Escola Tècnica Superior d'Enginyeria
Universitat Rovira i Virgili

[DΣIM]

Approximate Reasoning

Aida Valls

Introduction to Evidence Theory: Dempster-Shafer model

MESIIA – MIA

Dempster - Shafer Theory of Evidence

- What is the probability that it will rain tomorrow?
  - 55.3678%?
- Basic assumption of probability theory
  - The probabilities of the atomic events can be set with a unique number of arbitrary precision in the range 0 to 1
- However
  - Humans can rarely, if ever, allocate such precise probabilities
  - Humans can rarely, if ever, give their exact independence assumptions
  - Probability assignments from humans are almost always inconsistent
- Dempster - Shafer Theory attempts to deal with this problem
  - Works with degrees of Belief and Plausability
  - Models the way in which the beliefs are transmitted between the various hypotheses involved

Frames of Discernment

- Dempster - Shafer theory assumes a fixed, exhaustive set of mutually exclusive events
  - E = {E1, E2,...En }
    - Same assumption as probability theory
  - Dempster - Shafer theory is concerned with the set of all subsets of E, known as the **Frame of Discernment**
  - A subset {E1, E2, E3} implicitly represents the proposition that one of E1, E2 or E3 is the true event
  - The complete set E represents the proposition that one of the exhaustive set of events is true
    - So E is always true
  - The empty set ϕ represents the proposition that none of the exhaustive set of events is true
    - So ϕ always false

Mass Assignments

- Mass assignment can be viewed as a game
  - The user is given a mass of total weight 1
    - The total amount of belief available
  - The user divides the mass amongst the subsets of E
    - According to the degree of belief that each corresponding proposition is true
      - To the extent that the user is completely uncertain, that proportion of belief is just allocated to the overall environment E, since E by definition must be true
    - Mathematically, a mass assignment m is a function from the power set (set of all subsets) of E to [0,1]
    - The function must satisfy two constraints:
      - $m(\phi) = 0$
      - The sum of m(A), over all subsets A of E, is 1
  - This model generalises Probability theory:
    - In Probability we assume: $m(X) = 0$ unless X is a single-element set

Example (I)

- Example: “The cat in the box is dead.” The observations lead to an assignment of masses: m(Alive)=0.2 and m(Dead)=0.5.
- The remaining mass of 0.3 (the gap between the 0.5+0.2 and 1) is “indeterminate”. This interval represents the level of uncertainty based on the evidence in your system.
- Mass is used to calculate the Belief and Plausibility

Hyphotesis	Mass	Belief	Plausibility
Null	0.0	0.0	0.0
Alive	0.2	0.2	0.5
Dead	0.5	0.5	0.8
Alive or Dead	0.3	1.0	1.0

Belief Functions

- The Belief function of a subset X of A is the total belief that at least X must be true, ie

$$Bel(A) = \sum_{X \subseteq A} m(X)$$

  - Suppose you want to find out how much you should believe that either the cat is "Alive or Dead" is true:
    - you should add in any evidence that you may have that just one of A or D is true, and similarly any information that just A on its own is true, etc

Hypothesis	Mass	Belief	
Null	0.0	0.0	
Alive	0.2	0.2	
Dead	0.5	0.5	
Alive or Dead	0.3	1.0	

Plausibility Functions

- We can define the “doubt” on a proposition A as:

  Doubt(A) = Belief(~A)

- And then the plausibility is the

  Plausibility(A) = 1 - Doubt(A) = 1 - Belief(~A)

- Belief is the minimum degree of support on A
- Plausibility is the upper bound on the degree of support to A

Hypothesis	Mass	Belief	Plausibility
Null	0.0	0.0	0.0
Alive	0.2	0.2	0.5
Dead	0.5	0.5	0.8
Alive or Dead	0.3	1.0	1.0

Confidence interval

Belief and Plausibility

$$\text{Belief}(\emptyset) = 0 \quad \text{Plausibility}(\emptyset) = 0$$

$$\text{Belief}(E) = 1 \ \text{Plausibility}(E) = 1$$

$$\text{Plausibility}(A) >= \text{Belief}(A)$$

- The higher the interval, the higher the uncertainty
- Intervals gives information about the uncertainty level (quality) as well as about the probability (quantity)

[0,1] maximum uncertainty
[0.8, 0.9] low uncertainty, high belief
[0.1, 0.2] low uncertainty, low belief

Combining Evidence

- During the reasoning process we may need to combine the information from different pieces.
- For a subset $A \neq \emptyset$:

  the combined evidence is just the sum of all the different ways in which m1 and m2 combine to give evidence for exactly A

$$[m_1 \oplus m_2](A) = \sum_{X \cap Y = A} m_1(X) * m_2(Y)$$

Problems in Combining Evidence

- Unfortunately, the above approach doesn't work:
  - Because some subsets X and Y don't intersect, so their intersection is the empty set
  - So when we apply the formula, we end up with non-zero mass assigned to the empty set
- The solution is to introduce a denominator, called Conflict Degree:

$$\kappa = \sum_{X \cap Y = \emptyset} m_1(X) * m_2(Y)$$

- Finally, $[m1 \oplus m2](A) = (\Sigma \quad m1(X) * m2(Y)) \; / \; (1 - \kappa)$

Example of evidence combination

- The propagation is made on the basis of the mass function we had m1 and the combination rule

Hyph	m1	m2	m	Bel	Pl
Alive	0.2	0.7	0.41/0.65 =0.63	0.63	0.77
Dead	0.5	0.0	0.15/0.65 =0.23	0.23	0.37
A,D	0.3	0.3	0.09/0.65 =0.14	1.0	1.0

m1+m2	Alive	Dead	Alive,Dead
Alive	0.14 A	0.35 Ø	0.21 A
Dead	0.0 Ø	0.0 D	0.0 D
Alive,Dead	0.06 A	0.15 D	0.09 AD

New info:

m2(Alive) = 0.7

m2(A,D) = 0.3

Summary

- Strong points:
  - Consistent, systematic treatment of lack of knowledge with confidence intervals
  - Gives information about the uncertainty and also about the belief on the different hypothesis
  - Several hypothesis are managed at the same time
- Weak points:
  - Difficulty on designing the “mass assignment” function for each real problem.

Bibliography

- U. Rakowsky Fundamentals of the Dempster-Shafer theory and its applications to system safety and reliability modelling - RTA # 3-4, 2007, December - Special Issue