Similarity measures with attributes selection for Case-Based Reasoning in TAVI

Hélène FEUILLATRE[1,2], Vincent AUFFRET[1,2,3], Miguel CASTRO[1,2], Hervé Le Breton[1,2,3], Mireille GARREAU[1,2] and Pascal HAIGRON[1,2]

[1]INSERM U1099, 35000 Rennes, France
[2]Université de Rennes 1, LTSI, 35000 Rennes, France
[3]CHU Rennes, Service de Cardiologie et Maladies Vasculaires, 35000 Rennes, France

Case-Based Reasoning (CBR) system makes the assumption that past experiences may be useful in solving similar current problems. In the case of Transcatheter Aortic Valve Implantation (TAVI), the CBR could help practitioner to plan the procedure. In this study, we focus on retrieving the most similar cases to the current patient, and especially on the similarity measure. We study new similarity measures combining weighted heterogeneous metric and the selection of attributes and cases through the clinical decision tree.

1 Introduction

The proposed Case-Based Reasoning system is a part of the European project EurValve which develops a Clinical Decision Support System to improve the management of valvular heart disease. The developed CBR module focuses on TAVI procedure. The objective of the CBR is to provide the practitioner with a selection of the cases most relevant to the current candidate patient to plan the procedure (which vascular access, which prosthesis).
The CBR solving cycle is composed on four steps [1] (Figure 1):

- RETRIEVE: to get a set of similar cases through a similarity measure.
- REUSE: to take the decision about the best solution (i.e. to adapt similar cases to the current patient).
- REVISE: to get the result after application of the solution and to complete the case.
- RETAIN: to update the Case-Base (useful experience is retained for future reuse).

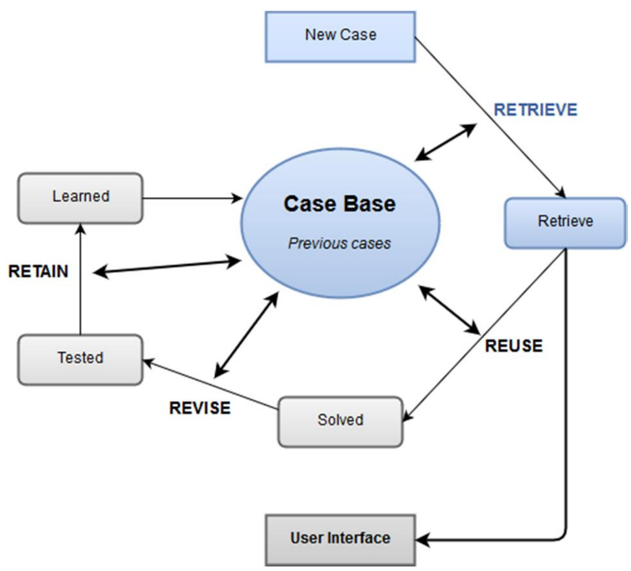

Figure 1: *The four* **ñrò** *of the CBR cycle.*

A case is structured in three categories: attributes from patient characteristics and medical imaging (problem description), procedure characteristics (solution) and procedure outcome.
The feasibility of designing a CBR for TAVI has been recently shown [2], but this work did not focus on investigating similarity functions and only simple representation of cases were considered. Our work mainly deals with case retrieving, which represents the most computational part of the CBR. The other steps are realized through a graphical user interface in order to leave the final choice for the decision making to the practitioner. In the retrieve step, defining a convenient similarity measure (SM) is essential. This study aims to analyse the performance of different SMs and attribute selection thanks to a clinical decision tree (Figure 2).

2 Methods

Generally a standard weighted heterogeneous similarity measure is used in association with the k-nearest neighbour algorithm (kNN) to retrieve similar cases among the Case-Base [3-5]. In our weighted heterogeneous similarity measure, different normalized metrics are considered according to the attribute type (Euclidean distance for numerical data, Hamming distance for Boolean data and a similarity matrix for ordinal data). The management of missing values is also an important issue in the SM. A neutral approach is included to deal with missing information [6]. Each attribute does not have the same importance in the decision-making process. To deal with this issue, we introduced in the SM a clinical decision tree (CDT) related to the searched solution (Figure 2). CDT which is built from guidelines and clinical expertise, allows to select a limited number of attributes and to determine their weight. The higher the weight of an attribute, the more relevant it is.

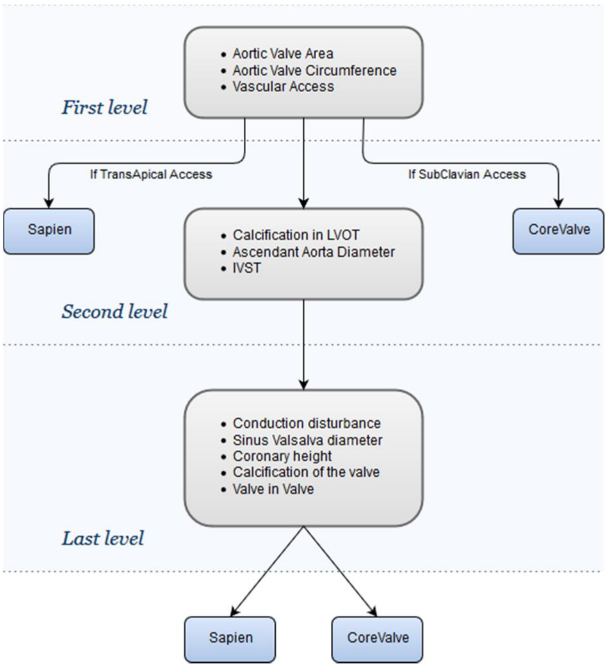

Figure 2: *An example of decision tree for the prosthesis (size and type).*

Four different similarity measures are considered in this study. The first SM (SM1) is based only on the standard weighted heterogeneous similarity measure where all attributes are taken into account. The second SM (SM2) involves the selection of a limited number of attributes through the CDT. The third and the fourth SMs are similar to the previous one, but the weight of attributes is fixed thanks to the CDT. For the third SM (SM3), weights are computed according to the attribute level in the CDT and the total number of levels. As for the fourth SM, a hierarchical similarity measure (HSM) is proposed based on the CDT. The formulation of the metric constituting the HSM is adapted according to each level of the CDT.
First, only attributes present in the first level of the CDT are considered in the similarity measure (Figure 2). Next, a selection of cases is made. Only 50% of the most similar cases are kept to compute

the next value of the similarity measure which takes into account the attributes present both in the previous and current levels of the CDT. This HSM allows selecting the most relevant attributes and the most similar cases.

3 Results

The similarity measures were evaluated for 100 patients with a leave-one-out cross validation. In Table 1, the CBR is performed with the four similarity measures to help practitioner for the decision about the size and the type of the prosthesis. When three most similar patients are selected (k=3 for the kNN algorithm) the correct solution about the size and type of the prosthesis appears at least once in almost 90% of cases, when the CBR is performed with the fourth hierarchical similarity measure. However, with the third SM a loss of precision is shown, contrary to the second SM which uses also the same attributes but with the same weight. Some attributes end up with an insignificant weight value. The method using only the ratio between the level of the attribute in the CDT and the total number of level is not really suitable for our problem.

Table 1: Percentage of correct solutions which appear at least once for the size and the type of the prosthesis when the three most similar patients are selected.

	Percentage of correctly classified cases	
Similarity measure	Prosthesis type	Prosthesis size
SM1	84%	82%
SM2	88%	88%
SM3	88%	82%
HSM	89%	90%

Results show that the similarity measure can be improved with the selection of attributes and cases through a clinical decision tree. In the future, we will evaluate the CBR with the hierarchical similarity measure not only for the choice of the prosthesis but also for the choice of the vascular access.

Acknowledgments

This work was supported by the European Unionôs Horizon 2020 research and innovation programme under grant agreement No 689617 (EurValve).

References

[1] Aamodt A. and Plaza E. (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Communication, 7(1):39-59

[2] El-Fakdi A., Gamero F., Meléndez J., Auffret V. and Haigron P. (2014). eXiTCDSS: A framework for a workflow-based CBR for interventional Clinical Decision Support Systems and its application to TAVI. Expert Systems with Applications, 41(2):284ï 94

[3] Avramenko Y. and Kraslawski A. (2006). Similarity Concept for Case-Based Design in Process Engineering. Computers & Chemical Engineering 30(3):548ï 57

[4] Wilson D.R. and Martinez T.R. (1997). Improved Heterogeneous Distance Functions. Journal of Artificial Intelligence Research 6:1ï 34

[5] Gu D., Liang C. and Zhao H. (2017). A Case-Based Reasoning System Based on Weighted Heterogeneous Value Distance Metric for Breast Cancer Diagnosis. Artificial Intelligence in Medicine. 77:31-47

[6] Guessoum S., Laskri M.T. and Lieber J. (2014). RespiDiag: A Case-Based Reasoning System for the Diagnosis of Chronic Obstructive Pulmonary Disease. Expert Systems with Applications. 41(2):267ï 73