Dempster-Shafer Theory

Dempster-Shafer Theory starts by assuming a *Universe of Discourse* θ, also called a *Frame of Discernment*, which is a set of mutually exclusive alternatives. Given the example of determining the disease of a patient, θ would be the set consisting of all possible diseases.

Elements of 2^θ, i.e. subsets of θ are the class of general propositions in the domain. For example, the proposition "The disease is infectious" corresponds to the set of the elements of θ which are infectious, i.e. {"Influenza", "Small Pox",…}.

A function $m:2^\theta \rightarrow [0,1]$ is called a *basic probability assignment* if it satisfies m(∅)=0 and

$$\sum_{A \subseteq \theta} m(A) = 1$$

The quantity m(A) is defined as A's *basic probability number*. It represents the strength of some evidence; our exact belief in the proposition represented by A.

A function $m:2^\theta \rightarrow [0,1]$ is called a *belief function* if it satisfies Bel(∅)=0, Bel(θ)=1, and for any collection $A_1 \ldots A_n$ of subsets of θ

$$Bel(A_1 \cup \ldots \cup A_n) \geq \sum_{\substack{I \subseteq \{1 \ldots n\} \\ I \neq \phi}} (-1)^{|I|+1} Bel\left(\bigcap_{i \in I} A_i\right)$$

A belief function assigns to each subset of θ a measure of our total belief in the proposition represented by the subset.

There corresponds to each belief function one and only one basic probability assignment. Conversely, there corresponds to each basic probability assignment one and only one belief function. They are related by the following two formulae:

$$Bel(A) = \sum_{B \subseteq A} m(B) \text{, for all } A \subseteq \theta$$

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

Thus a belief function and a basic probability assignment convey exactly the same information. Corresponding to each belief function are three other commonly used quantities that convey the same information:

A function $Q:2^\theta \rightarrow [0,1]$ is called a *commonality function* if there is a basic probability assignment, m, such that

$$Q(A) = \sum_{A \subseteq B} m(B) \text{ for all } A \subseteq \theta$$

The *doubt function* is given by

$$Dou(A) = Bel(\neg A)$$

and the *upper probability function* is given by

$$P^*(A) = 1 - Dou(A)$$

This expresses how much we should believe in A if all currently unknown facts were to support A.
This the true belief in A will be somewhere in the interval $[Bel(A), P^*(A)]$
Evidence may be combined using Dempster-Shafer theory. Let m_1 and m_2 be basic probability assignments on the same frame, θ, and define $m=m_1 \oplus m_2$, their *orthogonal sum*, to be $m(\varnothing)=0$ and

$$m(A) = K \sum_{X \cap Y = A} m_1(X) \bullet m_2(Y)$$
$$K^{-1} = 1 - \sum_{X \cap Y = \phi} m_1(X) \bullet m_2(Y) = \sum_{X \cap Y \neq \phi} m_1(X) \bullet m_2(Y)$$

when $A \neq \varnothing$. The function m is a basic probability assignment if $K^{-1} \neq 0$; if $K^{-1}=0$ then $m_1 \oplus m_2$ does not exist and m_1 and m_2 are said to be *totally* or *flatly contradictory*.

The quantity Log K = Con(Bel_1,Bel_2) is called the *weight of conflict* between Bel_1 and Bel_2.

The formula for the orthogonal sum of more than two belief functions follows. Let $m=m_1 \oplus \ldots \oplus m_n$, then $m(\varnothing)=0$ and

$$m(A) = K \sum_{\cap A_i = A} \prod_{1 \le i \le n} m_i(A_1)$$
$$K^{-1} = 1 - \sum_{\cap A_i = \phi} \prod_{1 \le i \le n} m_i(A_1) = \sum_{\cap A_i \neq \phi} \prod_{1 \le i \le n} m_i(A_1)$$

Dempster-Shafer Example

The above formulae are best illustrated through an example. Assume θ is {H, C, P} and the following probability assignments are given:

m({H}) = 0.3
m({H, C}) = 0.2
m({H, C, P}) = 0.5

The belief, Bel(A) of any subset of θ is calculated by adding all m(B) where B is a subset of A, so the beliefs for a few subsets may be calculated as:

Bel({H}) = 0.3
Bel({H, C}) = 0.3 + 0.2 = 0.5
Bel({H, P}) = 0.3
Bel({H, C, P}) = 0.3 + 0.2 + 0.5 = 1.0

The commonality function, Q(A) of any subset of θ is calculated by adding all m(B) where A is a subset of B:

Q({H}) = 0.3 + 0.2 + 0.5 = 1
Q({H, C}) = 0.2 + 0.5 = 0.7
Q({H, P}) = 0
Q({H, C, P}) = 0.5

The doubt function may be calculated as:

Dou({H}) = Bel({C,P}) = 0
Dou({H, C}) = Bel({P}) = 0
Dou({H, P}) = Bel({C}) = 0
Dou({H, C, P}) = Bel(∅) = 0

Finally, the upper probability function calculated for these subsets is:
P^*({H}) = 1 - Dou({H}) = 1
P^*({H, C}) = 1 - Dou({H, C}) = 1
P^*({H, P}) = 1 - Dou({H, P}) = 1
P^*({H, C, P}) = 1 - Dou({H, C, P}) = 1

Combining beliefs is somewhat more complex. Taking a less complex example, assume the following two sets of probability assignments are given for a universe of discourse, θ = {D, D'}:

m_1({D})=0.8	m_2({D})=0.9
m_1 ({D'})=0	m_2 ({D'})=0
m_1 ({D, D'})=0.2	m_2 ({D, D'})=0.1

The solution can be more clearly illustrated if a table is created with rows and columns named by subsets of θ:

		m_2		
		{D}: 0.9	{D'}: 0	{D, D'}: 0.1
m_1	{D}: 0.8	0.72	0	0.08
	{D'}: 0	0	0	0
	{D, D'}: 0.2	0.18	0	0.02

Firstly K is calculated. By definition

$$K^{-1} = 1 - \sum_{X \cap Y = \phi} m_1(X) \bullet m_2(Y)$$

= 1 - (0 + 0)
= 1

There are only two cases where $X \cap Y = \varnothing$: The case $m_1(\{D\}) \cap m_2(\{D'\})$, and the case $m_1(\{D'\}) \cap m_2(\{D\})$. From the table above, the product of both of these is zero. Therefore $K^{-1}=1$, so K=1.

For each probability we want to combine the following formula is used:

$$m(A) = K \sum_{X \cap Y = A} m_1(X) \bullet m_2(Y)$$

Therefore:

$m_1 \oplus m_2(\{D\}) = (1)\ (0.72 + 0.08 + 0.18) = 0.98$
$m_1 \oplus m_2(\{D'\}) = (1)\ (0) = 0$
$m_1 \oplus m_2(\{D, D'\}) = (1)\ (0.02) = 0.02$

So given the evidence presented by m_1 and m_2, we can state that the most probable belief for this universe of discourse is D.

We can also state that the weight of conflict between m_1 and m_2 is Log 1 = 0. Therefore the evidence given by m_1 and m_2 does not contradict.