Chapter 1
Information Fusion and Aggregation Operators
Introduction

Vicenç Torra

April 25, 2007

Vicenç Torra, Yasuo Narukawa (2007) Modeling Decisions: Information Fusion and Aggregation Operators, Springer. http://www.springer.com/3-540-68789-0;
http://www.mdai.cat/ifao

Introduction

- Information Fusion and Aggregation Operators
  - To produce the most comprehensive and specific datum about an entity from data supplied by several information sources.

Introduction

- Examples,

$$\sum_{i=1}^{N} a_i/N$$

$$\sum_{i=1}^{N} p_i \cdot a_i$$

Introduction

- Information fusion studies all aspects related to the combination of information

  $\longrightarrow$ it studies information fusion techniques

  – the properties of these techniques, and
  – how to build techniques from properties

- In short,

  – systematize and formalize the fusion process
  – characterize the existing methods

  $\longrightarrow$ with the goal of applying old (and new) methods to new problems

Applications

- Economics
  - Indices to summarize economic information.
    E.g., Retail Price Index, Human Development Index (HDI).
- Biology
  - Fusion of DNA and RNA sequences
  - Combination of taxonomies. E.g. dendrograms, partitions
- Education
  - Assessment of students
  - Evaluation of educational institutions and researchers.
- Computer Science
  - Hardware and software evaluation

Applications

- Robotics
  - Fusion of data from sensors
- Vision
  - Fusion of images
- Knowledge based systems
  - Integration of different kinds of knowledge, verification of correctness (computation of a golden standard), defuzzification (fuzzy KBS), evaluation of solutions (considering different criteria)
- Data mining / machine learning
  - Ensemble methods

Applications

Example. A and B teaching a tutorial+training course w/ constraints

- The total number of sessions is six.
- Professor A will give the tutorial, which should consist of about three sessions; three is the optimal number of sessions; a difference in the number of sessions greater than two is unacceptable.
- Professor B will give the training part,
  consisting of about two sessions.
- Both professors should give more or less the same number of sessions.
  A difference of one or two is half acceptable; a difference of three is unacceptable.

Applications

Example. Formalization

- Variables
  - x_A: Number of sessions taught by Professor A
  - x_B: Number of sessions taught by Professor B
- Constraints
  - the constraints are translated into
    - C_1: $x_A + x_B$ should be about 6
    - C_2: x_A should be about 3
    - C_3: x_B should be about 2
    - C_4: $|x_A - x_B|$ should be about 0
  - using fuzzy sets, the constraints are described ...

Applications

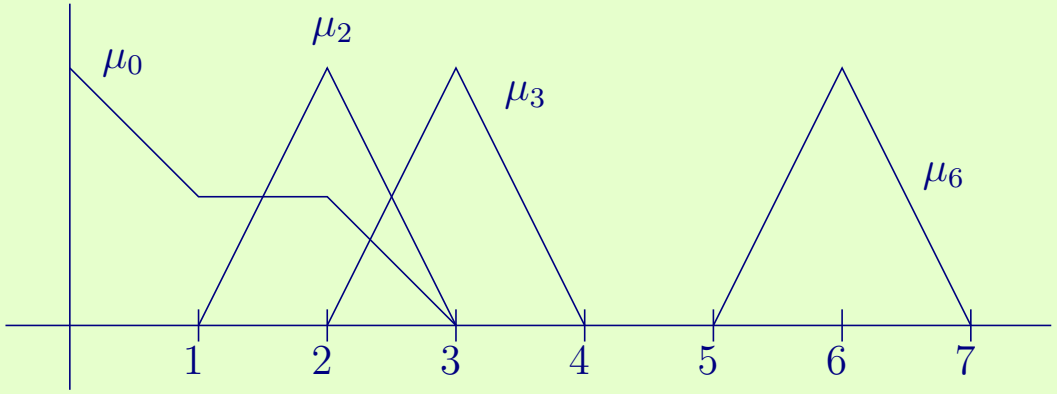
Example. Formalization

- Constraints
  - if fuzzy set μ_6 expresses "about 6," then,
    we evaluate "$x_A + x_B$ should be about 6" by $\mu_6(x_A + x_B)$.
    $\longrightarrow$ given μ_6, μ_3, μ_2, μ_0,
  - Then, given a solution pair (x_A, x_B), the degrees of satisfaction:
    * $\mu_6(x_A + x_B)$
    * $\mu_3(x_A)$
    * $\mu_2(x_B)$
    * $\mu_0(|x_A - x_B|)$

Applications

Example. Formalization

- Membership functions for constraints



Applications

Example. Application

alternative	Satisfaction degrees	Satisfaction degrees			
(x_A, x_B)	$(\mu_6(x_A + x_B),\ \mu_3(x_A),$ $\mu_2(x_B),\ \mu_0(\lvert x_A - x_B\rvert))$	C_1	C_2	C_3	C_4
$(2, 2)$	$(\mu_6(4),\ \mu_3(2),\ \mu_2(2),\ \mu_0(0))$	0	0.5	1	1
$(2, 3)$	$(\mu_6(5),\ \mu_3(2),\ \mu_2(3),\ \mu_0(1))$	0.5	0.5	0.5	0.5
$(2, 4)$	$(\mu_6(6),\ \mu_3(2),\ \mu_2(4),\ \mu_0(2))$	1	0.5	0	0.5
$(3.5, 2.5)$	$(\mu_6(6),\ \mu_3(3.5),\ \mu_2(2.5),\ \mu_0(1))$	1	0.5	0.5	0.5
$(3, 2)$	$(\mu_6(5),\ \mu_3(3),\ \mu_2(2),\ \mu_0(1))$	0.5	1	1	0.5
$(3, 3)$	$(\mu_6(6),\ \mu_3(3),\ \mu_2(3),\ \mu_0(0))$	1	1	0.5	1

Applications

- Two ultimate goals for information fusion applications

(i) To make decisions
(ii) To have a better understanding of the application domain

Applications: To Make Decisions

(i) Decision making:

- Alternative selection:
  - different alternatives / different criteria $\longrightarrow$ select an alternative
    $\longrightarrow$ multicriteria decision making (MCDM) problem
    several criteria (or utility functions) or a single multivalued pref.

Criteria Satisfaction on:

alt	Price	Quality	Comfort
FordT	0.2	0.8	0.3
206	0.7	0.7	0.8
...	...		

alt	Consensus
FordT	0.35
206	0.72
...	...

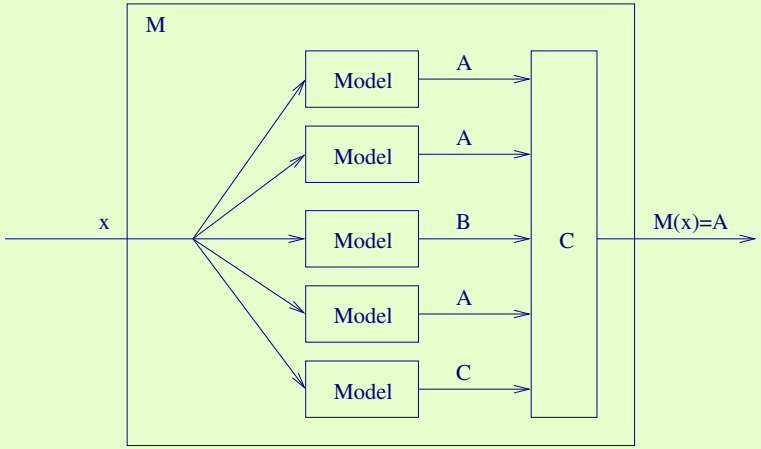
alt	Ranking
206	0.72
FordT	0.35
...	...

  - Solved:
    (i) Aggregate degrees of satisfaction of all criteria
    (ii) Rank the alternatives w.r.t. the global satisfaction degree

Applications: To Make Decisions

(i) Decision making:

- Alternative construction (from a set of them):
  - Several alternatives $\longrightarrow$ a new one
    reliability and constraints of the alternatives
  - Examples
    - Plan merging
      $\longrightarrow$ preconditions and effects as constraints
    - Ensemble methods



Applications: Improving Understanding of the Domain

(ii) Improve understanding of the application domain

- Single source: inconveniences caused by insufficient data quality
  - lack of accuracy (errors by the source / in transmission)
  - lack of reliability
  - too narrow information

  $\longrightarrow$ multiple sources

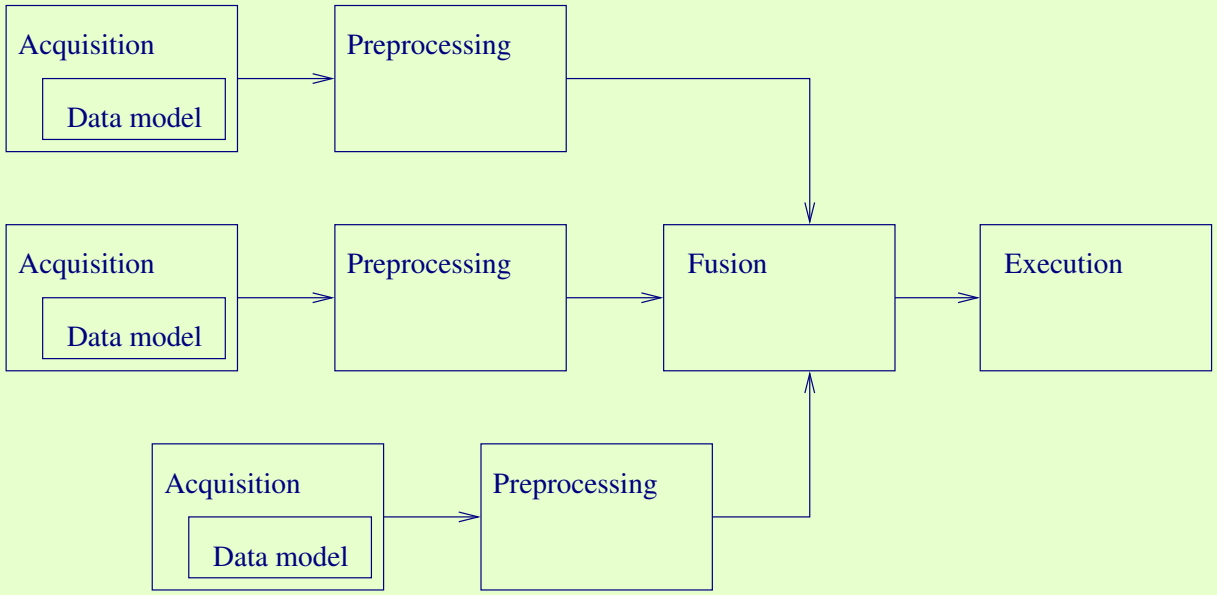
Fusion, integration and aggregation

- Terms
  - Information integration
  - Information fusion: particular functions, methods, procedures, algorithms
  - Aggregation operators: $\mathbb{C} : D^N \to D$ ($\mathbb{C}$ from *Consensus*)
    $\to$ and $\mathbb{C}$ with parameters (background knowledge): $\mathbb{C}_P$

- Aggregation:
  - Unanimity (idempotency): $\mathbb{C}(a, \ldots, a) = a$ for all a
  - Monotonicity: $\mathbb{C}(a_1, \ldots, a_N) \geq \mathbb{C}(a'_1, \ldots, a'_N)$ when $a_i \geq a'_i$
  - Symmetry: for any π permutation on $\{1, \ldots, N\}$,
    $\mathbb{C}(a_1, \ldots, a_N) = \mathbb{C}(a_{\pi(1)}, \ldots, a_{\pi(N)})$
  - Unanimity + monotonicity $\to$ internality:
    $\min_i a_i \leq \mathbb{C}(a_1, \ldots, a_N) \leq \max_i a_i$

An Architecture for Information Integration

- Architecture:



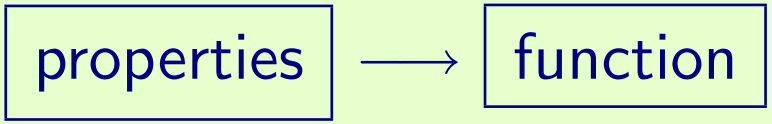
- Stages:
  - Acquisition
  - Preprocessing
  - Fusion
  - Execution

Information Fusion Methods

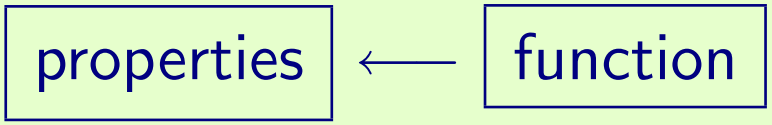
- Perspectives:
  - Type of information:
    - redundant vs. complementary
  - Type of data representation:
    - numerical, ordinal, fuzzy sets, preference relations, dendrograms, partitions, ...
  - Level of abstraction:
    - low level vs. high level

Function Construction

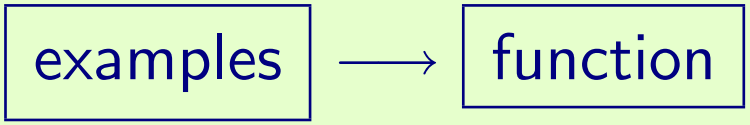
- Definition from properties



- Heuristic definition

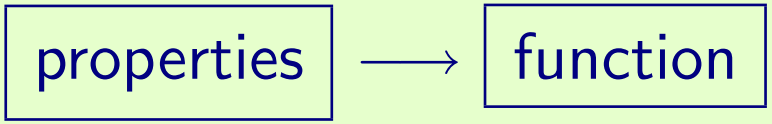


- Definition from examples



Function Construction

- Definition from properties



- Examples

  a) Using functional equations

  b) Aggregation of $a_1, a_2, \ldots, a_N \in D$, as the c located at the minimum distance of the a_i:

$$\mathbb{C}(a_1, a_2, \ldots, a_N) = argmin_c\{\sum_{a_i} d(c, a_i)\},$$

  d is a distance over D.

Goals of Information Fusion

- There are two main goals:
  - Formalization of aggregation processes
  - Study of existing methods

Goals of Information Fusion

- There are two main goals:
  - Formalization of aggregation processes
    - Function definition
    - Function selection
    - Parameter determination
  - Study of existing methods
    - Function characterization
    - Determination of function's modeling capabilities
    - Relationship between operators and parameters