The t-norm valuation-based system for the database attributes configuration in the Dempster-Shafer's theory of evidence

Anna Bryniarska

Institute of Mathematics and Informatics

Opole University

abryniarska@math.uni.opole.pl

Agenda

- Motivation, problem and the aim of research
- Valuations for the database attributes configuration
- Marginalization
- Combination valuations
- Attribute valuations network (AVN)
- Atomic valuations
- Relation of part order in AVN
- T-norm valuation system
- Logic Algebra LA
- Conclusion

Motivation

AI Magazine Volume 7 Number 2 (1986) (© AAAI)

A Simple View of the Dempster-Shafer Theory of Evidence and its Implication for the Rule of Combination

Lotfi A. Zadeh

Computer Science Division, University of California, Berkeley, California 94720

Appeared in : Yager, R. R., M. Federizzi, and J. Kacprzyk, eds., Advances in the Dempster-Shafer Theory of Evidence, 1994, pp. 395–414, John Wiley & Sons, New York, NY.

USING DEMPSTER-SHAFER'S BELIEF-FUNCTION THEORY IN EXPERT SYSTEMS

by

Prakash P. Shenoy

School of Business, University of Kansas, Lawrence, KS 66045-7585, USA
Tel: (785)-864-7551, Fax: (785)-864-5328, Email: pshenoy@ku.edu

ABSTRACT
The main objective of this paper is to describe how Dempster-Shafer's (DS) theory of belief functions fits in the framework of valuation-based systems (VBS). Since VBS serve as a framework for managing uncertainty in expert systems, this facilitates the use of DS belief-function theory in expert systems.

Keywords: Dempster-Shafer's theory of belief functions, valuation-based systems, expert systems

Granular Computing—Computing with Uncertain, Imprecise and Partially True Data

Lotfi A. Zadeh

Computer Science Division
Department of EECS
UC Berkeley

ISSDQ'07
Enschede, The Netherlands
June13, 2007

(IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 2, No.9, 2013

The Paradox of the Fuzzy Disambiguation in the Information Retrieval

Anna Bryniarska
Department of Electrical and Computer Engineering
Opole University of Technology
Opole, Poland
Email: a.bryniarska@doktorant.po.edu.pl

Problem

- In the Dempster-Shafer's theory of evidence, for incorporating uncertainty, the valuation assigns to the data tables the degrees of belief for these data.
- Is there a valuation-based system (VBS) in which combination and marginalization operate on valuations and has this system properties analogical to the t-norm system?

The aim of research

- For the t-norm system of the valuation for the specific database attributes configuration (analogously to [*]) describe the algebra in which can be interpreted the Information Retrieval Logic [*].

[*] Bryniarska A., The Paradox of the Fuzzy Disambiguation in the Information Retrieval. (IJARAI) International Journal of Advanced Research in Artificial Intelligence, s. 55–58, vol. 2, No. 9, September 2013.

Valuations for the database attributes configuration

ES=<***U***, ***D***, ***A***> is called the **evidence system** for objects of the universe *U*.

- ***U*** is a finite set of all possible data indicating the objects of the universe *U*.
- ***D*** is a finite data set of all data used to represent knowledge about the object from the universe *U*.
- ***A***= $\{a_1, a_2, \ldots, a_n\}$ is a set of all used and numerated functions a:$\mathbf{U}\rightarrow\mathbf{D}$ *called* ***attributes***. The set $D_i \subseteq \mathbf{D}$, with index *i*, is a set of the attribute values a_i.
- ***The configuration of attributes (configuration)*** is any subset of attributes $s = \{a_{s1}, a_{s2}, \ldots, a_{sk}\}$, where $s_1 < s_2 < \ldots < s_k$.
- ***The frame of data configuration*** *for s* is the set $D_s = D_{s1} \times D_{s2} \times \ldots \times D_{sk}$.
- **The table** is any set $T \subseteq D_s$ for configuration *s*,
- **The table elements** of *T*, are the *data configurations* of this table.
- The set of tables is the ***database***.

Valuations for the database attributes configuration

- Analogously to the Dempster-Shafer theory of evidence, it is assumed that the degrees of belief for possibility of occurrence the data representation s *in the table* $T \in 2^{D_s}$ are defined by the functions $m_s: 2^{D_s} \to [0, 1]$. These functions are called the ***valuations*** for *s* (also called a mass or a basic probability assignment BPA function). They satisfied the conditions:

  a) $m_s(\varnothing) = 0$

  b) $\sum\{m_s(X): X \subseteq D_s\} = 1$

  when exists the table $X \subseteq D_s$ *such that* $m_s(X) \neq 0$.

- **The empty valuation** $m_\varnothing$, is for the configuration $s=\varnothing$.
- $\mathbf{V}_s$ is the **set of all valuations** m_s for $s \subseteq A$ and $\mathbf{V}_i$ is for $s=\{a_i\}$.

Marginalization

(Shenoy, 1994)

- The operation ${}^{\downarrow h}: D_s \to D_{s\setminus h}$, is called the **marginalization of the data** for configuration *s* to the configuration for *s\h, where h, s* $\subseteq$ ***A*** and $s\setminus h \neq \varnothing$, if: for all data configurations $\mathbf{d}_s \in D_s$,

$$\mathbf{d}_s^{\downarrow h} = \mathbf{d}_{s\setminus h} \in D_{s\setminus h},$$

is a subsequence *of* $\mathbf{d}_s$ with values for attributes from the configuration $s\setminus h$.

- ***The table marginalization*** *$T \subseteq D_s$, is described by:* $T^{\downarrow h} = \{\mathbf{d}: \mathbf{d} = \mathbf{d}_s^{\downarrow h}, \mathbf{d}_s \in D_s\}$.
- Using the data and table marginalization, can be described the ***marginalization of valuation***:

$$\text{for } m_s \in \mathbf{V}_s,\ m_s^{\downarrow h} \in \mathbf{V}_{s\setminus h} \text{ and for any tables } T_{s\setminus h} \subseteq D_{s\setminus h},$$

$$m_s^{\downarrow h}(T_{s\setminus h}) = \Sigma\{m_s(T_s): T_s \subseteq D_s \text{ such that } T_s^{\downarrow h} = T_{s\setminus h}\}.$$

- ***Theorem1.*** *The marginalization of valuation is a valuation.*
- ***The zero valuation:*** $0_s \in \mathbf{0}$ iff $\{X \subseteq D_s: 0_s(X) > 0\} = \varnothing$
- ***The neutral valuation:*** $1_s \in \mathbf{1}$ iff $\{X \subseteq D_s: 0_s(X) > 0\} = \{D_s\}$
- ***Statement 1.*** *For the sets h, s and attributes* $s\setminus h \neq \varnothing$, $(0_s)^{\downarrow h} = 0_{s\setminus h}$, $(1_s)^{\downarrow h} = 1_{s\setminus h}$.

Combination of valuations
(Dempster, 1967)

The operation $\bullet: \mathbf{V} \times \mathbf{V} \to \mathbf{V}$ called the **combination**, is defined as follow:

- For any configuration s, h, any valuations $m_1 \in \mathbf{V}_s$, $m_2 \in \mathbf{V}_h$ and any table $T \in D_{s \cup h}$,

  if $T=\varnothing$, then $\quad m_1 \bullet m_2(T) = 0,$

  if $T \neq \varnothing$, then $\quad m_1 \bullet m_2(T) = [m_1,m_2] \Sigma\{m_1(T_3)m_2(T_4)$: exist such tables $T_1, T_2 \in D_{s \cup h}$, that $T_1 \cap T_2 = T$, $T \neq \varnothing$, $T_3 = T_1^{\downarrow(h \setminus s)}$, $T_4 = T_2^{\downarrow(s \setminus h)}\}$,

- where the **normalization factor** is:

  $[m_1,m_2] = (K(m_1,m_2))^{-1}$, when $K(m_1,m_2) \neq 0$

  $[m_1,m_2] = 0$, when $K(m_1,m_2) = 0$,

  $$K(m_1,m_2) = \Sigma\{m_1(T_3)m_2(T_4): \text{exist such tables } T_1, T_2 \in D_{s \cup h}, \text{ that } T_1 \cap T_2 \neq \varnothing, T_3 = T_1^{\downarrow(h \setminus s)}, T_4 = T_2^{\downarrow(s \setminus h)}\}.$$

  If for any $T_1, T_2 \in D_{s \cup h}$, such that $T_1 \cap T_2 \neq \varnothing$, $m_1(T_1^{\downarrow(h \setminus s)}) = 0$ or $m_2(T_2^{\downarrow(s \setminus h)}) = 0$, then $K=0$. Then, $m_1 \bullet m_2 = 0_{s \cup h}$.

- ***Theorem 2.*** *The combination of valuation is a valuation.*

Attribute valuations network (AVN) (Shenoy, 1994)

- In the evidence system ES=<$\boldsymbol{U}$, $\boldsymbol{D}$, $\boldsymbol{A}$>, the **attribute valuation network** is defined as:

$$\textbf{AVN} = < \boldsymbol{V}_{AVN}, \boldsymbol{V}_{\boldsymbol{gen}}, \bullet, \downarrow, \boldsymbol{1}, \boldsymbol{0}>,$$

where $\boldsymbol{V}_{AVN} \subseteq \boldsymbol{V}$ is a set of all valuations generated from a set of **generators** $\boldsymbol{V}_{\boldsymbol{gen}} = \{\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots, \sigma_k\}$, using operations $\bullet, \downarrow, \sigma_i \notin \boldsymbol{0}, \sigma_i \notin \boldsymbol{1}, \sigma_i \in \boldsymbol{V}, i=1,2,\ldots,k.$

- **Notation**

$$m^1 = m,$$

$$m^{k+1} = m^k \bullet m, \text{ for } k>0$$

For the valuation sequence $M=\{m_1, m_2, \ldots, m_j\} \subseteq \boldsymbol{V}$:

$$\bullet M = \bullet\{m_1, m_2, \ldots, m_j\} = m_1 \bullet m_2 \bullet \ldots \bullet m_j$$

Atomic valuations Decompositon of the valuation

- ***The atomic valuation*** are called the generators, the neutral valuations, the zero valuations and the valuations which cannot be marginalized. The set of all atomic valuations is described as ***Atom***

Statement 2.

a) $\mathbf{0} \subseteq \mathbf{Atom}$ *and* $\mathbf{1} \subseteq \mathbf{Atom}$

b) $V_{gen} \subseteq \mathbf{Atom}$

c) $V_i \cap V_{AVN} \subseteq \mathbf{Atom}$, *i=1,2,…n, (valuation for singleton configuration).*

Theorem 3.

Any valuations *m* is the atomic valuations or can be decomposed into some atomic valuations. There exist a finite sequence of the atomic valuations *m(M) with values from the set*:

$$M=\{m_1, m_2, \ldots, m_j\} \subseteq \mathbf{Atom}, \text{ such that } m=\bullet M$$

The finite sequence of atomic valuations *m(M)* from the set *M* is called the ***decompositon of the valuation*** *m,* if *m=•M*.

Relation of part order in AVN

- $m^{\downarrow}(M)$ is a subsequence of the decomposiotion $m(M)$, such that its values are all elements of the set $M \backslash (\boldsymbol{V}_{gen} \cup \mathbf{1})$.
- $m^{\uparrow}(M)$ is a subsequence of the $m(M)$, such that its values are all elements of the set $M \cap (\boldsymbol{V}_{gen} \cup \mathbf{1})$.
- The pair $< m^{\downarrow}(M), m^{\uparrow}(M) >$ is called the **canonical decomposition of the valuation** m with values from the set M.
- For any valuations $m_1, m_2, \in \boldsymbol{V}_{AVN}$,

$$m_1 \prec m_2 \text{ iff}$$

There exist such canonical decompositions

$$< m_1^{\downarrow}(M), m_1^{\uparrow}(M)>, < m_2^{\downarrow}(M), m_2^{\uparrow}(M)>$$

of the valuations $m_1, m_2,$ that

$m_2^{\downarrow}(M)$ is a subsequence of $m_1^{\downarrow}(M)$ and

$m_1^{\uparrow}(M)$ is a subsequence of $m_2^{\uparrow}(M)$.

Theorem 4. The relation $\prec$ is a **relation of part order** between the valuations.

T-norm valuation system

- The system ***t-VN*** = < $\boldsymbol{V}_{AVN}$, $\boldsymbol{V}_{gen}$, •, $\preceq$, ***1***, ***0***>, is called the ***t-norm valuation system*** for the configuration of the database attributes. *The combination operation in the t-norm system is called the **t-norm valuation**.*

Theorem 5.

*In the system **t-VN**, for any valuation m_1, m_2, $m_3 \in$ **V** , for some attribute configurations s:*

- $m_1 \bullet 1_s = m_1$ and $m_1 \bullet 0_s = 0_s$, (boundary conditions),
- $m_1 \bullet m_2 = m_2 \bullet m_1$ (commutativity),
- $(m_1 \bullet m_2) \bullet m_3 = m_1 \bullet (m_2 \bullet m_3)$ (associativity),
- *if* $m_1 \preceq m_2$, *then* $(m_1 \bullet m_3) \preceq (m_2 \bullet m_3)$ (monotonicity of valuation combination),

Logic Algebra *LA* =< V_{AVN}, $\otimes$, $\oplus$, $\rightarrow$, $\leftrightarrow$, `, **0**, **1**> (Hájek, 1998)

Implication operation

$$m_3 \{ (m_1 \rightarrow m_2) \quad \text{iff } m_1 \bullet m_3 \{ m_2$$

Conjunction operation

$$m_1 \otimes m_2 = m_1 \bullet (m_1 \rightarrow m_2)$$

Alternative operation

$$m_1 \oplus y = ((m_1 \rightarrow m_2) \rightarrow m_2) \otimes ((m_2 \rightarrow m_1) \rightarrow m_1)$$

Negation operation

$$m_1' = (m_1 \rightarrow 0_s), \text{ when } m_1 \in \mathbf{V_s},$$

Equivalence operation

$$m_1 \leftrightarrow m_2 = (m_1 \rightarrow m_2) \otimes (m_2 \rightarrow m_1)$$

Conclusion

In the t-norm valuation system for the configuration of database attributes, analogously to [3], can be described the algebra **LA** =< V_{AVN}, $\otimes$, $\oplus$, $\rightarrow$, ***0***, ***1***>, in which the Information Retrieval Logic can be interpreted.

References

[1] Dempster, A. P. Upper and lower probabilities induced by a multivalued mapping, in: *Annals of Mathematical Statistics*, 38, 1967, 325-339.

[2] Shenoy P. P., Using Dempster-Shafer's belief-function theory in expert systems, in: Yager, R. R., M. Federizzi, and J. Kacprzyk, eds., Advances in the Dempster-Shafer Theory of Evidence, John Wiley & Sons, New York, NY, 1994, pp. 395–414.

[3] Bryniarska A., The Paradox of the Fuzzy Disambiguation in the Information Retrieval. (IJARAI) International Journal of Advanced Research in Artificial Intelligence, s. 55–58, vol. 2, No. 9, September 2013.

[4] Hájek P., Metamathematics of Fuzzy Logic. Volume 4 of Trends in Logic, Studia Logica Library, Kluwer Academic Publishers, Dordrecht 1998.

Thank you for your attention