The Dempster–Shafer Theory

Malcolm J. Beynon
Cardiff University, UK

INTRODUCTION

The initial work introducing Dempster-Shafer (D-S) theory is found in Dempster (1967) and Shafer (1976). Since its introduction the very name causes confusion, a more general term often used is belief functions (both used intermittently here). Nguyen (1978) points out, soon after its introduction, that the rudiments of D-S theory can be considered through distributions of random sets. More furtive comparison has been with the traditional Bayesian theory, where D-S theory has been considered a generalisation of it (Schubert, 1994). Cobb and Shenoy (2003) direct its attention to the comparison of D-S theory and the Bayesian formulisation. Their conclusions are that they have the same expressive power, but that one technique cannot simply take the role of the other.

The association with artificial intelligence (AI) is clearly outlined in Smets (1990), who at the time, acknowledged the AI community has started to show interest for what they call the Dempster-Shafer model. It is of interest that even then, they highlight that there is confusion on what type of version of D-S theory is considered. D-S theory was employed in an event driven integration reasoning scheme in Xia *et al.* (1997), associated with automated route planning, which they view as a very important branch in applications of AI. Liu (1999) investigated Gaussian belief functions and specifically considered their proposed computation scheme and its potential usage in AI and statistics. Huang and Lees (2005) apply a D-S theory model in natural-resource classification, comparing with it with two other AI models.

Wadsworth and Hall (2007) considered D-S theory in a combination with other techniques to investigate site-specific critical loads for conservation agencies. Pertinently, they outline its positioning with respect to AI (p. 400);

The approach was developed in the AI (artificial intelligence) community in an attempt to develop systems that could reason in a more human manner and par-ticularly the ability of human experts to "diagnose" situations with limited information.

This statement is pertinent here, since emphasis within the examples later given is more towards the general human decision making problem and the handling of ignorance in AI. Dempster and Kong (1988) investigated how D-S theory fits in with being an artificial analogy for human reasoning under uncertainty.

An example problem is considered, the murder of Mr. White, where witness evidence is used to classify the belief in the identification of an assassin from considered suspects. The numerical analyses presented exposit a role played by D-S theory, including the different ways it can act on incomplete knowledge.

BACKGROUND

The background section to this article covers the basic formulisations of D-S theory, as well as certain developments. Formally, D-S theory is based on a finite set of p elements $\Theta = \{s_1, s_2, ..., s_p\}$, called a frame of discernment. A *mass value* is a function m: $2^\Theta \to [0, 1]$ such that $m(\varnothing) = 0$ ($\varnothing$ - the empty set) and:

$$\sum_{s \in 2^\Theta} m(s) = 1$$

(2^Θ - the power set of Θ). Any proper subset s of the frame of discernment Θ, for which $m(s)$ is non-zero, is called a focal element and represents the exact belief in the proposition depicted by s. The notion of a proposition here being the collection of the hypotheses represented by the elements in a focal element.

In the original formulisation of D-S theory, from a single piece of evidence all assigned mass values sum to unity and there is no belief in the empty set. In the case of the Transferable Belief Model (TBM), a fundamental development on the original D-S theory (see Smets and Kennes, 1994), a non-zero mass value

can be assigned to the empty set allowing $m(\varnothing) \geq 0$. The set of mass values associated with a single piece of evidence is called a *body of evidence* (BOE), often denoted $m(\cdot)$. The mass value $m(\Theta)$ assigned to the frame of discernment Θ is considered the amount of ignorance within the BOE, since it represents the level of exact belief that cannot be discerned to any proper subsets of Θ.

D-S theory also provides a method to combine the BOE from different pieces of evidence, using Dempster's rule of combination. This rule assumes these pieces of evidence are independent, then the function $(m_1 \oplus m_2): 2^\Theta \to [0, 1]$, defined by:

$$(m_1 \oplus m_2)(x) =$$

$$\begin{cases} 0 & x = \varnothing \\ \dfrac{\sum\limits_{s_1 \cap s_2 = x} m_1(s_1) m_2(s_2)}{1 - \sum\limits_{s_1 \cap s_2 = \varnothing} m_1(s_1) m_2(s_2)} & x \neq \varnothing \end{cases}$$

$$(1)$$

is a mass value, where s_1 and s_2 are focal elements from the BOEs, $m_1(\cdot)$ and $m_2(\cdot)$, respectively. The denominator part of the combination expression includes:

$$\sum\limits_{s_1 \cap s_2 = \varnothing} m_1(s_1) m_2(s_2)$$

that measures the level of conflict in the combination process (Murphy, 2000). It is the existence of the denominator part in this combination rule that separates D-S theory (includes it) from TBM (excludes it). Benouhiba and Nigro (2006) view this difference as whether considering the conflict mass:

$$\left(\sum\limits_{s_1 \cap s_2 = \varnothing} m_1(s_1) m_2(s_2) \right)$$

as a further form of ignorance mass is an acceptable point of view.

D-S theory, along with TBM, also differs to the Bayesian approach in that it does not necessarily produce final results. Moreover, partial answers are present in the final BOE produced (through the combination of evidence), including focal elements with more than one element, unlike the Bayesian approach where probabilities on only individual elements would be accrued. This restriction of the Bayesian approach to consider singleton elements is clearly understood through the 'Principle of insufficient Reason', see Beynon *et al.* (2000) and Beynon (2002, 2005).

To enable final results to be created with D-S theory, a number of concomitant functions exist with D-S theory, including;

i)   The Belief function,

$$\text{Bel}(s_i) = \sum\limits_{s_j \subseteq s_i} m(s_j)$$

for all $s_i \subseteq \Theta$, representing the confidence that a proposition y lies in s_i or any subset of s_i,

ii)   The Plausibility function,

$$\text{Pls}(s_i) = \sum\limits_{s_j \cap s_i \neq \varnothing} m(s_j)$$

for all $s_i \subseteq \Theta$, represents the extent to which we fail to disbelieve s_i,

iii)   The Pignistic function (see Smets and Kennes, 1994),

$$\text{BetP}(s_i) = \sum\limits_{s_j \subseteq \Theta, s_j \neq \varnothing} m(s_j) \frac{|s_i \cap s_j|}{|s_j|},$$

for all $s_i \subseteq \Theta$, represents the extent to which we fail to disbelieve s_i.

From the definitions given above, the Belief function is cautious of the ignorance incumbent in the evidence, where as the Plausibility function is more inclusive of its presence. The Pignistic function acts more like a probability function, partitioning levels of exact belief (mass) amongst the elements of the focal element it is associated with.

Related Content

Machine Learning and Deep Learning for Applications: A Hands-On Study With Python
Naciye Celebi, Tze-Li Hsuand Qingzhong Liu (2022). *Applications of Machine Learning and Artificial Intelligence in Education (pp. 1-47).*
www.irma-international.org/chapter/machine-learning-and-deep-learning-for-applications/299219

Applying CI in Biology through PSO
Rojalina Priyadarshini, Nilamadhab Dash, Brojo Kishore Mishraand Rachita Misra (2016). *Handbook of Research on Computational Intelligence Applications in Bioinformatics (pp. 119-143).*
www.irma-international.org/chapter/applying-ci-in-biology-through-pso/157485

Socrates Digital™ System Architecture: Overview of the Socrates Digital™ Module
 (2022). *Socrates Digital™ for Learning and Problem Solving (pp. 87-105).*
www.irma-international.org/chapter/socrates-digital-system-architecture/290565

Conceptually Advancing "Application Mobility" Towards Design: Applying a Concept-Driven Approach to the Design of Mobile IT for Home Care Service Groups
Dan Johanssonand Mikael Wiberg (2012). *International Journal of Ambient Computing and Intelligence (pp. 20-32).*
www.irma-international.org/article/conceptually-advancing-application-mobility-towards/68842

A Hybrid Model for Service Selection in Semantic Web Service Composition
Sandeep Kumarand R.B. Mishra (2008). *International Journal of Intelligent Information Technologies (pp. 55-69).*
www.irma-international.org/article/hybrid-model-service-selection-semantic/2443