Belief Measures

Properties of Belief Measures

- Say, the sets $A_1, A_2, \ldots A_n$ in X are pair wise disjoint, i.e. $A_i \cap A_j = \phi$ for all $i, j \in \{1,2,\ldots,n\}$, $i \neq j$
- Subadditivity: $Bel(A_1 \cup A_2) \cup \ldots \cup A_n) \geq Bel(A_1)+Bel(A_2)+\ldots\ldots+Bel(A_n)$
- For n=2; $A_1 \equiv A$ and $A_2 \equiv \sim A$, $Bel(A_1 \cup A_2)=Bel(A \cup \sim A)$

So, $Bel(A \cup \sim A) \geq Bel(A) + Bel(\sim A) - Bel(A \cap \sim A)$

- $A \cup \sim A = X$ and $A \cap \sim A = \phi$, so $Bel(X) \geq Bel(A) + Bel(\sim A)$
  So $Bel(A)+Bel(\sim A) \leq 1$ as $Bel(X)=1$
- Lack of belief in $x \in A$ does not imply a strong belief in $x \in \sim A$
- Plausibility measure Pl is defined as $Pl(A)=1- Bel(\sim A)$
- $Bel(A) = 1- Pl(\sim A)$; $Pl(A) \geq Bel(A)$
- Bel and Pl measures are mutually dual

Basic Probability Assignment

- A Plausibility measure is a function Pl: $\Re \rightarrow [0,1]$ satisfy axioms of fuzzy measures and subadditivity axiom.

- $Pl(A)+Pl(\sim A) \geq 1$ when n=2 and $A \cap \sim A = \phi$

- Belief and Plausibility are expressed in terms of a set function *m*, called a *basic probability assignment*.

- m:P(X)$\rightarrow$[0,1] such that $m(\phi) = 0$ and $\sum_{A \in P(X)} m(A) = 1$

- m(A) is interpreted as the degree of evidence indicating that a specific element of X belongs to the set A
- Any additional evidence of the subsets of A, i.e. $B \subseteq A$, expressed by m(B).

Relationship: *m*(A), Bel(A) and Pl(A)

- *m*(A) measures the belief that one commits to the set A exactly.
- To obtain total belief that one commits to A, one must add to *m*(A) the quantities *m*(B) for all proper subsets B of A.
- Given a *m*, belief and plausibility measures are determined by

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(B) \quad \text{and} \quad \mathrm{Pl}(A) = \sum_{B \cap A \neq \phi} m(B); \text{ for all sets } A \in P(X)$$

$$m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} \mathrm{Bel}(B) \quad ; \ |A-B| \text{ cardinality of the set difference of A and B}$$

Relationship: m(A), Bel(A) and Pl(A)

- m(A) measures the belief that the element ($x \in$A) in question belongs to the set alone, not the total belief that the element commits to A.

- Bel(A) indicates the total evidence that the element belongs to the set A and to the various special subset of A.

- Pl(A) represents the total evidence or belief that the element belongs to the set A or to any of the various special subsets of A plus the additional evidence or belief associated with sets that *overlap* with A.

$$\mathrm{Pl}(A) \geq \mathrm{Bel}(A) \geq m(A) \quad \text{for } \forall A \in P(X)$$

Focal Elements

- For every set A for which $m(A)>0$ is called a *focal element* of *m*
- *Focal elements* are subsets of X on which the available elements focuses
- When X is finite, m can be fully characterized by a list of focal elements A and corresponding values of $m(A)$
- Total ignorance is expressed in terms of the basic assignment by $m(X)=1$ and $m(A)=0$ for all $A \neq X$
- Ignorance states that we know that the element is in the universal set, but we have no evidence about its location in any subset of X

Single Source based Measurement

- Assume a doctor is trying to diagnose a disease by determining whether the patient belongs to one of the sets of people with pneumonia (P), bronchitis (B) or emphysema (E).
- Universal set X denote the set of all possible diseases and say, $X=P \cup B \cup E$
- The doctor after examining the patient, provide the basic assignments to the focal elements such that sum of those equal to 1:

$$P,\ B,\ E,\ P \cup B,\ P \cup E,\ B \cup E,\ P \cup B \cup E$$

Uncertainty and Ignorance

- Bayesian Theory is only concerned on Single evidence.
- Bayesian theory is not able to distinguish between uncertainty and ignorance owing to incomplete information.
- When we do not have any evidence about X, we assign Bel(X) = 0 as well as Bel(¬X) = 0
- It is customary in Dempster-Shafer theory to think about the *degree of belief in evidence* as analogous to the **mass** of a physical object. That is, the mass of evidence supports a belief.
- The reason for the analogy with an object of mass is to consider belief as a quantity that can move around, be split up, and combined.

- A fundamental difference between Dempster-Shafer theory and probability theory is the treatment of **ignorance**.
- Probability theory must distribute an equal amount of probability even in ignorance.
- If you have no prior knowledge, then you must assume the probability p of each possibility is

  $p = 1/N$ where N is the number of possibilities
- The Dempster-Shafer theory does not force belief to be assigned to ignorance or refutation of a hypothesis.
- The mass is assigned only to those subsets of the environment to which you wish to assign belief.
  - Consider a set of propositions as a whole
  - Assign a set of propositions an interval [belief, plausibility] to constraint the degree of belief for each individual propositions in the set

Rules of Combination

- Combine multiple basic assignments in order to get a joint basic probability assignment which is equivalent to a group decision from various experts based on the evidence or belief from each of them independently.
- Conjunction operation (AND, $\cap$): All sources are reliable
- Disjunctive operation (OR, $\cup$): Only one reliable source among many

DST is an evidence theory, it combines all possible outcomes of the problem. Hence it is used to solve problems where there may be a chance that a different evidence will lead to some different result.

The **uncertainity in this model** is given by:-

1. Consider all possible outcomes.
2. **Belief will lead to believe in some possibility by bringing out some evidence.**
3. Plausibility will make evidence compatible with possible outcomes.

Dempster's Rule Of Combinations

- In 1976, Dempster proposed a standard way of combining evidence from two independent sources (B ad C), termed as Dempster's rule of combination
- Q – an exhaustive set of mutually exclusive hypotheses
- A – a subset of Q
- Given two basic assignments m_1 and m_2 on $\wp$(X), from two independent sources.
- Total mass in support of A: $\sum_{B \cap C = A} m_1(B) * m_2(C)$
- Mass of Conflict: $\sum_{B \cap C = \phi} m_1(B) * m_2(C)$

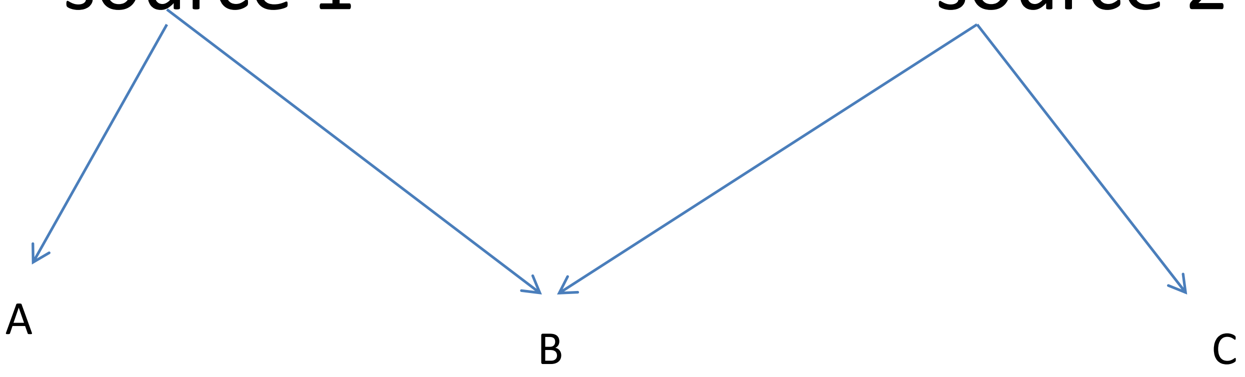
Dempster's Rule

- The combined value of evidence $m_{1,2}(A)$ according to this rule is given by:

$$m_{12}(A) = \frac{\sum_{B \cap C = A} m_1(B) * m_2(C)}{1 - K}$$

- Conflict is normalized out by dividing 1– K, where K is the conflict

- source 1 source 2



Source 1: $m_{1(}A) = 0.99$ (component A)
Source 1: $m_1(B) = 0.01$(component B)

Source 2: $m_2(B) = 0.01$ (component B)
Source 2: $m_2(C) = 0.99$ (component C)

Dempster 's Rule: Support: $m_1(B) * m_2(B)=0.0001$
Conflict: $m_1(A)m_2(B)+m_1(A)m_2(C)+m_1(B)m_2(C)=0.99*0.01+0.99*0.01+0.99*0.99= 0.9999$
Answer: $m_{12}(B)=0.01*0.01/1-0.9999=1$

Problem

- Let us consider a room where four people are present, A, B, C and D. Suddenly the lights go out and when the lights come back, B has been stabbed in the back by a knife, leading to his death. No one came into the room and no one left the room. We know that B has not committed suicide. Now we have to find out who the murderer is.

Solution

- To solve these there are the **following possibilities**:
  - Either {A} or {C} or {D} has killed him.
  - Either {A, C} or {C, D} or {A,D} have killed him.
  - Or the three of them have killed him i.e; {A, C, D}
  - None of them have killed him {o} (let's say).

- There will be the possible evidence by which we can find the murderer by measure of plausibility.
  Using the above example we can say:
  Set of possible conclusion (P): {p1, p2….pn}
  where P is set of possible conclusions and cannot be exhaustive, i.e. at least one (p)i must be true.
  (p)i must be mutually exclusive.
  Power Set will contain 2^n elements where n is number of elements in the possible set.
  For eg:-
  If P = { a, b, c}, then Power set is given as
  {o, {a}, {b}, {c}, {a, b}, {b, c}, {a, c}, {a, b, c}}= 2^3 elements.

- **Mass function m(K):** It is an interpretation of m({K or B}) i.e; it means there is evidence for {K or B} which cannot be divided among more specific beliefs for K and B.
- **Belief in K:** The belief in element K of Power Set is the sum of masses of element which are subsets of K. This can be explained through an example
  Lets say K = {a, b, c}
  Bel(K) = m(a) + m(b) + m(c) + m(a, b) + m(a, c) + m(b, c) + m(a, b, c)
- **Plausibility in K:** It is the sum of masses of set that intersects with K.
  i.e; Pl(K) = m(a) + m(b) + m(c) + m(a, b) + m(b, c) + m(a, c) + m(a, b, c)

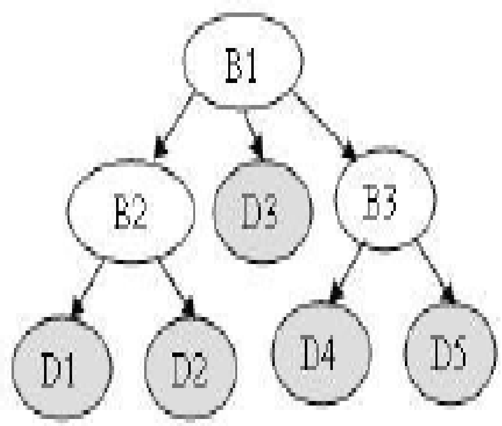
- Ignorance is reduced in this theory by adding more and more evidences.
- Combination rule is used to combine various types of possibilities.

Diagnose Diseases

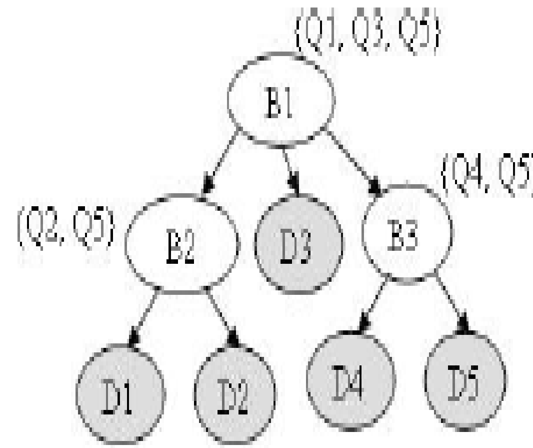
- Due to the combined effect of dearth of experts and diversity of symptoms in diseases, an increasing need of knowledge based decision making system has been felt in the health care domain.
- The existing knowledge-based query-oriented systems consist of static set of queries which are manually affixed.
- When the patients come with subjective symptoms, a large number of questions are asked to diagnose the underlying disease properly.
- During each stage of questioning, the doctor in a systematic way utilizes knowledge to differentiate particular disease(s) from a broader domain of possible diseases.
- In reality, based on the response of a patient to a particular query, the doctor asks the next query, most relevant to the previous response.

- Framework that arranges queries hierarchically to facilitate selection of most suitable query dynamically, based on the patient’s response.
- The sets of queries selected in hierarchy with enough differentiating factors, eventually lead to a fair provisional diagnosis with promising degree of accuracy.
- Fuzzy measures for each possible option relevant to a particular query are calculated using b. p. a.
- Multiple experts’ opinions are fused using Dempster’s rule of combination to obtain fuzzy measures used for selecting the most suitable next query.

- The B_i node represents the broader domains of diseases which can be subdivided while the D_i node denotes the fields representing final diagnosis of disease.



Disease Hierarchy



Query Hierarchy with Disease Hierarchy

- Suppose, for a query Q_1, the *b. p. a.* values are provided by two experts, given in Table 1 and Table 2.
- Each O_j $(1 \leq j \leq 3)$ represents all possible options for query Q_1 and each D_i $(1 \leq i \leq 5)$ is a representative disease.
- *Now* combined b. p. a. *is calculated and Table 3 shows final b. p. a. for query* Q_1.

Table1. Expert 1

	O_1	O_2	O_3
D_1	m_{11}	m_{12}	m_{13}
D_2	m_{21}	m_{22}	m_{23}
D_3	m_{31}	m_{32}	m_{33}
D_4	m_{41}	m_{42}	m_{43}
D_5	m_{51}	m_{52}	m_{53}

Table2. Expert 2

	O_1	O_2	O_3
D_1	n_{11}	n_{12}	n_{13}
D_2	n_{21}	n_{22}	n_{23}
D_3	n_{31}	n_{32}	n_{33}
D_4	n_{41}	n_{42}	n_{43}
D_5	n_{51}	n_{52}	n_{53}

Table3. Combined b. p. a.

	O_1	O_2	O_3
D_1	v_{11}	v_{12}	v_{13}
D_2	v_{21}	v_{22}	v_{23}
D_3	v_{31}	v_{32}	v_{33}
D_4	v_{41}	v_{42}	v_{43}
D_5	v_{51}	v_{52}	v_{53}