POSSIBILITY-BASED ADAPTATION OF CASUISTIC REASONING MODELS

Rolanda PREDESCU
National Commission of Informatics
Piata Victoriei 1, PO Box 54-048
71201 Bucharest ROMANIA
Fax: (+40)(1) - 3121219

Abstract

One of the major difficulties in case-based reasoning systems, when trying to classify and solve new cases, is the selection of similar, already known cases, mostly appropriate to offer adequate solutions.

The paper introduces practical methods for the evaluation and comparison of the relevance of known cases.

Estimations are made concerning the power and coherence of the explanations they may offer for the newly encountered cases.

Conditional possibility and plausability measures are used as estimating tools, in order to provide more accurate selection methods, due to a quite consistent meaning of the concepts of explanation and coherence.

Key words: abduction, case/explanation-based reasoning, coherence, possibility theory, conditional possibility, belief functions, conditional plausability

1. Introduction

Case-based reasoning systems /1,2,7/ are usually substituting the modus ponens type inference model, by a specific interpretation of the abductive type inference model:

$$\frac{\begin{array}{r} A \longrightarrow B \\ B \end{array}}{\text{possibly } A}$$

In the special case of explanation oriented reasoning /8, 10/, the concepts of "implication" and "consistency" are generally replaced by "explanation" and "coherence" :

A explains B
B

B is coherent with [explanation] A

The possibility to find several explanations for the same fact implicitly triggers the necessity to attach various coherence degrees to it, according to each alternative explanation. A is considered an "explanatory support" for B, which, in its turn, inherits an "explanatory power", due to its coherence with A.

Generally, case-based reasonong models do correspond to a diagnostic-like sheme, of abductive type:

diagnostic explains features
features
----
features are coherent with diagnostic

The diagnostic is used as an explanatory support for the features and reflects how the case is classified. The features are usually divided into two categories: relevant and supplimentary ones.

New cases (simply, new sets of features) are classified by their ressemblance to already known cases, namely by:
- the matching degree between relevant features;
- the coherence degree between the features of the new case and tthe explanatory support of the known cases.

A generalized prototype model is produced on the basis of known cases, which is then adjusted, according to the restrictions propagated by the features of the new case.

Finally, the adjusted model, containing a diagnostic and the adequate treatment, is stored into the memory of cases, as a hew, complete case structure.

2. The possibilistic approach to explanation-based reasoning

a) The intention of the present paper is to introduce a more rigurous definition for "explanation" and "coherence", using the specific concepts of conditional possibility and plausability /5,6/. The aim is to offer better means for the comparison of known cases relevance during solving processes.

Thus, we shall define "A explains B", by the possibility of B to occur, when the condition A is present, namely by the conditional possibility measure Pos $(B|A)$ /6/.

According to the theory of possibility /11/, if U is a universe of possible values for a variable X,then $\text{posX}: U \longrightarrow [0,1]$ represents a distribution of possibilities over U, namely, for each element $u \in U$, posX(u) is the possibility that X takes the value u. For a subset $A \subset U$, posX reflects the possibility that $X \in A$ and induces a possibility measure attached to A:

$$\mathrm{Pos}(A) = \max_{u \in A} \mathrm{pos}X(u)$$

A conditional possibility distribution $\mathrm{pos}Y|X : U\times V \to [0,1]$ reflects how the values taken by Y within $B \subseteq V$ directly depend on the values taken by X within $A \subseteq U$, and is defined by:

$$\mathrm{pos}Y|X(v, u) = \min(\mathrm{pos}X(u), \mathrm{pos}Y(v))$$

Hence, in the sentence "A explains B", the explanatory power of B due to A, that we shall denote by EXPL(B|A), is measured by

$$\mathrm{Pos}(BA) = \max_{u,v \in A\times B} \mathrm{pos}Y|X(v, u)$$

b) As well, we sustain the definition of "B is coherent with A", by the plausability of B, in the presence of the explanation A, namely by the conditional plausability Pl(B|A) /5, 6/.

According to the Dempster-Shafer theory of belief functions /3, 9/, if E is a set of evidences, with an associated probability distribution $p: E \to [0,1]$ and H is a set of hypotheses, depending on E, then p induces a basic probability(mass) distribution $m: \mathcal{P}(H) \to [0,1]$, with $m(\emptyset)=0$ and $\sum_{A \subseteq H} m(A) = 1$.

Consequently, a set of "focal" elements $A \subset H$ can be derived, for positive masses: $m(A) > 0$. The plausability of an subset B of H is given by:

$$\mathrm{Pl}(B) = \sum_{A \cap B \neq \emptyset} m(A)$$

when B is a crisp set, and by:

$$\mathrm{Pl}(B) = \sum_{A} m(A) \times \mathrm{Pos}(B|A)$$

when B is imprecise (the sum of masses of focal elements that make B possible).

Dubois and Prade /4/ offer an extension of this definition:

$$\mathrm{Pl}(B) = \sum_{A} m(A) \times \mathrm{Pl}(B|A)$$

with $\mathrm{Pl}(B|A) = \mathrm{Pl}(A \cap B)/\mathrm{Pl}(A)$.

Hence, in the sentence "B is coherent with A", we consider the degree of coherence between B and A as being expressed by Pl(B|A) and we shall denote it by COHE(B|A).

The new mass of B will be:

$$m1(B|A) = \left[\sum_{C \cap A = B} m(C)\right] / \mathrm{Pl}(A)$$

c) A more special case is represented by explicative chains of the form:

A1 explains A2 .... explains An explains B

and coherence chains:

B is coherent to An .... is coherent to A1

We shall denote the final explanatory power and coherence of B by EXPL(B | (Ai)i) and COHE(B|(Ai)i), respectively.

The problem is, of course, how the two parametres are propagated through the chain.

In the first case, we consider an extension of the conditional possibility, linking the values of the variable Y to the values taken by several variables X1,..,Xn , known to be non-interactive.

Thus, the final explanatory power propagated towards B, will be given by:

$$\mathrm{EXPL}(B|(Ai)i) = \max_{u1,..,un,v \in A1x..xAnxB} \mathrm{posY}|X1,..,Xn(v,u1,..,un)$$

where :

$$\mathrm{posY}|X1,..,Xn(v,u1,..,un) = \min(\mathrm{posX1}(u1),.., \mathrm{posXn}(un))$$

In the second case, for the coherence degree, we build an extension of the conditional plausability, reflecting its iterative updating (re-conditioning) through the chain:

$$\mathrm{COHE}(B|(Ai)i) = 1/\mathrm{Pl}(A1 \cap .. \cap An) \times \left[\mathrm{Sum}_i\, \mathrm{Pl1}(Ai|Ai-1) + \mathrm{Pl1}(B|An)\right]$$

where $\mathrm{Pl1}(B2|B1) = \mathrm{m1}(B2|B1) \times \mathrm{Pl}(B2|B1)$.

3. The power and coherence of explanations

We shall simply call explanations of a concept B, both the "singleton" explanatory supports "A explains B", as well as the explicative chains supports. Let EB = {E1,..,En} be a set of such possible explanations.

We shall define an ordering relation on EB. Thus, we shall assume that "Ei explains B better than Ej" iff :

EXPL(B|(Aik)k) > EXPL(B|(Ajl)l)

with k,l >= 1.

In a similar manner, if we consider CB = {C1,..,Cn} to be a set of coherence chains attached to B, we may define an ordering relation on CB. We assume that "B is more coherent with Ci than with Cj", iff:

COHE(B|(Aik)k) > COHE(B|(Ajl)l)

with k,l >= 1.

4. Conclusions

The paper introduces practical methods for the evaluation and comparison of case relevance, in casuistic reasoning models. The differentiation criteria are the power and coherence of the explanations offered by each alternative case.

The use of the possibility theory framework allows for the development of more accurate and reliable tools, dedicated to the refinement and up-grading of solutions in case-based reasoning systems.

The application areas are significantly wide and cover case analysis requirements in both technical and extra-technical fields.

References

1. Aamodt A. - A computational model of knowledge-intensive learning and problem solving, in Current Trends in Knowledge Aquisition, Wielinga, Boose, Gaines, Schreiber, Van Sommeren, eds., IOS Publishers, 1990

2. Bareiss E.R., Porter R., Wier C.C. - PROTOS: An exemplar-based learning apprentice, 1987

3. Dempster A.P. - Upper and lower probabilities induced by a multi-valued mapping, Ann. Math. Stat., 38, 325-339, 1967

4. Dubois D., Prade H. - On the unicity of Dempster rule of combination, Int. J. of Intelligent Systems, 1, 133-142,1986

5. Dubois D., Prade H. - Measuring and updating information in Conditioning, updating, non-standard uncertainty models and non-monotonic logic, Rep. IRIT/90-52/R, IRIT, Toulouse, 1990

6. Dubois D., Prade H. - Updating with belief functions, ordinal conditional functions and possibility measures, in Conditioning, updating, non-standard uncertainty models and non-monotonic logic, Rep. IRIT/90-52/R, IRIT, Toulouse, 1990

7. Kolodner J. - Case-based reasoning, in Proc. IJCAI, 1985

8. Kaas A. - Adaptation-based explanations, in Proc. IJCAI, 1989

9. Shafer G. - A Mathematical Theory of Evidence, Princeton Unniv. Press., Princeton, 1976

10. Thagard P. - Explanatory Coherence, CSL Rep. 16, Princeton Univ., 1989

11. Zadeh L.A. - Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems, 1, 1, 1978