A Case Base Approach to Cardiovascular Diseases using Chest X-ray Image Analysis

Ricardo Faria¹, Victor Alves¹, Filipa Ferraz¹, João Neves², Henrique Vicente¹,³ and José Neves¹*

¹Centro Algoritmi, Universidade do Minho, Braga, Portugal
²Mediclinic Arabian Ranches, PO Box 282602, Dubai, U.A.E.
³Departamento de Química, Escola de Ciências e Tecnologia, Universidade de Évora, Évora, Portugal

Keywords: Chest X-ray Images, Knowledge Representation and Reasoning, Mathematical Logic and Logic Programming, Case-based Reasoning, Similarity Analysis.

Abstract: Cardiovascular Disease (CVD) also known as heart and circulatory disease comprises all the illnesses of the heart and the circulatory system, namely coronary heart disease, angina, heart attack, congenital heart disease or stroke. CVDs are, nowadays, one of the main causes of death. Indeed, this fact reveals the centrality of prevention and how important it is to be aware on these kind of situations. Thus, this work will focus on the development of a decision support system to help to prevent these events from happening, centred on a formal framework based on Mathematical Logic and Logic Programming for Knowledge Representation and Reasoning, complemented with a Case Based Reasoning approach to computing that caters to the handling of incomplete, unknown or even self-contradictory information or knowledge.

1 INTRODUCTION

Chest X-ray is a painless and non-invasive medical procedure to get images of different structures inside the thorax zone, turning easy the access to body parts like heart, lungs or blood vessels. It stands for a symptomatic approach to look at different kinds of illnesses, namely pneumonia, heart failure, lung cancer, lung tissue scarring or sarcoidosis. In this study the X-ray images will be used to evaluate cardiovascular problems, disease that cause 31.5% of the overall deaths in the world every year. Indeed, this work is focused on the development of a hybrid methodology for problem solving, aiming at the elaboration of a decision support systems to detect cardiovascular problems based on parameters obtained from chest X-ray images, like the Cardiac Width (Figure 1(a)), the Thoracic Width (Figure 1(b)) and the Aortic Knuckle Perimeter (AKP) (Figure 1(c)), according to a historical dataset, under a Case Based Reasoning (CBR) approach to problem solving (Aamodt and Plaza, 1994; Richter and Weber, 2013). Undeniably, CBR provides the ability of solving new problems by reusing knowledge acquired from past experiences (Aamodt and Plaza, 1994), i.e., CBR is used especially when similar cases have similar terms and solutions, even when they have different backgrounds (Richter and Weber, 2013). Its use may be found in many different arenas, like in Online Dispute Resolution (Carneiro et al. 2013) or Medicine (Begum et al. 2011; Blanco et al. 2013), just to name a few.

This article is subdivided into five sections. In the former one a brief introduction to the problem is made. Then a mathematical logic approach to Knowledge Representation and Reasoning and a CBR view to computing are introduced. In the third and fourth sections a case study is set. Finally, in the last section the most relevant attainments are described and possible directions for future work are outlined.

2 BACKGROUND

2.1 Knowledge Representation and Reasoning

Many approaches to Knowledge Representation and Reasoning have been proposed using the Logic
Programming (LP) epitome from Mathematical Logic, namely in the area of Model Theory (Kakas et al. 1998; Pereira and Anh, 2009) and Proof Theory (Neves, 1984; Neves et al. 2007). In the present work the Proof Theoretical approach in terms of an extension to the LP language is followed. An Extended Logic Program, or Logic Program, for short, is a finite set of clauses, given in the form:

\[
\{ \\
\neg p \leftarrow \text{not } p, \text{not exception}_p \\
p \leftarrow p_1, \ldots, p_n, \text{not } q_1, \ldots, \text{not } q_m \\
? (p_1, \ldots, p_n, \text{not } q_1, \ldots, \text{not } q_m) (n, m \geq 0) \\
\text{exception}_{p_1} \\
\ldots \\
\text{exception}_{p_j} (0 \leq j \leq k), \text{ being } k \text{ an integer} \\
\}
\]

where the first clause stand for predicate’s closure, “\(\neg\)” denotes logical and, while “?" is a domain atom denoting falsity. The “\(p, q\) and \(p\)” are classical ground literals, i.e., either positive atoms or atoms preceded by the classical negation sign “\(\neg\)” (Neves, 1984. Indeed, “\(\neg\)” stands for a strong declaration that speaks for itself, and not denotes negation-by-failure, or in other words, a flop in proving a given statement, once it was not declared explicitly. Every program is also associated with a set of abducibles (Kakas et al. 1998; Pereira and Anh, 2009), given here in the form of exceptions to the extensions of the predicates that make the overall program, i.e., clauses of the form:

\[\text{exception}_{p_1}, \ldots \text{exception}_{p_j} (0 \leq j \leq k), \text{ being } k \text{ an integer}\]

that stands for data, information or knowledge that cannot be ruled out. On the other hand, clauses of the type:

\[? (p_1, \ldots, p_n, \text{not } q_1, \ldots, \text{not } q_m) (n, m \geq 0)\]

also named invariants or restrictions, allow one to set the context under which the universe of discourse has to be understood. The term \(\text{scoring}_{value}\) stands for the relative weight of the extension of a specific predicate with respect to the extensions of the peers ones that make the inclusive or global program.

Aiming to set one’s approach to Knowledge Representation and Reasoning, two metrics were set, namely the Quality-of-Information (QoI) and the Degree-of-Confidence (DoC). The QoI of a logic program should be understood as a mathematical function that will return a truth-value ranging between 0 and 1, once it is fed with the extension of a given predicate, i.e., \(\text{QoI} = 1\) when the information is known (positive) or false (negative) and \(\text{QoI} = 0\) if the information is unknown. For situations where the extensions of the predicates that make the program also include abducible sets, its terms (or clauses) present a \(\text{QoI} \in [0, 1]\) (Fernandes et al. 2016).

Figure 1: The Chest X-ray’s parameters that were taken into account in this study, i.e., Cardiac Width (a), Thoracic Width (b), and Aortic Knuckle Perimeter (c).
The DoCs, in turn, stand for one’s confidence that the argument values or attributes of the terms that make the extension of a given predicate, having into consideration their domains, are, in a given interval (Neves et al. 2015). The DoC is figured out using \( \text{DoC} = \sqrt{1 - \Delta t^2} \), where \( \Delta t \) stands for the argument interval length, which was set to the interval \([0, 1]\), since the ranges of attributes values for a given predicate and respective domains were normalized, in terms of the expression \( (Y - Y_{\text{min}}) / (Y_{\text{max}} - Y_{\text{min}}) \), where the \( Y_i \) stand for themselves.

Thus, the universe of discourse is engendered according to the information presented in the extensions of such predicates, according to productions of the type:

\[
predicate_i = \bigcup_{1 \leq j \leq m} \text{clause}_j \left( \left( (A_{x_1, j}, B_{x_1, j}) \langle \text{QoI}_{x_1, j}, \text{DoC}_{x_1, j} \rangle \right), \ldots, \right. \]
\[
\left. \ldots, \left( (A_{x_n, j}, B_{x_n, j}) \langle \text{QoI}_{x_n, j}, \text{DoC}_{x_n, j} \rangle \right) : \langle \text{QoI}_j, : \text{DoC}_j \rangle \right)
\]

where \( \cup, m \) and \( l \) stand, respectively, for set union, the cardinality of the extension of predicate, and the number of attributes of each clause (Neves et al. 2015). On the other hand, either the subscripts of the QoI and the DoC, or those of the pairs \( (A_i, B_i) \), i.e., \( x_1, \ldots, x_n \), stand for the attributes’ clauses values ranges.

### 2.2 Case based Computing

The CBR approach to computing stands for an act of finding and justifying a solution to a given problem based on the consideration of the solutions of similar past ones, either using old solutions, or by reprocessing and generating new data or knowledge from the old ones (Aamodt and Plaza, 1994; Richter and Weber, 2013). In CBR the cases are stored in a Cases repository, and those cases that are similar (or close) to a new one are used in the problem solving process.

The typical CBR cycle presents the mechanism that should be followed to have a consistent model. The first stage entails an initial description and a reprocessing of the problem’s data or knowledge. The new case is defined and it is used to retrieve one or more cases from the repository, i.e., at this point it is imperative to identify the characteristics of the new case and retrieve cases with a higher degree of similarity to it. Thereafter, a solution to the problem emerges, on the Reuse phase, based on the blend of the new case with the retrieved ones. The suggested solution is reused (i.e., adapted to the new case), and a solution is provided (Aamodt and Plaza, 1994; Richter and Weber, 2013). However, when adapting the solution it is crucial to have feedback from the user, since automatic adaptation in existing systems is almost impossible. This is the Revise stage, in which the suggested solution is tested by the user, allowing for its correction, adaptation and/or modification, originating the test repaired case’s phase that sets the solution to the new problem. Thus, one is faced with an iterative process since the solution must be tested and adapted, while the result of considering that solution is inconclusive. During the Retain (or Learning) stage the case is learned and the repository is updated, by inserting the new case (Aamodt and Plaza, 1994; Richter and Weber, 2013).

On the other hand, and despite promising results, the current CBR systems are neither complete nor adaptable for all domains. In some cases, the user cannot choose the similarity(ies) method(s) used in the retrieval phase and is required to follow the system defined one(s), even if they do not meet their needs. Moreover, in real problems, the access to all necessary information is not always possible, since existing CBR systems have limitations related to the capability of dealing, explicitly, with unknown, incomplete, and even self-contradictory information. To make a change, a different CBR cycle was induced (Figure 2). It takes into consideration the case’s QoI and DoC metrics. It also contemplates a cases optimization process present in the Case-base, whenever they do not comply with the terms under which a given problem as to be addressed (e.g., the expected DoC on a prediction was not attained). In this process may be used Artificial Neural Networks (Haykin, 2009; Vicente et al. 2012), Particle Swarm Optimization (Mendes et al. 2003) or Genetic Algorithms (Neves et al. 2007), just to name a few. Indeed, the optimization process generates a set of new cases which must be in conformity with the invariant:

\[
\bigcap_{\text{new cases} \neq \emptyset} \bigcap_{i=1}^{n} (B_{i, \text{E}_i}) \neq \emptyset \tag{1}
\]

that states that the intersection of the attribute’s values ranges for the cases’ set that make the Case-base or their optimized counterparts \( (B_i) \) (being \( n \) its cardinality), and the ones that were object of a process of optimization \( (E_i) \), cannot be empty.
3 METHODS

Aiming to develop a predictive model to estimate the risk of cardiovascular diseases, a database was set, built on 388 health records of patients from a major health care institution in the North of Portugal.

The patients included in this study aged between 19 to 93 years old, with an average of 49 ± 12 years old. The gender distribution was 43.8% and 56.2% for male and female, respectively.

After having collected the data it is possible to build up a knowledge database given in terms of the extensions of the relations or predicates depicted in Figure 3, which stand for a situation where one has to manage information aiming to access the cardiovascular disease predisposing. The tables include features obtained by both objective and subjective methods. The physicians may populate some issues and others may be perceived by additional tests. The software imageJ (Rasband, 2016) was used to extract the necessary features from X-ray images (Figure 1). Under this scenario some incomplete and/or default data is also present. For instance, the Triglycerides in case 2 are unknown (depicted by the symbol ⊥), while the Risk Factors range in the interval [1, 2]. The values presented in the Risk Factors column of Cardiovascular Diseases Predisposing table is the sum of the correspondent table values, ranging between 0 and 4. The CTR column is the Cardiac Thoracic Ratio computed using cardiac and thoracic width. The Descriptions column stands for free text fields that allow for the registration of relevant patient features.

Applying the algorithm presented in Neves et al. (2015) to the fields that make the knowledge base for Cardiovascular Diseases Predisposing (Figure 3), excluding at this stage of such a process the Description one, and looking to the DoCs’ values, it is possible to set the arguments of the predicate cardiovascular diseases predisposing (cdp) referred to below, whose extension denote the objective function with respect to the problem under analyze:

\[
\text{cdp: Age, SystolicBloodPressure, CholesterolLDL, CholesterolHDL, Triglycerides Cardiac Ratio, AorticAnklePerimeter, RiskFactors} \rightarrow \{0, 1\}
\]

where 0 (zero) and 1 (one) denote, respectively, the truth values false and true.
The application of the algorithm presented in Neves et al. (2015) comprises several phases. In the former one the clauses or terms that make extension of the predicate under study are established. In the next stage the boundaries of the attributes intervals are set in the interval \([0, 1]\) according to a normalization process in terms of the expression \((Y_s - Y_{\text{min}}) / (Y_{\text{max}} - Y_{\text{min}})\), where the \(Y_s\) stand for themselves. Finally, the \(\text{DoC}\) is evaluated as described in section 2.1. Exemplifying the application of the algorithm referred to above, to a term (patient) that presents the feature vector \(\text{Age} = 64, \text{SBP} = L, \text{Chol}_{\text{LDL}} = 128, \text{Chol}_{\text{HDL}} = 47, \text{Trigly} = 203, \text{CTR} = 0.45, \text{AKP} = 124, \text{RF} = [1, 2]\), one may have:

\[
\begin{align*}
\neg \text{cdp} \left( \left( \text{Age}, \text{Age} \right)(\text{Qol}_{\text{Age}}, \text{DoC}_{\text{Age}}), \left( \text{SBP}, \text{SBP} \right)(\text{Qol}_{\text{SBP}}, \text{DoC}_{\text{SBP}}) \right), \ldots, \\
\left( \text{RF}, \text{RF} \right)(\text{Qol}_{\text{RF}}, \text{DoC}_{\text{RF}}) \right) \\
\neg \text{not cdp} \left( \left( \text{Age}, \text{Age} \right)(\text{Qol}_{\text{Age}}, \text{DoC}_{\text{Age}}), \left( \text{SBP}, \text{SBP} \right)(\text{Qol}_{\text{SBP}}, \text{DoC}_{\text{SBP}}) \right), \ldots, \\
\left( \text{RF}, \text{RF} \right)(\text{Qol}_{\text{RF}}, \text{DoC}_{\text{RF}}) \right) \\
\text{cdp} \left( \left( \left( 64, 64 \right)(\text{DoC}_{\left[64, 64\right]}), \left( 70, 200 \right)(\text{DoC}_{\left[70, 200\right]}), \ldots, \\
\left( 1, 2 \right)(\text{DoC}_{\left[1, 2\right]}), \left( 19, 93 \right)(\text{DoC}_{\left[19, 93\right]} \ldots \left( 70, 200 \right)(\text{DoC}_{\left[70, 200\right]}), \ldots \right) \right) :: 1 :: \text{DoC} \\
\end{align*}
\]

The predicate’s extension that sets the Universe-of-Discourse for the term under observation is fixed.

\[
\begin{align*}
\{ & \\
\neg \text{cdp} \left( \left( \text{Age}, \text{Age} \right)(\text{Qol}_{\text{Age}}, \text{DoC}_{\text{Age}}), \left( \text{SBP}, \text{SBP} \right)(\text{Qol}_{\text{SBP}}, \text{DoC}_{\text{SBP}}) \right), \ldots, \\
\left( \text{RF}, \text{RF} \right)(\text{Qol}_{\text{RF}}, \text{DoC}_{\text{RF}}) \right) \\
\neg \text{not cdp} \left( \left( \text{Age}, \text{Age} \right)(\text{Qol}_{\text{Age}}, \text{DoC}_{\text{Age}}), \left( \text{SBP}, \text{SBP} \right)(\text{Qol}_{\text{SBP}}, \text{DoC}_{\text{SBP}}) \right), \ldots, \\
\left( \text{RF}, \text{RF} \right)(\text{Qol}_{\text{RF}}, \text{DoC}_{\text{RF}}) \right) \\
\text{cdp} \left( \left( \left( 64, 64 \right)(\text{DoC}_{\left[64, 64\right]}), \left( 70, 200 \right)(\text{DoC}_{\left[70, 200\right]}), \ldots, \\
\left( 1, 2 \right)(\text{DoC}_{\left[1, 2\right]}), \left( 19, 93 \right)(\text{DoC}_{\left[19, 93\right]} \ldots \left( 70, 200 \right)(\text{DoC}_{\left[70, 200\right]}), \ldots \right) \right) \} :: 1 :: \text{DoC} \\
\end{align*}
\]
%The attribute’s boundaries are set to the interval \([0, 1]\), according to a normalization process that uses the expression \((Y - Y_{\text{min}})/(Y_{\text{max}} - Y_{\text{min}})\)\%

\[
\neg \text{cdp} \left( \left( (A_{\text{Age}}, B_{\text{Age}})(QoI_{\text{Age}}, DoC_{\text{Age}}) \right), \left( (A_{\text{SBP}}, B_{\text{SBP}})(QoI_{\text{SBP}}, DoC_{\text{SBP}}) \right), \cdots, \left( (A_{\text{RF}}, B_{\text{RF}})(QoI_{\text{RF}}, DoC_{\text{RF}}) \right) \right) \\
\neg \text{not cdp} \left( \left( (A_{\text{Age}}, B_{\text{Age}})(QoI_{\text{Age}}, DoC_{\text{Age}}) \right), \left( (A_{\text{SBP}}, B_{\text{SBP}})(QoI_{\text{SBP}}, DoC_{\text{SBP}}) \right), \cdots, \left( (A_{\text{RF}}, B_{\text{RF}})(QoI_{\text{RF}}, DoC_{\text{RF}}) \right) \right) \\
\text{cdp} \left( \left( (0.61, 0.61)(1_{[0.61, 0.61]} DoC_{[0.61, 0.61]}), \left( (0, 1)(1_{[0, 1]} DoC_{[0, 1]}), \cdots, \left( (0.25, 0.5)(1_{[0.25, 0.5]} DoC_{[0.25, 0.5]}) \right) \right) \right) \bowtie 1 :: \text{DoC} \\[0, 1\] \[0, 1\] \[0, 1\] \hspace{1cm} \text{attribute's domains once normalized} \\
\text{vdash} \quad 1 \\
%

%The DoC’s values are evaluated% 

\[
\neg \text{cdp} \left( \left( (A_{\text{Age}}, B_{\text{Age}})(QoI_{\text{Age}}, DoC_{\text{Age}}) \right), \left( (A_{\text{SBP}}, B_{\text{SBP}})(QoI_{\text{SBP}}, DoC_{\text{SBP}}) \right), \cdots, \left( (A_{\text{RF}}, B_{\text{RF}})(QoI_{\text{RF}}, DoC_{\text{RF}}) \right) \right) \\
\neg \text{not cdp} \left( \left( (A_{\text{Age}}, B_{\text{Age}})(QoI_{\text{Age}}, DoC_{\text{Age}}) \right), \left( (A_{\text{SBP}}, B_{\text{SBP}})(QoI_{\text{SBP}}, DoC_{\text{SBP}}) \right), \cdots, \left( (A_{\text{RF}}, B_{\text{RF}})(QoI_{\text{RF}}, DoC_{\text{RF}}) \right) \right) \\
\text{cdp} \left( \left( (0.61, 0.61)(1_{[0.61, 0.61]} DoC_{[0.61, 0.61]}), \left( (0, 1)(1_{[0, 1]} DoC_{[0, 1]}), \cdots, \left( (0.25, 0.5)(1_{[0.25, 0.5]} DoC_{[0.25, 0.5]}) \right) \right) \right) \bowtie 1 :: 0.87 \\[0, 1\] \[0, 1\] \[0, 1\] \hspace{1cm} \text{attribute's values ranges once normalized and respective QoI and DoC values} \\
\text{vdash} \quad 1 \\
%

End

4 A CASE BASED REASONING APPROACH TO COMPUTING

The framework presented previously shows how the information comes together and how it is processed. In this section, a soft computing approach was set to model the universe of discourse, where the computational part is based on a CBR approach to computing. Contrasting with other problem solving strategies (e.g., those that use Decision Trees or Artificial Neural Networks), relatively little work is done offline. Undeniably, in almost all the situations the work is performed at query time. The main difference between this approach and the typical CBR one relies on the fact that not only all the cases have their arguments set in the interval \([0, 1]\), a situation that is complemented with the prospect of handling incomplete, unknown, or even self-contradictory data, information or knowledge. Thus, the classic CBR cycle was changed (Figure 2), being the Case-base given in terms of the pattern:

\[
\text{Case} = \{ \text{Raw}_{\text{data}}, \text{Normalized}_{\text{data}}, \text{Description}_{\text{data}} \} \quad (2)
\]
where the Description data field will not be object of attention in this study.

Undeniably, when confronted with a new case, the system is able to retrieve all cases that meet such a case structure and optimize such a population, having in consideration that the cases retrieved from the Case-base must satisfy the invariant present in equation (1), in order to ensure that the intersection of the attributes range in the cases that make the Case-base repository or their optimized counterparts, and the equals in the new case cannot be empty. Having this in mind, the algorithm given in Neves et al. (2016) is applied to the new case that presents the feature vector \( Age = 57, SBP = 118, CholLDL = 1, CholHDL = 1, Trigly = 1, CTR = 0.43, AKP = 127, RF = [1, 3], \) with the outcome:

\[
\begin{align*}
\text{cdp}_{\text{new case}} & \left( \left( 0.51, 0.51 \right)(1, 1), \ldots, \right. \\
& \left. \left( 0.25, 0.75 \right)(1, 0.87) \right) :: 1 :: 0.61
\end{align*}
\]

Now, the new case may be portrayed on the Cartesian plane in terms of its \( QoI \) and DoC, and by using clustering methods (Neves et al. 2016) it is feasible to identify the cluster(s) that intermingle with the new one (epitomized as a square in Figure 4). The new case is compared with every retrieved case from the clusters using a similarity function \( \text{sim} \), given in terms of the average of the modulus of the arithmetic difference between the arguments of each case of the selected cluster and those of the new case. Thus, one may have:

\[
\begin{align*}
\text{retrieved case}_1 & \left( \left( 0.57, 0.57 \right)(1, 1), \ldots, \right. \\
& \left. \left( 0.5, 0.5 \right)(1, 1) \right) :: 1 :: 0.73 \\
\text{retrieved case}_2 & \left( \left( 0.49, 0.49 \right)(1, 1), \ldots, \right. \\
& \left. \left( 0.25, 0.5 \right)(1, 0.97) \right) :: 1 :: 0.62 \\
& \vdots \\
\text{retrieved case}_j & \left( \left( 0.55, 0.55 \right)(1, 1), \ldots, \right. \\
& \left. \left( 0.75, 0.75 \right)(1, 1) \right) :: 1 :: 0.88
\end{align*}
\]

Assuming that every attribute has equal weight, for the sake of presentation, the \( d\text{is}(\text{similarity}) \), in terms of DoC, between new case and the retrieved case \( j \), may be computed as follows:

\[
\begin{align*}
\text{dist}_{\text{new case} \rightarrow j} &= \frac{\|1 - 1\| + \cdots + \|0.87 - 1\|}{8} \\
&= 0.16
\end{align*}
\]  

Thus, one may have:

\[
\begin{align*}
\text{sim}_{\text{new case} \rightarrow 1} &= 1 \times 0.84 = 0.84 \quad (4)
\end{align*}
\]

These procedures should be applied to the remaining cases of the retrieved clusters in order to obtain the most similar ones, which may stand for the possible solutions to the problem. This approach allows users to define the most appropriate similarity threshold to address the problem (i.e., it gives the user the possibility to narrow the number of selected cases with the increase of the similarity threshold).

The proposed model was tested on a real data set with 388 examples. Thus, the dataset was divided in exclusive subsets through a ten-folds cross validation (Haykin, 2009). In the implementation of the respective dividing procedures, ten executions were performed for each one of them. Table 1 presents the coincidence matrix of the CBR model, where the values presented denote the average of 25 (twenty five) experiments. A perusal to Table 1 shows that the model accuracy was 91.8% (i.e., 356 instances correctly classified in 388). Thus, the predictions made by the CBR model are satisfactory, attaining accuracies higher than 90%. The sensitivity and specificity of the model were 92.8% and 90.5%, while Positive and Negative Predictive Values were 91.9% and 91.5%, respectively. The ROC curve is shown in Figure 5. The area under ROC curve (0.92) denotes that the model exhibits a good performance in the assessment of cardiovascular diseases predisposing.

\[\text{Figure 4: A case's set divided into clusters.}\]

\[
\begin{align*}
\text{sim}_{\text{new case} \rightarrow 1} &= 1 \times 0.84 = 0.84 \quad (4)
\end{align*}
\]

Therefore, the \( \text{sim}(\text{ilarity}) \), i.e., \( \text{sim}_{\text{new case} \rightarrow 1} \) is set as \( 1 - 0.16 = 0.84 \). Regarding \( QoI \) the procedure is similar, returning \( \text{sim}_{\text{QoI, new case} \rightarrow 1} = 1 \). Thus, one may have:

\[
\begin{align*}
\text{sim}_{\text{QoI, new case} \rightarrow 1} &= 1 \times 0.84 = 0.84 \quad (4)
\end{align*}
\]
5 CONCLUSIONS

This work presents a Logic Programming based Decision Support System to estimate the cardiovascular diseases predisposing, i.e., it is centred on a formal framework based on LP for Knowledge Representation and Reasoning, complemented with a CBR approach to computing that caters for the handling of incomplete, unknown, or even self-contradictory information. The proposed model is able to provide adequate responses, once the overall accuracy is higher than 90%. The computational framework presented above uses powerful knowledge representation and reasoning methods to set the structure of the information and the associate inference mechanisms. Indeed, it has also the potential to be disseminated across other prospective areas, therefore validating an universal attitude. Additionally, it gives the user the possibility to narrow the search space for similar cases at runtime by choosing the most appropriate strategy to address the problem.

ACKNOWLEDGEMENTS

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2013.

REFERENCES

Figueiredo, M., Esteves, L., Neves, J., Vicente, H., 2016. A data mining approach to study the impact of the methodology followed in chemistry lab classes on the weight attributed by the students to the lab work on learning and motivation. Chemistry Education Research and Practice, 17: 156-171.

