Kybernetika

Vladimír Mařík; Hana Krautwurmová

An extension of the pseudobayesian model by making use of Dempster-Shafer's theory

Kybernetika, Vol. 25 (1989), No. Suppl, 59--64

Persistent URL: http://dml.cz/dmlcz/125415

AN EXTENSION OF THE PSEUDOBAYESIAN MODEL BY MAKING USE OF DEMPSTER-SHAFER'S THEORY

VLADIMÍR MAŘÍK, HANA KRAUTWURMOVÁ

Dempster-Shafer's theory of evidence is appealing because it does suggest a coherent approach for dealing with uncertain knowledge used in expert systems. We present an adaptation of the Prospector pseudobayesian model for inexact reasoning with relevance to the D-S theory and to the algebraic theory developed by P. Hájek and J. J. Valdés.

1. INTRODUCTION

The pseudobayesian model for uncertainty handling previously used by Prospector [1] is very widely used in diagnostic expert systems. In such systems the knowledge representation is based on uncertain production rules ($E \rightarrow H$ rules) having the following form:

$$\text{IF } \langle \text{evidence } E \rangle \text{ THEN } \langle \text{hypothesis } H \rangle \text{ WITH } \langle \text{probability } P(H \mid E) \rangle$$
$$\text{ELSE } \langle \text{hypothesis } H \rangle \text{ WITH } \langle \text{probability } P(H \mid -E) \rangle$$

where ⟨evidence E⟩ and ⟨hypothesis H⟩ are propositions, ⟨probability $P(H \mid E)$⟩ and ⟨probability $P(H \mid -E)$⟩ are subjective uncertainty measures (not probabilities in an exact mathematical meaning!). Their values are given by the expert. The pseudobaesian model for uncertainty handling requires prior probability to be assigned to each proposition by the expert.

In the pseudobayesian (Prospector) model of uncertainty handling the "classical" Bayesian formulae form the theoretical background for

a) computing the posterior probabilities $P(H \mid E')$ making use of a piece of evidence $P(E \mid E')$ where E' denotes the user's observation,

b) combining the influence of several rules with the same hypothesis on the right-hand side (the independence of all pieces of evidence being presumed).

The pseudobayesian model has been described in detail for instance in [1], [2] and [3]. Let's only underline that to ensure the model consistency (without knowing the complete probabilistic distribution) a heuristic interpolation function $P(H \mid E') = f(P(E/E'))$ has to be established. Three points of this function are fixed by the expert, namely $[0, P(H \mid -E)]$, $[P(E), P(H)]$, $[1, P(H \mid E)]$. The most common and natural interpolation function f is the piecewise-linear one (see Fig. 1). Such interpolation function plays the same role as the CTR function in Hájek's sense of algebraic theory.

The uncertainty (no matter what is the type or source of it) of each proposition is involved in just one parameter — the posterior probability.

Dempster-Shafer's (D-S) theory. The Dempster theory [4] may be considered as a generalization of the Bayesian probability theory. This theory has been extended by Shafer in [5]. Later on it has been thoroughly investigated by P. Hájek, the aspects of possible applications in the area of rule-based diagnostic expert systems being

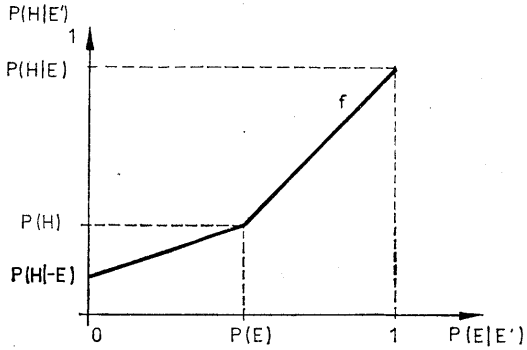

Fig. 1.

taken into account. P. Hájek and J. Valdés [6] discovered important properties of algebraic structures based on D-S theory and they have distinguished *certainty* and *vagueness* as two quite different independent aspects of uncertain information.

A combined approach. The goal of our approach is to properly extend the pseudo-bayesian model by making use of the D-S theory. In a modified model for uncertainty handling each production rule $E \rightarrow H$ is considered as a Dempster source. The pseudobayesion philosophy is preserved but the uncertainty has a two-dimensional character and may be expressed as a Dempster pair $(a, b) = (P(N), P(-N))$, where $P(N)$ is the probability that the given proposition N holds, $P(-N)$ is the probability that the given proposition N does not hold, $1 - P(N) - P(-N)$ remains unassi gned. Of course, $P(N) + P(-N) \leqq 1$. A Dempster pair may be easily converted into the pair (certainty, vagueness), by applying Hájek-Valdés's h and g mapping (see [6]).

2. EXTENSION OF THE PSEUDOBAYESIAN MODEL

2.1. $E \rightarrow H$ rule and Dempster source

A Dempster space (X, T, μ, Γ) may be defined like this: X and T are sets containing these elements $X = \{E, -E, U\}$, (U ... "unknown"), $T = \{H, -H\}$, μ is a probability measure on X that holds:

$$\mu(E) = P(E), \quad \mu(-E) = P(-E),$$
$$\mu(U) = 1 - P(E) - P(-E)$$

and Γ is a multivalued mapping $X \rightarrow T$.

Let's define a (prior) Dempster information source Z:

$$p_H = \mu(x \in X,\ \Gamma(x) = \{\ H\}) = P(H \mid E) . P(E) + P(H \mid -E) . P(-E)$$
$$p_{-H} = \mu(x \in X,\ \Gamma(x) = \{-H\}) = P(-H \mid E) . P(E) + P(-H \mid -E) . P(-E)$$
$$p_0 = 0$$
$$p_{H,-H} = 1 - p_H - p_{-H} .$$

The set $T = \{H, -H\}$ contains only two elements, that's why it is possible to write:

$$a = P_*(\{H\}) = p_H ,\quad b = P_*(\{-H\}) = p_{-H} ,$$
$$P^*(\{H\}) = 1 - P_*(\{-H\}) = 1 - b ,$$
$$P^*(\{-H\}) = 1 - P_*(\{H\}) = 1 - a .$$

To express the prior Dempster source Z the following probabilities must be given by the expert: $P(E)$, $P(-E)$, $P(H \mid E)$, $P(H \mid -E)$, $P(-H \mid E)$, $P(-H \mid -E)$. That's why a production rule $E \rightarrow H$ has to be interpreted as a fourtuple-rule in the following way:

IF ⟨evidence E⟩ THEN ⟨hypothesis H⟩ WITH ⟨probability $P(H \mid E)$⟩
IF ⟨evidence $-E$⟩ THEN ⟨hypothesis H⟩ WITH ⟨probability $P(H \mid -E)$⟩
IF ⟨evidence E⟩ THEN ⟨hypothesis $-H$⟩ WITH ⟨probability $P(-H \mid E)$⟩
IF ⟨evidence $-E$⟩ THEN ⟨hypothesis $-H$⟩ WITH ⟨probability $P(-H \mid -E)$⟩

where $P(-H \mid E) + P(H \mid E) \leqq 1$, $P(H \mid -E) + P(-H \mid -E) \leqq 1$. (Let's remark that – in contradiction with the new, combined model – in the "clasical" pseudobayesian model the dependencies $P(-H \mid E) = 1 - P(H \mid E)$ and $P(-H \mid -E) = 1 - P(H \mid -E)$ are assumed.)

The Dempster source Z is based on prior probabilities. With respect to the user's information $P(E \mid E')$ and $P(-E \mid E')$ the prior Dempster source has to be changed – it is converted into the posterior Dempster source Z'.

Analogically to the original pseudobayesian model, the consistency of the combined one is ensured by interpolation functions. (As a matter of fact, there are two interpolation functions in the combined model: $P(H \mid E') = f_{\text{I}}(P(E \mid E'))$, $P(-H \mid E') = f_{\text{II}}(P(-E/E'))$.) Both interpolation functions have to satisfy some conditions to ensure:

a) the pair $(P(H \mid E'), P(-H \mid E'))$ to be a Dempster one for $P(E \mid E') \in \langle 0, 1 \rangle$, $P(-E, E') \in \langle 0, 1 \rangle$ and $P(E \mid E') + P(-E \mid E') \leqq 1$,

b) their monotonicity.

Let's consider without any lost of generality the interpolation function f_{I}, f_{II} in the form shown in Fig. 2. (The case when one of them is decreasing and the other increasing has no practical significance, the case of both decreasing functions may be easily converted in that of increasing ones by negation of E.)

The conditions which should f_{I}, f_{II} satisfy are expressed in Propositions 1–3 without proofs.

Proposition 1. Let

a) $Z = (a, b)$ be a prior Dempster source.

b) it hold

$$P(H \mid -E) \leqq P(H) \quad \leqq P(H \mid E)$$
$$P(-H \mid E) \leqq P(-H) \leqq P(-H \mid -E)$$

c) $(P(E \mid E'), P(-E \mid E'))$ be a Dempster pair, then by the choice

$$\hat{P}(H) = P(H) + k_1(1 - P(H) - P(-H)), \quad k_1 \in \langle 0, 1 \rangle,$$
$$\hat{P}(-H) = P(-H) + k_2(1 - P(H) - P(-H)), \quad k_2 \in \langle 0, 1 \rangle,$$

the pair $(P(H \mid E'), P(-H \mid E')) = (a', b')$ is a Dempster pair (Dempster source Z').

Properties of the interpolation described in Proposition 1:

a) By the choice $k_1 = k_2 = 0$ the functions $f_{\mathrm{I}}, f_{\mathrm{II}}$ have a "zone of unsensitivness": a value $P(E \mid E') \in \langle P(E), 1 - P(-E) \rangle$ has no influence on $a' = P(H \mid E')$, analogically $P(-E \mid E') \in \langle P(-E), 1 - P(E) \rangle$ has no influence on $b' = P(-H \mid E')$.

b) By the choice $k_1 = k_2 = 1$ the non-vague user's answer $(P(E \mid E') + P(-E \mid E') = 1)$ within the interval $P(E \mid E') \in \langle P(E), 1 - P(-E)) \rangle$, $P(-E \mid E') \in \langle P(-E), 1 - P(E) \rangle$ results in a non-vague posterior Dempster source Z' $(a' + b' = 1)$. The smaller values of k_1, k_2, the greater vagueness of Z' appears by the user's answer within intervals mentioned above.

Proposition 2. Let the Dempster source Z be a non vague source $(a + b = 1)$ and the user's observation be a non-vague, one, too. Then the interpolation function f_{I} is identical to the piecewise-linear interpolation function in the original pseudo-bayesian model.

There is another condition concerning the functions $f_{\mathrm{I}}, f_{\mathrm{II}}$ which has to be satisfied: these functions should be non decreasing to preserve the "natural behaviour" of the functions.

Proposition 3. Let the conditions a), b) and c) of Proposition 1 be satisfied. Let

$$P(H \mid E) \geqq 1 - P(-H) \text{ or } P(-H \mid -E) \geqq 1 - P(H) \text{ respectively)}$$

or let $k_1, k_2 \in \langle 0, 1 \rangle$ be chosen in the following way:

$$k_1 \leqq \frac{P(H \mid E) - P(H)}{1 - P(H) - P(-H)} \quad \left(\text{or } k_2 \leqq \frac{P(-H \mid -E) - P(-H)}{1 - P(H) - P(-H)} \text{ respectively}\right).$$

Then the functions $f_{\mathrm{I}}, f_{\mathrm{II}}$ are non-decreasing within the interval $\langle 0, 1 \rangle$.

2.2. Combination of rules

Each rule $E \rightarrow H$ is considered as a Dempster source. If there are m rules with the same H on the right-hand side, $E_i \rightarrow H$, $i = 1, 2, \ldots, m$, the Dempster combining rule (cf. [7] – it accomplishes the $\oplus$ operation in Hájek's sense) is used to compute the total value $P(H \mid E'_1, E'_2, \ldots, E'_m)$. The algebraic structure with the $\oplus$ operation defined over the set of all Dempster pairs forms a *standard Dempster combining structure* defined in [6]. That's why it is possible to make use of results of [6] directly. The following formula may be used to combine the rules:

$$A = A_0 \oplus (Z'_1 \oplus -A_0) + (Z'_2 \oplus -A_0) \oplus \ldots \oplus (Z'_m \oplus -A_0)$$

where

A_0 is a prior pair $(P(H), P(-H))$,

A is a posterior pair $(P(H \mid E'_1, E'_2, \ldots, E'_m), P(-H \mid E'_1, E'_2, \ldots, E'_m))$

Z'_i is a posterior Dempster source connected with the rule $E_i \to H$, $i = 1, 2, \ldots, m$.

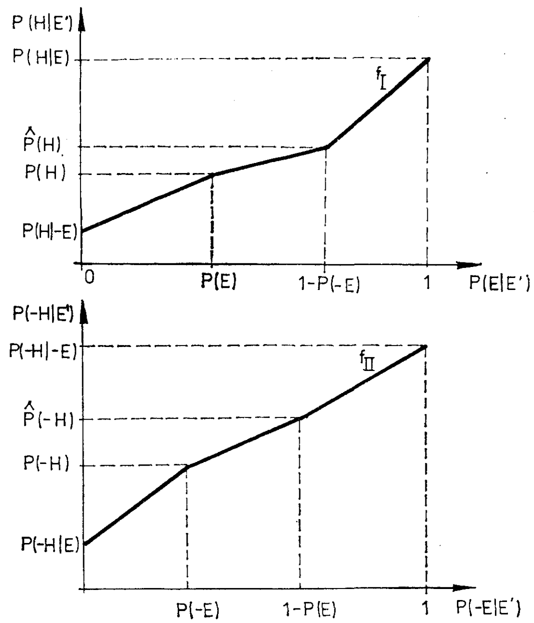

Fig. 2.

2.3. Logical functions and context links

To evaluate logical combinations of Dempster sources we employ the following formulae (cf. [8]) which have appeared as a natural extension of the Zadeh's formulae used in the original Prospector model:

$$(a_1, b_1) \wedge (a_2, b_2) = (\min(a_1, a_2), \max(b_1, b_2)),$$
$$(a_1, b_1) \vee (a_2, b_2) = (\max(a_1, a_2), \min(b_1, b_2)),$$
$$-(a, b) = (b, a).$$

A context link with parameters (α_1, α_2) is satisfied iff it holds for the context $A = (P(H \mid E'), P(-H \mid E'))$ (cf. [8]):

$$\alpha_1 \leqq P(H \mid E'), \quad \alpha_2 \geqq 1 - P(-H \mid E').$$

3. IMPLEMENTATIONS

The FEL-EXPERT family [9] of expert systems is based on the pseudobayesian model for uncertainty handling. The proposed combined model has been implemented in the FEL-EXPERT version 2.95 (see [8]). It has been extended to enable the processing of quantitative information (Q-nodes and S-nodes in the FEL-EXPERT notation (cf. [9]).

REFERENCES

[1] R. O. Duda, P. E. Hart and N. J. Nilsson: Subjective Bayesian Methods for Rule-Based Inference Systems. TN 124, SRI International, Stanford 1976.

[2] A. Patterson: AL/*X* — User's Manual. Intelligent Terminals Ltd., Oxford 1981.

[3] V. Mařík, Z. Zdráhal and A. Raymundo Mender: A Skeleton for Expert Systems: MEXEXP 1.0. Technical Report AHR-84-29, C.I.E.A. del I.P.N., México City, 1984, 31pp.

[4] A. P. Dempster: Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Statist. *38* (1967), 325—339.

[5] G. Shafer: A Mathematical Theory of Evidence. Princeton Univ. Press, Princeton N.J. 1976.

[6] P. Hájek and J. J. Valdés: Algebraic Foundations of Uncertainty Processing in Rule-Based Expert Systems. Report No. 28, MÚ ČSAV, Praha 1987.

[7] M. Ginsberg: Non-Monotonic Reasoning Using Dempster's Rule. HPP-84-30, Stanford Univ., Stanford 1984.

[8] H. Krautwurmová: Rozšíření modelu pro práci s neurčitými znalostmi u expertního systému FEL-EXPERT. Ph. D. Thesis, Czech Technical University, Prague 1988.

[9] V. Mařík and Z. Zdráhal: The FEL-EXPERT and MIFELEX Empty Expert Systems and Their Applications. Working Paper WP-86-61, IIASA, Laxenburg, Austria, 1986, pp. 46—56.

Doc. Ing. Vladimír Mařík, CSc., Ing. Hana Krautwurmová, elektrotechnická fakulta ČVUT (Faculty of Electrical Engineering — Czech Technical University), Suchbátarova 2, 166 27 Praha 6. Czechoslovakia.