Fuzzy Similarity-based Models in Case-based Reasoning

Francesc Esteva
Institut d'Investigació en
Intel·ligència Artificial, (CSIC)
Campus UAB,
08193 Bellaterra, Spain
esteva@iiia.csic.es

Pere Garcia
Institut d'Investigació en
Intel·ligència Artificial, (CSIC)
Campus UAB,
08193 Bellaterra, Spain
pere@iiia.csic.es

Lluís Godo
Institut d'Investigació en
Intel·ligència Artificial, (CSIC)
Campus UAB,
08193 Bellaterra, Spain
godo@iiia.csic.es

Abstract

This paper deals with fuzzy similarity-based models of the basic principle of case-based reasoning (CBR) stating that *"similar problems lead (or may lead) to similar outcomes"*. A stronger form of this principle stating that *"outcome attributes are at least as similar as problem description attributes"* has been studied in some previous works. In this paper another form of the basic principle stating that *"the more similar are the problem description attributes, the more similar are the outcome attributes"* is studied. These two forms of the CBR principle are used to infer possible outcomes for the current problem from the information stored in the memory of precedent cases.

Keywords: Case-based reasoning, fuzzzy similarity relations.

1 Introduccion

Let us first specify how case-based reasoning (CBR) is viewed in this paper (for a more general overview of case-based reasoning see e.g. [1]).

A case is viewed as a n-tuple of precise attribute values, this set of attributes being divided in two non-empty disjoint subsets: the problem description attributes, and the problem solution or outcome attributes, denoted by D and C respectively. These subsets are taken according to the problem we deal with. A *case* will be denoted as a tuple (s, t) where s and t stand for complete sets of precise attribute values of D and C respectively. In order to perform case-based reasoning we assume we have a finite set M of known cases or precedents, called *case base* or *memory* (M is thus a set of pairs (s, t)), and a current problem description, denoted by s_0, for which the precise values of all attributes belonging to D are known. Then, case-based reasoning aims at extrapolating or estimating the value t_0 of the attributes in C, for the current problem. In this paper it is assumed everywhere that the memory only includes completely informed cases.

In case-based reasoning it is assumed that the attributes belonging to the outcome set can be related, in some way, to the problem description attributes. Actually, the goal of case-based reasoning is to estimate the value of the outcome attribute(s) of the current case taking into account the case base of precedents M. To do this, in the following, it is assumed that a basic *principle* stating that "similar situations (problems) give (or may give) similar outcomes (solutions)" holds. Therefore, some kind of *similarity measure* between problem descriptions and a similarity measure between outcomes are needed.

In this paper, these similarity measures are suposed to be given by two fuzzy similarity relations S and T defined on the set of problem description attributes D and on the set of outcome attributes C respectively, i.e by functions S: $D \times D \rightarrow [0, 1]$, T: $C \times C \rightarrow [0, 1]$. We suppose that these fuzzy relations are primitive notions and known in advance (although we might think of learning them from the set of precedents stored in the memory). Methods for obtaining fuzzy relations, or checking their adequacy are out of the scope of this paper.

Expressed in terms of the fuzzy relations S and T, the implicit Case-Based Reasoning Principle can be expressed by the following rule:

"Similar problems in the sense of S have similar outcomes in the sense of T",

A problem in the framework of our case-based reasoning model is then denoted by a 4-tuple (M, S, T, s_0) where M stands for a case-base or memory, S and T stand for the similarity relations and s_0 stands for the current case. The goal of case-based reasoning is to estimate the outcome t_0 corresponding to the current problem s_0.

As in [5], we will refer to *deterministic case-based problems* (M, S, T, s_0) when the above principle is applicable, since in particular when two problems are identical, they have to have the same outcome as well; otherwise, we will refer to a CBR problem as *non-deterministic*, where only a weaker form of the principle applies, allowing to conclude only on the *possibility* that the outcome attributes be similar when the problem descriptions are indeed similar.

2 A Fuzzy Set Framework for Case-Based Reasoning: Deterministic Problems

In the deterministic setting, the above mentioned CBR principle can be modelled at least by one of the following constraints on the memory of cases M:

$$S(s_1,s_2) \leq T(t_1,t_2) \qquad (C1)$$

$$\text{if } S(s_1,s_2) \leq S(s_1,s_3) \text{ then } T(t_1,t_2) \leq T(t_1,t_3) \qquad (C2)$$

$$\text{if } S(s_1,s_2) < S(s_1,s_3) \text{ then } T(t_1,t_2) < T(t_1,t_3) \qquad (C3)$$

$$\text{if } S(s_1,s_2) < S(s_1,s_3) \text{ then } T(t_1,t_2) \leq T(t_1,t_3) \qquad (C4)$$

for any (s_1,t_1), (s_2,t_2), $(s_3,t_3) \in M$. Actually the fuzzy relations S and T are in principle required to be only reflexive, i.e. to be fuzzy proximity relations (see [10] and [12]).

The first one was considered in [5] and [9] and corresponds to interpreting the above principle as

"Outcome attributes are at least as similar as problem description attributes".

The other three constraints, which are proposed and studied in this paper, reflect (with small differences) another usual interpretation of the CBR principle in a fuzzy setting:

"The more similar are the problem description attributes in the sense of S, the more similar are the outcome attributes in the sense of T"

In fact, for instance, constraint C2 can be better described in terms of the following binary relations on M: $s \in D$ and $t \in C$ define

$$(s_1, t_1)\, R_s\, (s_2, t_2) \text{ iff } S(s,s_1) \leq S(s,s_2)$$

$$(s_1, t_1)\, R_t\, (s_2, t_2) \text{ iff } T(t,t_1) \leq T(t,t_2).$$

Then, constraint C2 can be equivalently expressed by stating:

$$R_t \supseteq R_s. \qquad (C2')$$

for any $(s, t) \in M$. If we denote by R'_s and R'_t the same relations but changing $\leq$ by $<$, then constraints C3 and C4 can be analogously described by c

$$R_t \supseteq R'_s. \qquad (C3')$$

and

$$R'_t \supseteq R_s. \qquad (C4')$$

respectively.

All these relations are transitive, but obviously only R_s and R_t are reflexive

If we also define the unary relation $R_{(s,\,t)}$ in M by

$$(s_1, t_1) \in R_{(s,\,t)} \text{ iff } S(s_1,s) \leq T(t_1,t)$$

for each $(s, t) \in D\times C$, then we can also express constraint C1 by the following inclusion:

$$R_{(s,\,t)} \supseteq M. \qquad (C1')$$

It is also easy to check that some constraints are stronger than others are, for instance, constraints C2 and C3 are stronger than constraint C4 but there is no implication between C2 and C3 and, clearly, there is no implication between C1 and the rest.

Of course, the properties fulfilled by a memory of cases M satisfying one of these constraints will vary. Indeed the following particular properties can be easily proved:

1. In the cases of constraints C1 and C2 we have,

✓ If s_1, s_2, are equivalent w.r.t. S (i.e., if $S(s_1, s_2) = 1$), then t_1, t_2 are equivalent w.r.t. T (i.e. $T(t_1, t_2) = 1$).

Thus, if for a current problem s_0 there is a case (s_1, t_1) in the memory M such that s_0 completely coincides with s_1, then the solution is to be found in class of outcomes equivalent to t_1 w.r.t. T. If T satisfies the separability property ($T(t, t') = 1$ iff $t = t'$), then the only solution is t_1.

2. In the case of constraint C3 we have:

✓ If s_1, s_2, are not equivalent w.r.t. S, then t_1,t_2 are not equivalent w.r.t. T.

3. In the cases of constraints C3 and C4 we have,

✓ If s_1, s_2, are equivalent w.r.t. S, then t_1,t_2 are either equivalent w.r.t. T or, (s_2, t_2) is the precedent of (s_1, t_1) in the preorder R_{t_1} (R'_{t_1}) and t_1 is the precedent of t_2 in the preorder R_{t_2}.

Summarizing, we have the following conditions depending of the constraint satisfied by the memory:

- ✓ With either constraints C1 or C2, it is not possible for a same problem (in the sense that $S(s_1, s_2) = 1$) to have different solutions (in the sense that $T(t_1,t_2) < 1$).
- ✓ With constraint C3 it is not possible for different problems (in the sense that $S(s_1, s_2) < 1$) to have same solutions (in the sense that $T(t_1,t_2) = 1$).
- ✓ With constraint C4 no further constraint exists for the outcomes of problems which are different ($S(s_1, s_2) < 1$).

Finally let us remark constraints C1 and C2 are not equivalent as we could easily find examples of memories and similarities satisfying C1 but not C2, or conversely satisfying C2 but not C1

3 Case-Based Inference

Let us now see how the satisfaction of the different deterministic constraints considered in the previous section by a memory of cases can be used to perform case-based inference. The idea is that, given a memory of cases M satisfying a constraint Ci (i = 1, 4), and given a new problem description s_0, then t_0 is considered a plausible solution for s_0 if $M' = M \cup \{(s_0, t_0)\}$ still satisfies the constraint Ci.

The case of constraint C1 was studied in [5] and the set of possible possible solutions for a new problem s_0 is:

$$E_{s0} = \{t \in C \,/\, S(s_i,s_0) \le T(t_i,t), \text{ for all } (s_i,t_i) \in M\} = \{t \in C \,/\, R_{(s0,\,t)} \supseteq M\}.$$

For the case of memories satisfying constraint C2, the set of possible values for t_0 is

$$E_{s0} = \{t \in C \,/\, T(t_i,t) \le T(t_j,t) \text{ if } S(s_i,s_0) \le S(s_j,s_0), \text{ for all } (s_i,t_i),(s_j,t_j) \in M\} = \{t \in C \,/\, R_t \supseteq R_{s0}\}.$$

One can proceed in an analogous way for the cases of constraints C3 and C4 with the obvious small changes.

All the elements of the sets E_{s0} are in principle equally plausible solutions for the current problem s_0. In the case there are several elements in E_{s0}, to select one out of them is a prcess that requires additional domain knowledge and it is not in the scope of this short paper.

On the other hand, it can happen that E_{s0} result in an empty set. Such a situation reflects the fact that the assumed constraint Ci may be indeed a too strong requirement for the current memory of cases, for which a deterministic set of solutions is then not possible. This happens when there is no t_0 for which $M \cup \{(s_0, t_0)\}$ satisfies constraint Ci, even though M does.

As an alternative inference method for this kind of situations, one can consider a weaker version of the Case-Based Reasoning Principle stating:

"The more similar s_1 and s_2 are, the more possible is that t_1 and t_2 are similar".

The formal expression of this principle requires to clarify the intended meaning of *possible* in this meta-rule. In [5] this *non-deterministic* principle was modeled by means of a kind of fuzzy rules called possibility rules, in contrast to gradual fuzzy rules which were used to model the deterministic CBR problems satisfying constraint C1. This non-deterministic fuzzy rule approach to CBR has been further explored in [7] and [8] (see also [11] for another use of fuzzy rules in CBR).

In this paper we propose another alternative: to consider a graded satisfiability of $M \cup \{(s_0, t_0)\}$ with respect to any constraint Ci, so that the more $M \cup \{(s_0, t_0)\}$ satisfies Ci, the more plausible (suitable) is t_0 a solution for s_0. In particular if deterministic solutions exist then the plausibility of these solutions have to be 1. Namely, given a continuous t-norm $\otimes$, for each $\alpha \in [0,1]$ we define generalized versions of the previous relations $R_{(s,\,t)}$, R_t, and R'_t as follows:

$$(s_1, t_1) \in {}^{\alpha}R_{(s,\,t)} \text{ iff } S(s_1, s) \otimes \alpha \le T(t_1,t)$$

$$(s_1, t_1)\, {}^{\alpha}R_t\, (s_2, t_2) \text{ iff } T(t, t_1) \otimes \alpha \le T(t, t_2)$$

$$(s_1, t_1)\, {}^{\alpha}R'_t\, (s_2, t_2) \text{ iff } T(t, t_1) \otimes \alpha < T(t, t_2)$$

Observe that the graded relations ${}^{\alpha}R$ satisfy the following properties:

1. ${}^{1}R = R$
2. ${}^{\alpha}R \subseteq {}^{\beta}R$, if $\beta \le \alpha$
3. ${}^{0}R = M$, for $R = R_{(s,\,t)}$ and
   ${}^{0}R = M \times M$, for $R = R_t$

Then, given a new problem s_0 we can define the α-level solution set $E_{s0}(\alpha)$ with respect to a constraint Ci using the above corresponding new graded relations. For instance, regarding C1, the α-level solution set is defined as

$$E_{s0}(\alpha) = \{t \in C \,/\, {}^{\alpha}R_{(s0,\,t)} \supseteq M\}.$$

while in the case of constraint C2 the α-level solution set is

$$E_{s0}(\alpha) = \{t \in C \,/\, {}^{\alpha}R_t \supseteq R_{s0}\}$$

while for the cases of constraints C3 and C4 the solution sets are

$$E_{s0}(\alpha) = \{t \in C \,/\, {}^{\alpha}R_t \supseteq R'_{s0}\}$$

and

$$E_{S0}(\alpha) = \{t \in C \,/\, {}^{\alpha}R'_t \supseteq R_{S0}\}$$

respectively.

Definition Let s_0 be a current situation. We define the *possibility* of $t \in C$ being a solution of s_0 as:

$$Pos[s_0](t) = Sup\{\alpha \in [0,1] \,/\, t \in E_{S0}(\alpha)\}.$$

Thus, $Pos[s_0]$ defines on C a *fuzzy set of possible solutions* for s_0, this fuzzy set being possibly non-normalized. In fact, it is easy to check that in the cases of M satisfying the constraints C1, C2 or C4 it holds that $Pos[s_0](t) = 1$ iff $t \in E_{S0}$, this is, $Pos[s_0]$ is normalized iff $E_{S0} \neq \emptyset$. In the case of satisfying constraint C3 it only holds that if $t \in E_{S0}$ then $Pos[s_0](t) = 1$.

Using well-known properties of continuous t-norms $\otimes$ and their residua $\otimes\!\!\rightarrow$, one can compute Poss(t) in the following ways:

$$Pos[s_0](t) = Inf\{S(s_i,s_0) \otimes\!\!\rightarrow T(t_i,t) \,/\, (s_i,t_i) \in M\}$$

in the case of constraint C1, and

$$Pos[s_0](t) = Inf\{T(t_i,t) \otimes\!\!\rightarrow T(t_j,t) \,/\, (s_i,t_i), (s_j,t_j) \in R_{S0}\}$$

in the case of constraint C2. This result would also be (partially) applicable to the cases of constraints C3 and C4.

Finally, in this approach it is natural to take as best solutions of the current problem s_0 those t with greatest possibility value in the above sense.

4 Conclusions and Further Work

In this paper we have been concerned with the fuzzy modelling of some case-based reasoning principles and their use in the CBR task of deriving possible solutions for current problems. The basic tool is the use of fuzzy similarity relations.

Of course the proposals contained in this paper are very preliminar and need to be tested and confronted in realistic applications CBR. In particular, it seems mandatory to incorporating fuzzy relations instead of the classical relations $R(s,t)$, R_s and R_t and their variants. Then one could use them in a fuzzy rule-based modelling of cases and thus to be able to exploit the full power of fuzzy inference machinery.

Acknowledgements

The authors acknowledge partial support from the CICYT project SMASH, TIC96-1138-C04-01/04, and CICYT Project e-INSTITUTOR, TIC 2000-1414.

References

[1] Aamodt A., Plaza E. (1994). Case-based reasoning: Foundational issues, methodological variations and system approaches. *AI Communications*, Vol. 7: 1, pp. 39-59.

[2] Bourrelly L., Chouraqui E., Ricard M. (1983). Formalisation of an approximate reasoning: The analogical reasoning. *Proc. of the IFAC Symp. on Fuzzy Inforn., Knowledge Representation and Decision Analysis,* Marseille, pp. 19-21.

[3] Davies T.R., Russell S.J. (1987). A logical approach to reasoning by analogy. *Proc. of IJCAI'87*, Milano, Italy, pp. 264-270.

[4] Dubois D., Prade H. (1996). What are fuzzy rules and how to use them. *Fuzzy Sets and Systems*, 84, pp. 169-185.

[5] Dubois D., Prade H., Esteva F., Garcia P., Godo L., López de Màntaras R. (1998). Fuzzy Set Modelling in Case-based Reasoning. *International Journal of Intelligent Systems*, Vol. 13, pp. 345-373.

[6] Dubois D., Esteva F., Garcia P., Godo L., Prade H. (1997). A logical approach to interpolation based on similarity relations. *Int. J. of Approximate Reasoning*, Vol. 17 (1), pp. 1-36.

[7] Dubois D., Hüllermeier E., Prade H. (2000). Formalizing case-based inference using fuzzy rules. *Soft Computing in Case-Based reasoning.* Springer-Verlag. pp. 47-72

[8] Hüllermeier E., Dubois D., Prade H. (2000). Knowledge based extrapolation of cases: A possibilistic approach. *Proceedings IPMU 2000,* pp. 1575-1582.

[9] Plaza E., Esteva F., Garcia P., Godo L., López de Màntaras R. (1998). A logical approach to Case-Based reasoning using fuzzy similarity relations. *Journal of Information Sciences* (Elsevier).Vol. 106, pp. 105-122.

[10] Ruspini E.H. (1991). On the semantics of fuzzy logics. *Int. Journal of Approximate Reasoning,* 5, pp. 45-88.

[11] Yager R.R. (1997). Case-based reasoning, fuzzy systems modelling and solution composition. *Proc. of ICCBR-97,* pp. 633-643.

[12] Zadeh L.A. (1971). Similarity relations and fuzzy orderings. *Journal of Information Sciences*, pp. 177-200.