Data Fusion and Evidence Theory

Chiara Foglietta
E-mail: fogliett@dia.uniroma3.it
site: www.dia.uniroma3.it/$\sim$fogliett
Room: 1.17 (first floor - MCIP lab)

University of "Roma Tre"

March, 2013

Purpose of this lesson

The objective of this lesson is:

1. provide a background in data fusion in general

Reading Materials

High Level Data Fusion, by Subrata Das, Artech House, 2008 - 393 pages

Sentz, Kari, and Scott Ferson. *Combination of evidence in Dempster-Shafer theory*. Albuquerque, New Mexico: Sandia National Laboratories, 2002.

Shafer, Glenn. *A mathematical theory of evidence*. Vol. 1. Princeton: Princeton university press, 1976.

Motivation

- Representing ignorance
- The problem of priors
- Symmetric treatment of prior belief & evidence
- Representing evidence:
  1. Evidential basis
  2. Weight of evidence
  3. Uncertain evidence

Introductory Example I

Mr. Jones was assassinated by the Mafia.

Evidence 1

An informer tells the police that the selection of the assassin was done as follow:

A fair coin is tossed:

1. *Head*: either Peter or Tom is selected
2. *Tail*: either Tom or Mary is selected

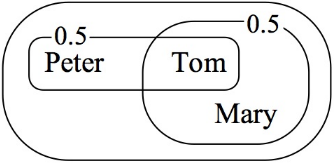
Evidence 2

The police finds the assassin's fingerprint. An expert states that it is male with 80% chance and female with 20% chance.

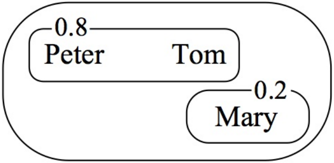
Introductory Example II

Using Evidence Theory

E1

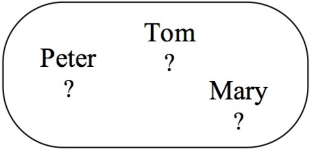


E2



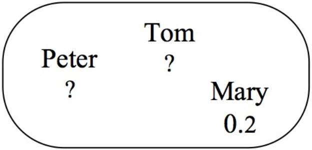
Using probabilities

E1



? $p_1(\text{Peter}) = p_1(\text{Mary}) = 0.25$
$p_1(\text{Tom}) = 0.5$

E2



? $p_2(\text{Peter}) = p_2(\text{Tom}) = 0.4$
$p_2(\text{Mary}) = 0.2$

Uncertainty

Aleatory Uncertainty is the type of uncertainty which results from the fact that a system can behave in random ways.
also known as: Stochastic uncertainty, Type A uncertainty, Irreducible uncertainty, Variability, Objective uncertainty.

Epistemic Uncertainty is the type of uncertainty which results from the lack of knowledge about a system and is a property of the analysts performing the analysis.
also known as: Subjective uncertainty, Type B uncertainty, Reducible uncertainty, State of Knowledge uncertainty, Ignorance

Probability

Laplace's Principle of Insufficient Reason. A probabilistic analysis requires that an analyst have information on the probability of all event. When this is not available, the uniform distribution function is often used.

Axiom of additivity. All probabilities that satisfy specific properties must add to 1. This forces the conclusion that knowledge of an event necessarily entails knowledge of the complement of an event, i.e., knowledge of the probability of the likelihood of the occurrence of an event can be translated into the knowledge of the likelihood of that event not occurring.

Beyond Probability

Where it is not possible to characterize uncertainty with a precise measure such as a precise probability, it is reasonable to consider a measure of probability as an interval or a set.

1. It is not necessary to elicit a precise measurement from an expert or an experiment if it is not realistic or feasible to do so.
2. The Principle of Insufficient Reason is not imposed. Statements can be made about the likelihood of multiple events together without having to resort to assumptions about the probabilities of the individual events under ignorance.
3. The axiom of additivity is not imposed. The measures do not have to add to 1. When the sum is less than 1, called the subadditive case, this implies an incompatibility between multiple sources of information, e.g. multiple sensors providing conflicting information. When the sum is greater than 1, the superadditive case, this implies a cooperative effect between multiple sources of information.

In a finite discrete space, Dempster-Shafer theory can be interpreted as a generalization of probability theory where probabilities are assigned to sets as opposed to mutually exclusive singletons. In traditional probability theory, evidence is associated with only one possible event.
One of the most important features of Dempster-Shafer theory is that the model is designed to cope with varying levels of precision regarding the information and no further assumptions are needed to represent the information.
It also allows for the direct representation of uncertainty of system responses where an imprecise input can be characterized by a set or an interval and the resulting output is a set or an interval.

Evidence

Evidence is a notion which probably can never be fully captured by a single formal theory. Here, “Evidence Theory”, or “Theory of Evidence”, will be understood in a narrow sense as the theory introduced by Dempster and Shafer, and variants thereof.
It is clear today that this theory can be given various different, but essentially equivalent mathematical forms. Some of them are based on probability theory, others are axiomatic theories, a priori without a reference to probabilistic approaches and non-probabilistic ones.

ROMA TRE UNIVERSITÀ DEGLI STUDI

Frame of Discernment

Define $\Omega = \{\omega_1, \cdots, \omega_n\}$ as the set of hypotheses that must be considered as the set of possible value of the variable ω. This set is called *frame of discernment*. For example, the possible causes of failures of a critical infrastructure could be a sabotage, the failure of an appliance, a fault due to the weather, or, for instance, a cyber attack. In the Theory of Evidence the hypotheses are assumed to be mutually exclusive.

Power Set

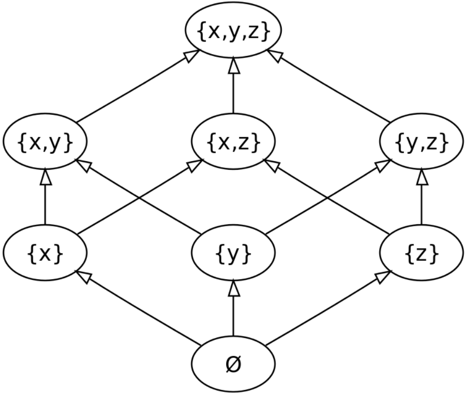
Starting from the frame of discernment, it is possible to define the *power set* as $\Gamma(\Omega) = \{\gamma_1, \cdots, \gamma_{2^{|\Omega|}}\}$, that has cardinality $|\Gamma(\Omega)| = 2^{|\Omega|}$. This set contains all possible subsets of Ω, including the empty set $\gamma_1 = \emptyset$ and the universal set $\gamma_{2^{|\Omega|}} = \Omega$.

ROMA TRE UNIVERSITÀ DEGLI STUDI

Hasse Diagram of Power Set

The Hasse diagram of the power set of three elements, partially ordered by inclusion.

$$\Gamma(\Omega) = \{\emptyset, \{x\}, \{y\}, \{z\}, \{x \cup y\}, \{x \cup z\}, \{y \cup z\}, \{x \cup y \cup z\}, \} \quad (1)$$



Basic Belief Assignment

The Trasferable Belief Model (TBM) [**?**] is based on the definition of a *basic belief mass* function: $m = \Gamma(\Omega) \rightarrow [0; 1.0]$. This function is a map that assigns to each element of the power set a value between 0 and 1. This function, also called *basic belief assignment* (BBA), shall respect the following constraint:

$$\sum_{\gamma_a \subseteq \Gamma(\Omega)} m(\gamma_a) = 1 \quad \text{with} \quad m(\emptyset) = 0 \tag{2}$$

Each element γ_a having $m(\gamma_a) \neq 0$ is named *focal set*.

ROMA TRE UNIVERSITÀ DEGLI STUDI

Understanding Basic Belief Assignment

In this framework, the interest is focused on quantifying the confidence of propositions of the form: "The true value of ω_i is in γ_a," with $\gamma_a \in \Gamma(\Omega)$.
For $\gamma_a \in \Gamma(\Omega)$, $m(\gamma_a)$ is the part of confidence that support exactly γ_a.
This means that the true value is in the set γ_a but, due to lack of further information, we are not able to better support any strictly subset of γ_a.
This is not a probability function, and it does not respect the property of additivity: $m(\gamma_a \cup \gamma_b) \neq m(\gamma_a) + m(\gamma_b)$.

Example I

Suppose two experts are consulted regarding a system failure. The failure could be caused by Component A, Component B or Component C.

The first expert believes that the failure is due to Component A with a probability of 0.99 or Component B with a probability of 0.01.
The second expert believes that the failure is due to Component C with a probability of 0.99 or Component B with a probability of 0.01.

Example II

Frame of Discernment

$$\Omega = \{A, B, C\} \tag{3}$$

Power Set

$$\Gamma(\Omega) = \{\{\emptyset\}, \{A\}, \{B\}, \{C\}, \{A \cup B\}, \{A \cup C\}, \{B \cup C\}, \{A \cup B \cup C\}, \} \tag{4}$$

Basic Probability Assignment

Expert 1 $\Longrightarrow m_1$

$m_1(A) = 0.99$ (failure due to Component A)

$m_1(B) = 0.01$ (failure due to Component B)

Expert 2 $\Longrightarrow m_2$

$m_2(B) = 0.01$ (failure due to Component B)

$m_2(C) = 0.99$ (failure due to Component C)

Belief and Plausibility

From the basic belief assignment, the upper and lower bounds of an interval can be defined. This interval contains the precise probability of a set of interest (in the classical sense) and is bounded by two nonadditive continuous measures called Belief and Plausibility.

The lower bound *Belief* for a set A is defined as the sum of all the basic probability assignments of the proper subsets B of the set of interest A, where $B \subseteq A$.

$$\mathrm{Bel}(A) = \sum_{B, B \subseteq A} m(B) \tag{5}$$

The upper bound, *Plausibility*, is the sum of all the basic probability assignments of the sets B that intersect the set of interest A, where $B \cap A \neq \emptyset$.

$$\mathrm{Pl}(A) = \sum_{B, B \cap A \neq \emptyset} m(B) \tag{6}$$

Example I

Evaluate Belief and Plausibility from BPAs:

	$\emptyset$	A	B	C	$A \cup B$	$A \cup C$	$B \cup C$	$A \cup B \cup C$
m_1	0.0	0.99	0.01	0.0	0.0	0.0	0.0	0.0
Bel_1	0.0	0.99	0.01	0.0	1.0	0.99	0.01	1.0
Pl_1	0.0	0.99	0.01	0.0	1.0	0.99	0.01	1.0
m_2	0.0	0.0	0.01	0.09	0.0	0.0	0.0	0.0
Bel_2	0.0	0.0	0.01	0.09	0.01	0.09	1.0	1.0
Pl_2	0.0	0.0	0.01	0.09	0.01	0.09	1.0	1.0

ROMA TRE UNIVERSITÀ DEGLI STUDI

Inverse Functions I

It is possible to obtain the basic probability assignment from the Belief measure with the following inverse function:

$$m(A) = \sum_{B, B \subseteq A} (-1)^{|A-B|} \mathrm{Bel}(B) \tag{7}$$

where $|A - B|$ is the difference of the cardinality of the two sets.

Inverse Functions II

In addition to deriving these measures from the basic probability assignment m, these two measures can be derived from each other. For example, Plausibility can be derived from Belief in the following way:

$$\mathrm{Pl}(A) = 1 - \mathrm{Bel}(\bar{A}) \tag{8}$$

where $\bar{A}$ is the classical complement of A.

$$\mathrm{Bel}(\bar{\mathrm{A}}) = \sum_{B, B \subseteq \bar{A}} m(B) = \sum_{B, B \cap A = \emptyset} m(B) \tag{9}$$

Dempster's Rule

The *Dempster's rule of combination* is a purely conjunctive operation. This rule strongly emphasises the agreement between multiple sources and ignores all the conflicting evidence through a normalisation factor. This has the effect to attribute null mass to the empty set. So the rule is formalized as:

$$\mathrm{Dempster}\{m_i, m_j\}(\emptyset) = 0$$

$$\mathrm{Dempster}\{m_i, m_j\}(\gamma_a) = \frac{\displaystyle\sum_{\gamma_b \cap \gamma_c = \gamma_a} m_i(\gamma_b) m_j(\gamma_c)}{1 - \displaystyle\sum_{\gamma_b \cap \gamma_c = \emptyset} m_i(\gamma_b) m_j(\gamma_c)} \qquad \forall \gamma_a \in \Gamma(\Omega) \tag{10}$$

Example I

Continuing the example...

		Expert (m_1)		
		$m(A) = 0.99$	$m(B) = 0.01$	$m(C) = 0.0$
	$m(A) = 0.0$	$m_1(A)m_2(A) = 0.0$	$m_1(B)m_2(A) = 0.0$	$m_1(C)m_2(A) = 0.0$
	$m(B) = 0.01$	$m_1(A)m_2(B) = 0.0099$	$m_1(B)m_2(B) = 0.0001$	$m_1(C)m_2(B) = 0.0$
	$m(C) = 0.99$	$m_1(A)m_2(C) = 0.9801$	$m_1(B)m_2(C) = 0.0099$	$m_1(C)m_2(C) = 0.0$

	$m(\emptyset)$	$m(A)$	$m(B)$	$m(C)$
	0.0	$\frac{0.0}{1-0.9999}$	$\frac{0.0001}{1-0.9999}$	$\frac{0.0}{1-0.9999}$
	0.0	0.0	1.0	0.0

Smets' Rule

Differently, the *Smets' rule of combination* allows to express explicitly the contradiction in the TBM, by letting $m(\emptyset) \neq 0$. This combination rule, compared to the Dempster's one, simply avoids the normalisation while preserving the commutativity and associativity properties. The formalization is as follows:

$$\mathrm{Smets}\{m_i, m_j\}(\gamma_a) = m_i(\gamma_a) \otimes m_j(\gamma_a) \qquad \forall \gamma_a \in \Gamma(\Omega) \tag{11}$$

where

$$m_i(\gamma_a) \otimes m_j(\gamma_a) = \sum_{\gamma_b \cap \gamma_c = \gamma_a} m_i(\gamma_b) m_j(\gamma_c) \qquad \forall \gamma_a \in \Gamma(\Omega) \tag{12}$$

ROMA TRE UNIVERSITÀ DEGLI STUDI

Conflict

The fact that $m(\emptyset) > 0$ can be explained in two ways: the open world assumption and the quantified conflict. The open world assumption, made by Dempster, reflects the idea that the frame of discernment must contain the true value. Necessarily, if the open world assumption is true, then the set of hypotheses must contains all possibilities. Under this interpretation, being $\emptyset$ the complement of Ω, the mass $m(\emptyset) > 0$ represents the case where the truth is not contained in Ω. The second interpretation of $m(\emptyset) > 0$ is that there is some underlying conflict between the sources that are combined in order to produce the BBA. Hence, the mass assigned to $m(\emptyset)$ represents the degree of conflict. In particular, it can be computed as follows:

$$m_i(\emptyset) \otimes m_j(\emptyset) = 1 - \sum_{\gamma_a \in \Gamma, \gamma_a \neq \emptyset} (m_i(\gamma_a) \otimes m_j(\gamma_a)) \tag{13}$$

ROMA TRE UNIVERSITÀ DEGLI STUDI

Example I

Continuing the example...

		Expert (m_1)		
		$m(A) = 0.99$	$m(B) = 0.01$	$m(C) = 0.0$
	$m(A) = 0.0$	$m_1(A)m_2(A) = 0.0$	$m_1(B)m_2(A) = 0.0$	$m_1(C)m_2(A) = 0.0$
	$m(B) = 0.01$	$m_1(A)m_2(B) = 0.0099$	$m_1(B)m_2(B) = 0.0001$	$m_1(C)m_2(B) = 0.0$
	$m(C) = 0.99$	$m_1(A)m_2(C) = 0.9801$	$m_1(B)m_2(C) = 0.0099$	$m_1(C)m_2(C) = 0.0$

	$m(\emptyset)$	$m(A)$	$m(B)$	$m(C)$
	0.9999	0.0	0.0001	0.0

ROMA TRE UNIVERSITÀ DEGLI STUDI

Decisions I

according to Smets' "Transferable Belief Model" (TBM)

- There is a *credal level* where beliefs are entertained and a *pignistic level* where beliefs are used to make decisions (from *pignus* = bet in Latin
- At the credal level beliefs are quantified by belief functions
- The credal level precedes the pignistic level in that at any time, belief are entertained and updated at the credal level. The pignistic level appears only when a decision needs to be made
- When decision must be made, beliefs at the credal level induce a probability measure at the pignistic level, i.e., there is a pignistic transformation from belief functions to probability functions.

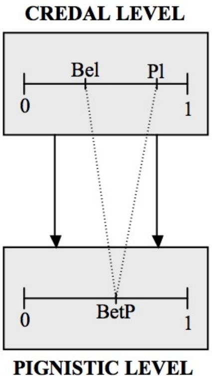
Decisions II

Pignistic Transformation

$$\mathrm{BetP}(\gamma_a) = \sum_{\gamma_b \subseteq \Gamma(\Omega)} \frac{|\gamma_a \cap \gamma_b|}{|\gamma_b|} \frac{m(\gamma_b)}{1 - m(\emptyset)} \qquad (14)$$

$$\mathrm{BetP}(\gamma_a) = \sum_{\gamma_b \subseteq \Gamma(\Omega)} \frac{|\gamma_a \cap \gamma_b|}{|\gamma_b|} m(\gamma_b) \qquad (15)$$

where $|\gamma_a|$ denotes the number of worlds in the set γ_a.



Example

Continuing the example...

Does Pignistic Transformation differ from Basic Probability Assignment?

Why?

Criticism

- The need to evaluate the Power Set
- How can we assign the BBA?

Closed World Assumption

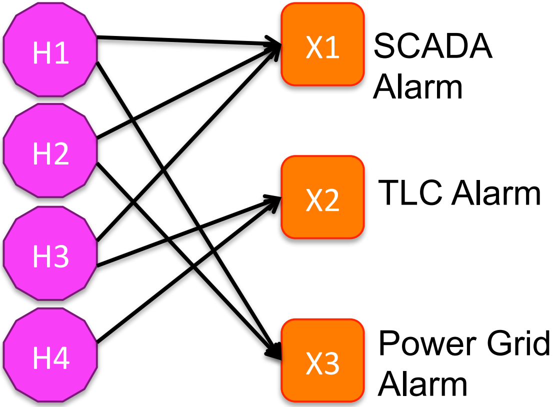
An issue is related to the assumption made in Dempster-Shafer framework: the closed world assumption. According to such an hypothesis, the considered situations have to be exhaustive and mutually exclusive. Smets, with the TBM approach, overcomes this limit including the possibility that empty set may have a non-zero mass; in this way it is possible to determine to which extent the estimation is contradictory. The event of an empty-set with non-zero mass may happen when several combined sources are in conflict, or when the frame of discernment Ω does not contain all possible situations, thus highlighting modeling errors and that the truth is not in Ω.

Exclusivity of Hypotheses

In the Dezert-Smarandache (DSmT) framework, one starts with a frame consisting only in a finite set of exhaustive hypotheses. This is the so-called *free DSm model*. The exclusivity assumption between elements (i.e. requirement for a refinement) of Θ is not necessary within DSmT. However, in DSmT any integrity constraints between elements of Θ can also be introduced, if necessary, depending on the fusion problem under consideration.

ROMA TRE
UNIVERSITÀ DEGLI STUDI

Knowledge Model



Hypotheses

The set of evidences is $[X1, X2, X3]$. Evidence Theory can help finding the most probable cause generating faults. The frame of discernment is composed of the following four hypotheses:

1. (H1) power grid failure;
2. (H2) transportation infrastructure failure;
3. (H3) cyber attack to telecommunication systems;
4. (H4) telecommunication network failure.

ROMA TRE
UNIVERSITÀ DEGLI STUDI

Basic Belief Assignment

For each spurious alarm j generated from the field and received by the control center, we considered the supporting subset Ψ_j, as the sub-set containing all the causes which have an outgoing edge that goes into the j-th failure node. At this subset Ψ_j, it has been assigned a mass α equal to the reliability of the sensor generating the alarm; and $1 - \alpha$ to the universal set $\{H1, H2, H3, H4\}$, as the set representing the maximum ignorance. In this paper, reliability values are pre-fixed values and never change during simulation. For reliability we means the probability that the alarm is a real alarm and not a false one.
The reliability value for each alarm is, respectively, α for X1, β as reliability value for sensor X2, and γ for sensor X3. The alarm X1 supports the subset $\{H1, H2, H3\}$; X2 alarm supports the subset $\{H3, H4\}$ and X3 supports $\{H1, H2\}$, as depicted in Figure **??**.

Right Behaviour I

T0: $v = [\alpha, \beta, 0]$
T1: $v = [\alpha, \beta, 0]$
T2: $v = [\alpha, 0, 0]$
T3: $v = [\alpha, 0, 0]$
where $alpha = 0.6$ and $\beta = 0.9$.

ROMA TRE UNIVERSITÀ DEGLI STUDI

Right Behaviour II

	T0	T1	T2	T3
$\{\emptyset\}$	0	0	0	0
$\{H1\}$	0	0	0	0
$\{H2\}$	0	0	0	0
$\{H3\}$	0.54	0.8316	0.9266	0.9647
$\{H4\}$	0	0	0	0
$\{H1, H2\}$	0	0	0	0
$\{H1, H3\}$	0	0	0	0
$\{H1, H4\}$	0	0	0	0
$\{H2, H3\}$	0	0	0	0
$\{H2, H4\}$	0	0	0	0
$\{H3, H4\}$	0.36	0.1584	0.0634	0.0253
$\{H1, H2, H3\}$	0.06	0.0084	0.0094	0.0097
$\{H1, H2, H4\}$	0	0	0	0
$\{H1, H3, H4\}$	0	0	0	0
$\{H2, H3, H4\}$	0	0	0	0
$\{H1, H2, H3, H4\}$	0.04	0.0016	0.0006	0.0003

ROMA TRE
UNIVERSITÀ DEGLI STUDI

Wrong Behaviour I

- T0: $v = [\alpha, \beta, 0]$
- T1: $v = [\alpha, \beta, 0]$
- T2: $v = [\alpha, 0, 0]$
- T3: $v = [\alpha, 0, 0]$
- T4: $v = [0, 0, \gamma]$
- T5: $v = [0, 0, \gamma]$

where $alpha = 0.6$ and $\beta = 0.9$ and $\gamma = 0.7$.

Wrong Behaviour II

	T0	T1	T2	T3	T4	T5
$\{\emptyset\}$	0	0	0	0	0.693	0.9009
$\{H1\}$	0	0	0	0	0	0
$\{H2\}$	0	0	0	0	0	0
$\{H3\}$	0.54	0.8316	0.9266	0.9647	0.2894	0.0868
$\{H4\}$	0	0	0	0	0	0
$\{H1, H2\}$	0	0	0	0	0.007	0.0091
$\{H1, H3\}$	0	0	0	0	0	0
$\{H1, H4\}$	0	0	0	0	0	0
$\{H2, H3\}$	0	0	0	0	0	0
$\{H2, H4\}$	0	0	0	0	0	0
$\{H3, H4\}$	0.36	0.1584	0.0634	0.0253	0.0076	0.0023
$\{H1, H2, H3\}$	0.06	0.0084	0.0094	0.0097	0.0029	0.0009
$\{H1, H2, H4\}$	0	0	0	0	0	0
$\{H1, H3, H4\}$	0	0	0	0	0	0
$\{H2, H3, H4\}$	0	0	0	0	0	0
$\{H1, H2, H3, H4\}$	0.04	0.0016	0.0006	0.0003	0.0001	0.0000

ROMA TRE
UNIVERSITÀ DEGLI STUDI

The End