j \neq 1$ gives

$$F^*(a_1, \ldots\ldots, a_n) = \mathrm{Max}_i[a_i]$$

- W_*: $w_n = 1$ & $w_j = 0$ for $j \neq n$ gives

$$F^*(a_1, \ldots\ldots, a_n) = \mathrm{Min}_i[a_i]$$

- W_N: $w_j = \frac{1}{n}$ for all j gives the simple average

$$F^*(a_1, \ldots\ldots, a_n) = \frac{1}{n}\sum_{i=1}^{n} a_i$$

Attitudinal Character of an OWA Operator

- $A\text{-}C(W) = \frac{1}{n-1} \sum_{j=1}^{n} w_j (n - j)$

- Characterization of type of aggregation

- $A\text{-}C(W) \in [0, 1]$

- $A\text{-}C(W^*) = 1 \quad A\text{-}C(W_N) = 0.5 \quad A\text{-}C(W_*) = 0$

- Weights symmetric $(w_j = w_{n-j+1}) \Rightarrow A\text{-}C(W) = 0.5$

An A-C value near **one** indicates a bias toward the **larger** values in the argument (**Or-like /Max-like**)

An A-C value near **zero** indicates a bias toward the **smaller** values in the argument (**And-like /Min-like**)

An A-C value **near 0.5** is an indication of a **neutral** type aggregation

Measure of Dispersion an OWA Operator

- $\text{Disp}(W) = - \sum_{j=1}^{n} w_j \ln(w_j)$

- Characterization amount of information used

- $\text{Disp}(W^*) = \text{Disp}(W_*) = 0$ (Smallest value)

  $\text{A-C}(W_N) = \ln(n)$ (Largest value)

- Alternative Measure

  $\text{Disp}(W) = \sum_{j=1}^{n} (w_j)^2$

Some Further Notable Examples

- **Median**: if n is **odd** then $w_{\frac{n+1}{2}} = 1$

  if n is **even** then $w_{\frac{n}{2}} = w_{\frac{n}{2}+1} = \frac{1}{2}$

- **kth best:** $w_k = 1$ then $F^*(a_1, \ldots.., a_n) = a_{id(k)}$

- **Olympic Average**: $w_1 = w_n = 0$, other $w_j = \frac{1}{n - 2}$

- **Hurwicz average**: $w_1 = \alpha$, $w_n = 1-\alpha$, other $w_j = 0$

OWA Operators Provide a Whole family of functions for the construction of mean like multi–Criteria decision functions

$$D(x) = F_W(C_1(x), C_2(x), \ldots\ldots, C_n(x))$$

Selection of Weighting Vector
Some Methods

1. Direct choice of the weights

2. Select a notable type of aggregation

3. Learn the weights from data

4. Use characterizing features

5. Linguistic Specification

Learning the Weights from Data

- Filev, D. P., & Yager, R. R. (1994). Learning OWA operator weights from data. Proceedings of the Third IEEE International Conference on Fuzzy Systems, Orlando, 468-473.

- Filev, D. P., & Yager, R. R. (1998). On the issue of obtaining OWA operator weights. Fuzzy Sets and Systems 94, 157-169.

- Torra, V. (1999). On learning of weights in some aggregation operators: the weighted mean and the OWA operators. Mathware and Softcomputing 6, 249-265

Algorithm for Learning OWA Weights

- Express OWA weights as $w_j = \dfrac{e^{\lambda_j}}{\sum_{k=1}^{n} e^{\lambda_k}}$
- Use data of observations to learn λ_i

  $(a_1, \quad , a_n)$ and aggregated value d
- Order arguments to get b_j for j = 1 to n
- Using current estimate of weights calculate

  $$\hat{d} = \sum_{j=1}^{n} w_j b_j$$
- Updated estimates of λ_j

  $$\lambda'_j = \lambda_j - \alpha w_j (b_i - \hat{d}) (\hat{d} - d)$$

Using Characterizing Features

- $A\text{-}C(W) = \frac{1}{n-1} \sum_{j=1}^{n} w_j (n - j)$

- A-C(W) = 1 "orlike"

  A-C(W) = 0 "andlike"

- $\alpha \in [0, 1]$ degree of "orness"

- Determine W with specified α

O'Hagan Method

- Specify α and determine weights to maximize the dispersion

- $$\text{Max} \; -\sum_{j=1}^{n} w_j \ln(w_j)$$

such that

1. $$\frac{1}{n-1} \sum_{j=1}^{n} w_j (n-j) = \alpha$$
2. $$\sum_{j=1}^{n} w_j = 1$$
3. $$w_j \geq 0$$

Linguistic Specification of Weights

1. Linguistically specify aggregation imperative of multiple criteria

2. Translate linguistic imperative into Fuzzy Set

3. Use fuzzy set to determine OWA weights

Computing with Information Specified in a Natural Language

Quantifier Guided Criteria Aggregation

- D = Min: **All** criteria must be satisfied

  D = Max: ***At least one*** *criteria must be satisfied*

  "Quantifier" criteria must be satisfied

- Other examples of linguistic quantifiers:

  most, almost all, at least half

  only a few, at least 1/3

- Monotonic quantifiers

Representation of Linguistic Quantifier

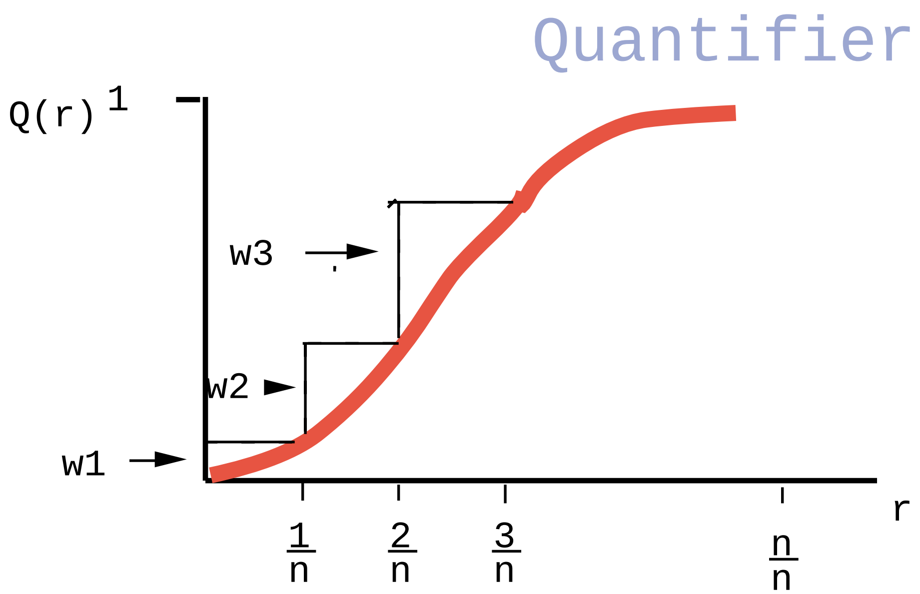
- Represent quantifier as fuzzy subset Q on unit interval

- Q(r) is the degree the proportion r satisfies the concept of the quantifier

- $Q : [0, 1] \rightarrow [0, 1]$
  1. $Q(0) = 0$
  2. $Q(1) = 1$
  3. $Q(r) \geq Q(p)$ if $r > p$

BUM Function

Obtaining OWA Weights from Quantifier



- $w_j = Q(\frac{j}{n}) - Q(\frac{j-1}{n})$

Functionally Guided Criteria Aggregation

- Specify a Bum function f: [0, 1] $\rightarrow$ [0, 1]

  1. $f(0) = 0$
  2. $f(1) = 1$
  3. $f(r) \geq f(p)$ if $r > p$

- $w_j = f(\frac{j}{n}) - f(\frac{j-1}{n})$

- Linear function $f(r) = r$ Quantifier $\Leftrightarrow$ Some

  $w_j = \frac{1}{n}$

Importance Weighted OWA Multi-Criteria Decision Functions

- Importance v_i associated criteria C_i
- Aggregation Agenda

  Quantifier Important Criteria are Satisfied

  Most Important Criteria are Satisfied

- $D(x) = F_{Q/V}(a_1, a_2, \ldots\ldots, a_n)$

  $$a_i = C_i(x)$$

Calculation of $D(x) = F_{Q/V}(a_1, a_2, \ldots\ldots, a_n)$

- Order the criteria satisfactions the a_i
- $a_{id(j)}$ is j^{th} largest & $v_{id(j)}$ its importance
- Calculate $S_j = \sum_{k=1}^{j} v_{id(k)}$ & $T = S_n = \sum_{k=1}^{n} v_{id(k)}$
- Determine OWA Weights

$$\tilde{w}_j = Q(\frac{S_j}{T}) - Q(\frac{S_{j-1}}{T})$$

- $D(x) = \sum_{j=1}^{n} \tilde{w}_j \, a_{id(j)}$

Some Methods of Obtaining Importances

- Fixed Specified Value

- Determined by Property of Alternative

$$v_j = E(x)$$

- Dependent upon Other Attribute in Aggregation

$$v_j = C_k(x)$$

Induces a prioritization

- Rule Based

Concept Based Hierarchical Formulation of Multi-Criteria Decision Functions Using OWA Operators

Definition of a Concept

- Concept is more abstract criteria

$$Con \equiv <C_1, C_2, \ldots, C_n : V : Q>.$$

- C_i are a collection of measurable criteria
- Q is an OWA Aggregation Imperative
- V vector where v_i is importance of C_i in concept
- $Con(x) = F_{Q/V}(C_1(x), C_2(x), \ldots, C_n(x))$

Concepts with Concepts as Components

$$\textbf{Con} = \langle Con_1, Con_2, \ldots, Con_q : V : Q \rangle.$$

$$Con(x) = F_{Q/V}(Con_1(x), Con_2(x), \ldots, Con_q(x))$$

Multi-Criteria Decision Function Viewed as Concept

Allows hierarchical structure for the multi-criteria decision functions

Decision function:

$(C_1 \text{ and } C_2 \text{ and } C_3) \text{ or } (C_3 \text{ and } C_4)$

Represent as concept: $<Con_1, Con_2 : V: Q>$.

Here Q is ***or*** and $V = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Additionally

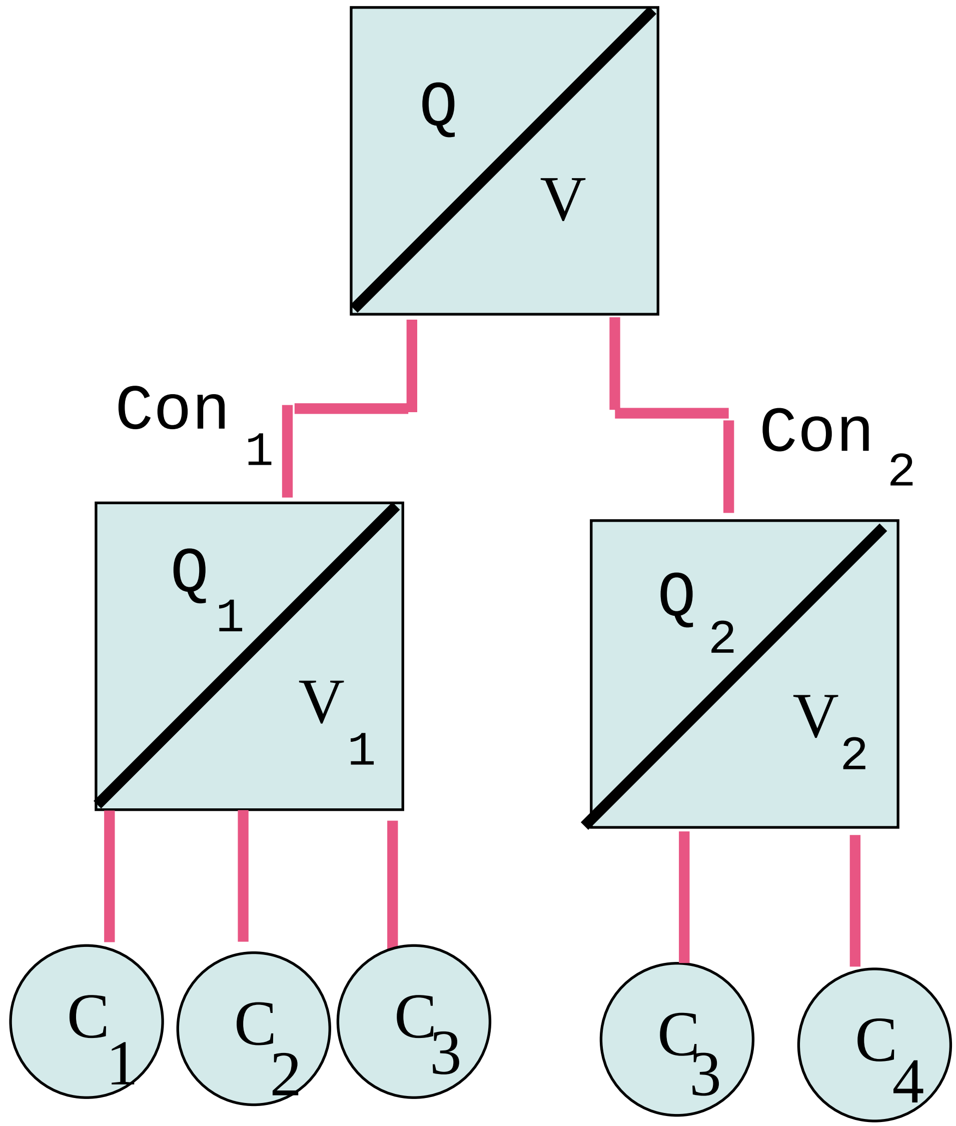
$Con_1 = <C_1, C_2, C_3: V_1: Q_1>$

$Con_2 = <C_3, C_4 : V_2: Q_2>$

Where $Q_1 = Q_2 =$ ***all***

$V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $V_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Hierarchical Formulation



Ordinal OWA Operator

- $Z = \{z_0, z_1, z_3, \ldots\ldots\ldots, z_m\}$ ordinal scale
- Mapping $F: Z^n \rightarrow Z$ with

$$F(a_1, \ldots\ldots, a_n) = \mathrm{Max}_j[w_j \wedge b_j]$$

  - b_j is the jth largest of the a_j
  - weights satisfy: 1. $w_j \in Z$

    2. $w_i \geq w_k$ if $i > j$

    3. $w_n = z_m$

- Allows mean like M-C decision functions with ordinal information

Multi-Criteria Decision Functions Using Choquet Aggregation Operators

- Provides wide class of M-C decision functions
- $C = \{C_1, C_2, \ldots\ldots\ldots, C_n\}$ "set of all criteria"
- Requires specification of monotonic measure μ over set of criteria
- $D(x) = G_\mu(a_1, a_2, \ldots\ldots, a_n)$

$$a_i = C_i(x)$$

Set Measure μ

- For any subset A of criteria, $\mu(A)$ indicates the acceptability of a solution that satisfies all the criteria in A

- $\mu: 2^C \rightarrow [0, 1]$ (subsets of C into the unit interval)
  1. $\mu(\emptyset) = 0$
  2. $\mu(C) = 1$
  3. $\mu(A) \geq \mu(B)$ if $B \subset A$

- $\mu(\emptyset) = 0$ & $\mu(A) = 1$ "any criteria is okay"
  $\mu(C) = 1$ & $\mu(A) = 0$ "all criteria are needed"

Evaluation of Choquet M-C Decision Function

- $D(x) = G_{\mu}(a_1, a_2, \ldots\ldots, a_n) \qquad a_i = C_i(x)$
- Order criteria satisfactions $\Rightarrow$ $a_{id(j)}$ is j^{th} largest
- $H_j = \{C_{id(k)} \mid k = 1 \text{ to } j\}$, j most satisfied criteria
- $w_j = \mu(H_j) - \mu(H_{j-1})$
- $D(x) = G_{\mu}(a_1, a_2, \ldots\ldots, a_n) = \sum_{j=1}^{n} w_j \, a_{id(j)}$

Uninorms

- t-norm operators

$$T(a_1, a_2, \ldots.., a_n) = T(a_1, a_2, \ldots.., a_n, 1)$$

Identity is One

$$T(a_1, a_2, \ldots.., a_n) \geq T(a_1, a_2, \ldots.., a_n, a_{n+1})$$

- t-conorm operators

$$S(a_1, a_2, \ldots.., a_n) \leq S(a_1, a_2, \ldots.., a_n, a_{n+1})$$

Identity is Zero

$$T(a_1, a_2, \ldots.., a_n) = T(a_1, a_2, \ldots.., a_n, 0)$$

- Uninorm operators

  Identity is **e** $\in$ [0, 1]

Uninorm operators with identity **e**

For $a_{n+1} < \mathbf{e}$

$$U(a_1, a_2, \ldots\ldots, a_n) \leq U(a_1, a_2, \ldots\ldots, a_n, a_{n+1})$$

For $a_{n+1} = \mathbf{e}$

$$U(a_1, a_2, \ldots\ldots, a_n) = U(a_1, a_2, \ldots\ldots, a_n, \mathbf{e})$$

For $a_{n+1} > \mathbf{e}$

$$U(a_1, a_2, \ldots\ldots, a_n) \geq U(a_1, a_2, \ldots\ldots, a_n, a_{n+1})$$

M-C Decision Functions Using Uninorms

- Multi-Criteria Decision Function

  $D(X) = U(C_1(x), \ldots, C_n(x))$

- Criteria with satisfaction greater then **e** have positive effect while those less then **e** have negative effect

- Introduces bipolar scale

- **e** acts like "0" in a zero in simple addition

Multi-Criteria Decision Functions Using Fuzzy Systems Modeling

- Set of Criteria C_1, C_2, ..., C_n
- Describe Decision Function D(x)
- If $S.C_1$ is A_{11} and ... $S.C_n$ is A_{1n} then D(x) is d_1

  If $S.C_1$ is A_{m1} and ... $S.C_n$ is A_{mn} then D(x) is d_m
- A_{ij} is fuzzy subset of unit interval

  d_i value in the unit interval

  $S.C_j$ denotes variable "satisfaction of Criteria C_j"

Evaluation of Decision Function by Alternative

- Determine Satisfaction of Rule i by alternative x

$$r_i(x) = \prod_{j=1}^{n} A_{ij}(C_j(x))$$

- Obtain overall satisfaction

$$D(x) = \frac{\sum_{i=1}^{m} r_i(x)\ d_i}{\sum_{i=1}^{m} r_i(x)}$$

Evaluating Criteria Satisfaction $C_j(x)$

- Scalar Number: $C_j(x) = 0.7$
- Ordinal Value: $C_j(x)$ = medium
- Interval Valued : $C_j(x) = [0.3, 0.7]$
- Fuzzy Set Valued: $C_j(x)$ is a fuzzy subset of $[0, 1]$
- Intuitionistic Values: $C_j(x) = (a, b)$ $/a + b \leq 1$

  a degree satisfaction/b degree not satisfaction
- Probabilistic Values: $C_j(x)$ is Probability distribution on $[0, 1]$

THE END