Screening a Case Base for Stroke Disease Detection

José Neves[1(✉)], Nuno Gonçalves[2], Ruben Oliveira[2], Sabino Gomes[2], João Neves[3], Joaquim Macedo[1], António Abelha[1], César Analide[1], José Machado[1], Manuel Filipe Santos[1], and Henrique Vicente[1,4]

[1] Centro Algorithmi, Universidade do Minho, Braga, Portugal
{jneves,macedo,abelha,analide,jmac}@di.uminho.pt,
mfs@dsi.uminho.pt

[2] Departamento de Informática, Universidade do Minho, Braga, Portugal
{pg24168,pg24166}@alunos.uminho.pt,
sabinogomes.antonio@gmail.com

[3] Drs. Nicolas and Asp, Dubai, United Arab Emirates
joaocpneves@gmail.com

[4] Departamento de Química, Escola de Ciências e Tecnologia,
Universidade de Évora, Évora, Portugal
hvicente@uevora.pt

Abstract. Stroke stands for one of the most frequent causes of death, without distinguishing age or genders. Despite representing an expressive mortality figure, the disease also causes long-term disabilities with a huge recovery time, which goes in parallel with costs. However, stroke and health diseases may also be prevented considering illness evidence. Therefore, the present work will start with the development of a decision support system to assess stroke risk, centered on a formal framework based on Logic Programming for knowledge representation and reasoning, complemented with a Case Based Reasoning (CBR) approach to computing. Indeed, and in order to target practically the CBR cycle, a normalization and an optimization phases were introduced, and clustering methods were used, then reducing the search space and enhancing the cases retrieval one. On the other hand, and aiming at an improvement of the CBR theoretical basis, the predicates` attributes were normalized to the interval 0…1, and the extensions of the predicates that match the universe of discourse were rewritten, and set not only in terms of an evaluation of its Quality-of-Information (QoI), but also in terms of an assessment of a Degree-of-Confidence (DoC), a measure of one's confidence that they fit into a given interval, taking into account their domains, i.e., each predicate attribute will be given in terms of a pair (QoI, DoC), a simple and elegant way to represent data or knowledge of the type incomplete, self-contradictory, or even unknown.

Keywords: Stroke Disease · Logic Programming · Knowledge Representation and Reasoning · Case Based Reasoning · Similarity Analysis

F. Martínez-Álvarez et al. (Eds.): HAIS 2016, LNAI 9648, pp. 3–13, 2016.
DOI: 10.1007/978-3-319-32034-2_1

1 Introduction

Stroke stands for a blood supply interruption that occurs in the brain, once a blood vessel is blocked, causing an ischaemic hit or bursts, leading to a hemorrhagic blow. Being a major factor related with mortality, this disease is closely followed with the main purpose of preventing it from happen, once, when diagnosed, it becomes less hazardous and more treatable, comparing with similar ones [1]. However, there are several factors associated with stroke, which transport a higher probability of occurrence, and may lead to such a happening. Some of these risk factors can be avoid or controlled, like high blood pressure [1, 2], cigarette smoking [2, 3], diabetes mellitus [4, 5], high blood cholesterol [6, 7], or the absence of physical activity [8, 9].

Despite these causes there are those who cannot be controlled, such as age (older people have more tendency to stroke [2, 10]), and gender (stroke is more common in men than in women, and the mere fact of having suffered a previous stroke represents an increased risk not controlled by any means [2, 11]), ethnicity [10, 11], among others. In this work it will be emphasized the prediction of a giving event, according to a historical dataset, under a Case Based Reasoning (CBR) approach to computing [12, 13]. Indeed, CBR provides the ability of solving new problems by reusing knowledge acquired from past experiences [12], i.e., CBR is used especially when similar cases have similar terms and solutions, even when they have different backgrounds [13]. Indeed, its use may be found in different arenas, namely in The Law, Online Dispute Resolution [14, 15] or Medicine [16, 17], just to name a few.

It must be also highlighted that up to present CBR systems have been unable to deal with incomplete, self-contradictory, or even unknown information. As a matter of fact the approach to CBR presented in this work will be a generic one and will have a focus on such a setting. It brings to evidence that the first step to be tackled is related with the construction of the Case Base. Thus, a normalization and optimization phases were introduced and clustering methods were used to distinguish and aggregate collections of historical data, in order to reduce the search space that speeds up the retrieve stage and all associated computational processes.

The article develops along five sections. In a former one a brief introduction to the problem is made. Then the proposed approach to knowledge representation and reasoning is introduced. In the third and fourth sections it is assumed a case study and presented a solution to the problem. Finally, in the last section the most relevant conclusions are described and possible directions for future work are outlined.

2 Knowledge Representation and Reasoning

Many approaches to knowledge representation and reasoning have been proposed using the Logic Programming (LP) paradigm, namely in the area of Model Theory [18, 19], and Proof Theory [20, 21]. In this work it is followed the proof theoretical approach in terms of an extension to LP. An Extended Logic Program is a finite set of clauses in the form:

$$
\begin{array}{l}
\{ \\
\quad p \leftarrow p_1, \cdots, p_n, not\ q_1, \cdots, not\ q_m \\
\quad ?\,(p_1, \cdots, p_n, not\ q_1, \cdots, not\ q_m)\ \ (n, m \geq 0) \\
\quad exception_{p_1}\ \ldots\ exception_{p_j}\ \ (j \leq m, n) \\
\} :: scoring_{value}
\end{array}
$$

where "*?*" is a domain atom denoting falsity, the p_i, q_j, and p are classical ground literals, i.e., either positive atoms or atoms preceded by the classical negation sign $\neg$ [20]. Under this formalism, every program is associated with a set of abducibles [18, 19], given here in the form of exceptions to the extensions of the predicates that make the program. The term $scoring_{value}$ stands for the relative weight of the extension of a specific *predicate* with respect to the extensions of the peers ones that make the overall program.

In order to evaluate the knowledge that stems from a logic program, an assessment of the *Quality-of-Information* (*QoI*), given by a truth-value in the interval [0, 1], inclusive in dynamic environments aiming at decision-making purposes, is set [22, 23]. Indeed, the objective is to build a quantification process of *QoI* and measure one's Degree of Confidence (*DoC*) that the argument values or attributes of the terms that make the extension of a given predicate with relation to their domains fit into a given interval [24]. Thus, the universe of discourse is engendered according to the information presented in the extensions of a given set of predicates, according to productions of the type:

$$predicate_i - \bigcup_{1 \leq j \leq m} clause_j((QoI_{x_1}, DoC_{x_1}), \cdots, (QoI_{x_m}, DoC_{x_m})) :: QoI_i :: DoC_i \quad (1)$$

where ∪ and m stand, respectively, for *set union* and the *cardinality* of the extension of *predicate*$_i$. QoI_i and DoC_i stand for themselves [24].

3 A Case Study

As a case study, consider a database given in terms of the extensions of the relations (or tables) depicted in Fig. 1, which stand for a situation where one has to manage information about stroke predisposing detection. The tables include features obtained by both objective and subjective methods, i.e., the physicians will fill the tables that are related to the *Stroke Predisposing* one while executing the health check. The clinics may populate some issues, others may be perceived by additional exams.

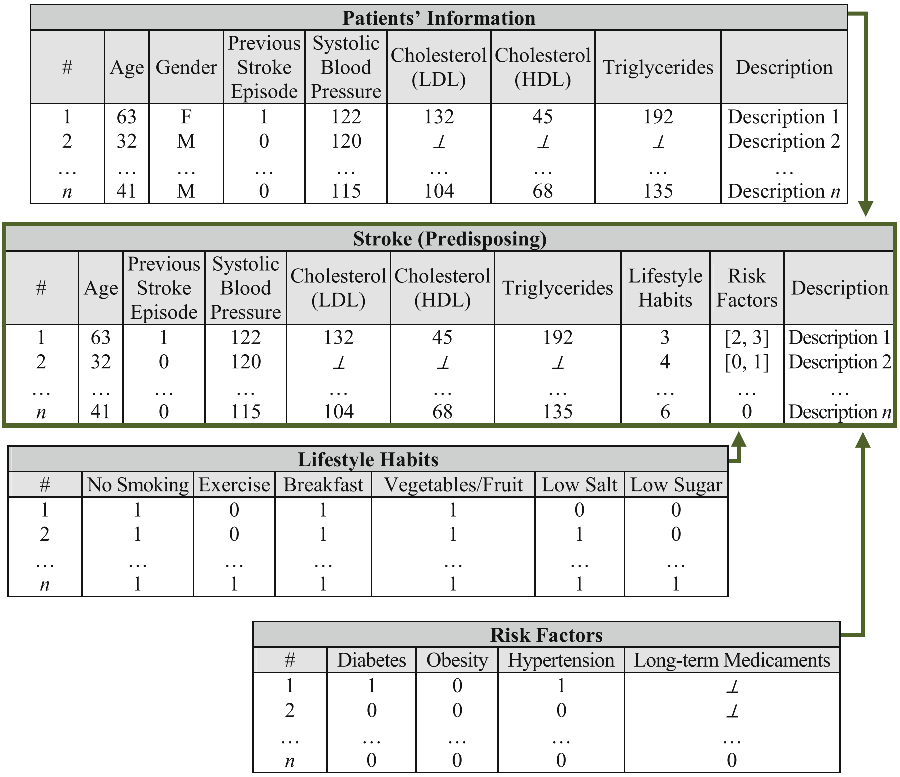
Under this scenario some incomplete and/or default data is also available. For instance, the *Triglycerides* in case 2 is unknown, while the *Risk Factors* range in the interval [0, 1]. In *Previous Stroke Episode* column 0 (zero) and 1 (one) denote, respectively, *nonoccurrence* and *occurrence*. In *Lifestyle Habits* and *Risk Factors* tables 0 (zero) and 1 (one) denote, respectively, *yes* and *no*. The values presented in the *Lifestyle Habits* and *Risk Factors* columns of *Stroke Predisposing* table are the sum of the correspondent table values, ranging between [0, 6] and [0, 4], respectively. The *Descriptions* column stands for free text fields that allow for the registration of relevant patient features.

Applying the rewritten algorithm presented in [24], to all the fields that make the knowledge base for *Stroke Predisposing* (Fig. 1), excluding of such a process the *Description* one, and looking to the *DoCs* values obtained in this manner, it is possible to set the arguments of the predicate referred to below, that also denotes the objective function with respect to the problem under analyze.

$$stroke : Age, P_{revious}S_{troke}E_{pisodes}, B_{lood}S_{ystolic}P_{ressure}, Chol_{esterol_{LDL}},$$
$$Chol_{esterol_{HDL}}, Trigly_{cerides}, L_{ifestyle}H_{abits}, R_{isk}F_{actors} \rightarrow \{0, 1\}$$

where 0 (zero) and 1 (one) denote, respectively, the truth values *false* and *true*.

Exemplifying the application of the rewritten algorithm presented in [24], in relation to the term that presents the feature vector $Age = 69$, $P_{revious}\ S_{troke}\ E_{pisodes} = 1$, $S_{ystolic}\ B_{lood}\ P_{ressure} = \perp$, $Chol_{esterolLDL} = 131$, $Chol_{esterolHDL} = 49$, $Trigly_{cerides} = 200$, $L_{ifestyle}\ H_{abits} = 4$, $R_{isk}\ F_{actors} = [1, 2]$, one may have:



Patients' Information

#	Age	Gender	Previous Stroke Episode	Systolic Blood Pressure	Cholesterol (LDL)	Cholesterol (HDL)	Triglycerides	Description
1	63	F	1	122	132	45	192	Description 1
2	32	M	0	120	⊥	⊥	⊥	Description 2
…	…	…	…	…	…	…	…	…
n	41	M	0	115	104	68	135	Description *n*

Stroke (Predisposing)

#	Age	Previous Stroke Episode	Systolic Blood Pressure	Cholesterol (LDL)	Cholesterol (HDL)	Triglycerides	Lifestyle Habits	Risk Factors	Description
1	63	1	122	132	45	192	3	[2, 3]	Description 1
2	32	0	120	⊥	⊥	⊥	4	[0, 1]	Description 2
…	…	…	…	…	…	…	…	…	…
n	41	0	115	104	68	135	6	0	Description *n*

Lifestyle Habits

#	No Smoking	Exercise	Breakfast	Vegetables/Fruit	Low Salt	Low Sugar
1	1	0	1	1	0	0
2	1	0	1	1	1	0
…	…	…	…	…	…	…
n	1	1	1	1	1	1

Risk Factors

#	Diabetes	Obesity	Hypertension	Long-term Medicaments
1	1	0	1	⊥
2	0	0	0	⊥
…	…	…	…	…
n	0	0	0	0

Fig. 1. A fragment of the knowledge base for Stroke Predisposing Diagnosis.

Begin, (DoCs evaluation),

The predicate's extension that maps the Universe-of-Discourse for the term under observation is set ←

$\{$

$$\neg\, stroke\left(\left(QoI_{Age}, DoC_{Age}\right),\ \cdots, \left(QoI_{SBP}, DoC_{SBP}\right),\ \cdots, \left(QoI_{RF}, DoC_{RF}\right)\right)$$

$$\leftarrow not\ stroke\left(\left(QoI_{Age}, DoC_{Age}\right),\ \cdots, \left(QoI_{SBP}, DoC_{SBP}\right),\ \cdots, \left(QoI_{RF}, DoC_{RF}\right)\right)$$

$$stroke\underbrace{\left(\left(1_{69}, DoC_{69}\right),\ \cdots, \left(1_{\perp}, DoC_{\perp}\right),\ \cdots, \left(1_{[1,\,2]}, DoC_{[1,\,2]}\right)\right)}_{attribute`s\ values} :: 1 :: DoC$$

$$\underbrace{[22, 95] \quad \cdots \quad [70, 200] \quad \cdots \quad [0, 4]}_{attribute`s\ domains}$$

$\} :: 1$

The attribute's values ranges are rewritten ←

$\{$

$$\neg\, stroke\left(\left(QoI_{Age}, DoC_{Age}\right),\ \cdots, \left(QoI_{SBP}, DoC_{SBP}\right),\ \cdots, \left(QoI_{RF}, DoC_{RF}\right)\right)$$

$$\leftarrow not\ stroke\left(\left(QoI_{Age}, DoC_{Age}\right),\ \cdots, \left(QoI_{SBP}, DoC_{SBP}\right),\ \cdots, \left(QoI_{RF}, DoC_{RF}\right)\right)$$

$$stroke\underbrace{\left(\left(1_{[69,69]}, DoC_{[69,69]}\right),\ \cdots, \left(1_{[70,200]}, DoC_{[70,200]}\right),\ \cdots, \left(1_{[1,\,2]}, DoC_{[1,\,2]}\right)\right)}_{attribute`s\ values}$$

$$:: 1 :: DoC$$

$$\underbrace{[22, 95] \qquad \cdots \qquad [70, 200] \qquad \cdots \qquad [0, 4]}_{attribute`s\ domains}$$

$\} :: 1$

The attribute's boundaries are set to the interval [0,1] ←

$\{$

$$\neg\, stroke\left(\left(QoI_{Age}, DoC_{Age}\right),\ \cdots, \left(QoI_{SBP}, DoC_{SBP}\right),\ \cdots, \left(QoI_{RF}, DoC_{RF}\right)\right)$$

$$\leftarrow not\ stroke\left(\left(QoI_{Age}, DoC_{Age}\right),\ \cdots, \left(QoI_{SBP}, DoC_{SBP}\right),\ \cdots, \left(QoI_{RF}, DoC_{RF}\right)\right)$$

$$stroke\underbrace{\left(\left(1_{[0.64,0.64]}, DoC_{[0.64,0.64]}\right), \cdots, \left(1_{[0,1]}, DoC_{[0,1]}\right), \cdots, \left(1_{[0.25,\,0.5]}, DoC_{[0.25,\,0.5]}\right)\right)}_{attribute`s\ values\ once\ normalized}$$

$$:: 1 :: DoC$$

$$\underbrace{[0, 1] \qquad \cdots \quad [0, 1] \qquad \cdots \qquad [0, 1]}_{attribute`s\ domains\ once\ normalized}$$

$\} :: 1$

***The DoC's values are evaluated* ←**

$\{$

$$\neg\, stroke\left(\left(QoI_{Age}, DoC_{Age}\right),\ \cdots,\left(QoI_{SBP}, DoC_{SBP}\right),\ \cdots,\left(QoI_{RF}, DoC_{RF}\right)\right)$$

$$\leftarrow not\ stroke\left(\left(QoI_{Age}, DoC_{Age}\right),\ \cdots,\left(QoI_{SBP}, DoC_{SBP}\right),\ \cdots,\left(QoI_{RF}, DoC_{RF}\right)\right)$$

$$stroke\,\underbrace{\left((1,\ 1),\qquad \cdots,\qquad (1,\ 0),\qquad \cdots,\qquad (1,\ 0.97)\right)}_{attribute`s\ quality-of-information\ and\ respective\ confidence\ values} :: 1 :: 0.89$$

$$\underbrace{[0.64, 0.64]\qquad \cdots\qquad [0, 1]\qquad \cdots\qquad [0.25, 0.5]}_{attribute`s\ values\ ranges\ once\ normalized}$$

$$\underbrace{[0, 1]\qquad \cdots\qquad [0, 1]\qquad \cdots\qquad [0, 1]}_{attribute`s\ domains\ once\ normalized}$$

$\} :: 1$

End.

It is now possible to represent the normalized case repository in a graphic form, showing each case in the Cartesian plane in terms of its *QoI* and *DoC* (Fig. 2). Furthermore, the retrieval stage can be improved by reducing the search space, using data mining techniques, like clustering, in order to obtain different groups to identify the one(s) that are more closed to the *New Case*, which is represented as a square in Fig. 2.

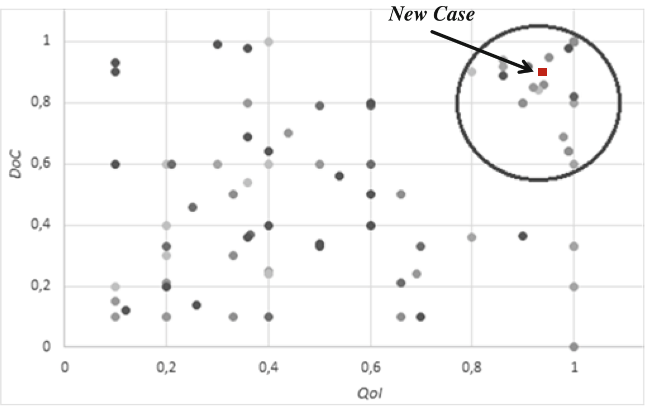


Fig. 2. A case's set split into clusters.

4 Case Based Reasoning

CBR methodology for problem solving stands for an act of finding and justifying the solution to a given problem based on the consideration of similar past ones, by reprocessing and/or adapting their data or knowledge [12]. In *CBR – the cases –* are stored in a *Case-Base*, 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. In fact, it is an iterative process since the solution must be tested and adapted while the result of applying that solution is inconclusive. In the final stage the case is learned and the knowledge base is updated with the new case [12, 13]. Despite promising results, the current CBR systems are neither complete nor adaptable enough for all domains. In some cases, the user is required to follow the similarity method defined by the system, even if it does not fit into their needs [25]. Moreover, other problems may be highlighted. On the one hand, the existent CBR systems have limitations related to the capability of dealing with unknown, incomplete and self-contradictory information. On the other hand, an important feature that often is discarded is the ability to compare strings. In some domains strings are important to describe a situation, a problem or even an event [12, 25].

Contrasting with other problem solving methodologies (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 new 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] but it also allows for the handling of incomplete, unknown, or even self-contradictory data or knowledge [25]. The classic CBR cycle was changed in order to include a normalization phase aiming to enhance the retrieve process (Fig. 3). The Case-Base will be given in terms of triples that follow the pattern:

$$Case = \{ < Raw_{case}, Normalized_{case}, Description_{case} > \}$$

where Raw_{case} and $Normalized_{case}$ stand for themselves, and $Description_{case}$ is made on a set of strings or even in free text, which may be analyzed with string similarity algorithms.

When confronted with a new case, (Fig. 4), the system is able to retrieve all cases that meet such a structure and optimize such a population, i.e., it considers the attributes DoC's value of each case or of their optimized counterparts when analysing similarities among them. Thus, under the occurrence of a new case, the goal is to find similar cases in the CaseBase. Having this in mind, the reductive algorithm given in [24] is applied to the new case, with the results:

$$\underbrace{stroke_{new}((1,1),(1,1),(1,1),(1,1),(1,1),(1,1),(1,1),(1,0.87)) :: 1 :: 0.98}_{new\,case}$$

After the normalization process, the new case is compared with every retrieved case from the cluster using a similarity function, *sim*, given in terms of the average of the modulus of the arithmetic difference between the arguments of the each case of the retrieved cluster and those of their counterparts in the problem (once *Description* stands for free text, its analysis is excluded at this stage). Thus, one may get:

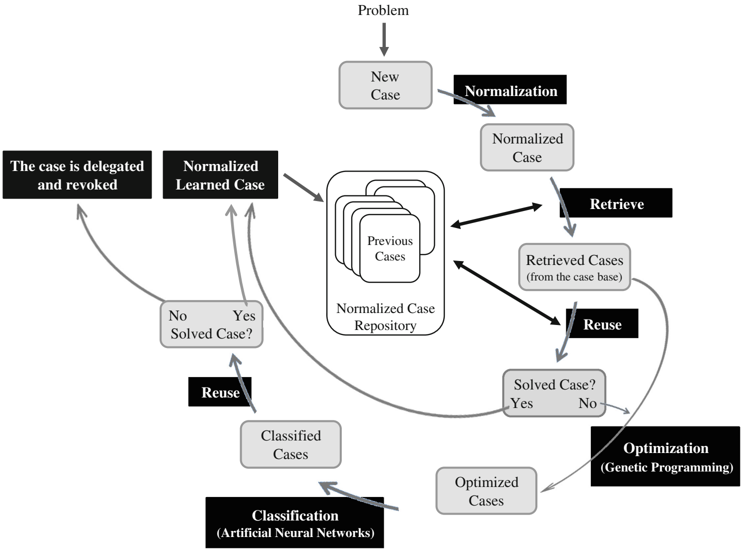


Fig. 3. The extended CBR cycle [25].

$$\underbrace{\begin{array}{l} stroke_1\big((1,1),(1,1),\ (1,0),\ (1,1),(1,1),(1,1),(1,\ 1),(1,\ 0.97)\big) :: 1 :: 0.89 \\ stroke_2\big((1,1),(1,1),(1,1),(1,0),\ (1,1),(1,1),,(1,0.8),(1,0.92)\big) :: 1 :: 0.84 \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ stroke_j\big((1,1),(1,1),\ (1,0),\ (1,1),(1,1),(1,1),(1,\ 1),(1,\ 0.96)\big) :: 1 :: 0.87 \end{array}}_{normalized\ cases\ from\ retrieved\ cluster}$$

$$stroke_{new\to 1}^{DoC} = \frac{\|1-1\|+\|1-1\|+\|1-0\|+\|1-1\|+\|1-1\|+\ \|1-1\|+\|1-1\|+\|0.87-0.97\|}{8} = 0.14$$

where $stroke_{new\to 1}^{DoC}$ denotes the dissimilarities between $stroke_{new}^{DoC}$ and the $stroke_{1}^{DoC}$. It was assumed that every attribute has equal weight. Thus, the similarity for $stroke_{new\to 1}^{DoC}$ is $1 - 0.14 = 0.86$. With respect to *QoI* the procedure is similar returning $stroke_{new\to 1}^{QoI} = 1$.

Stroke (Predisposing)									
#	Age	PSE	SBP	Cholesterol (LDL)	Cholesterol (HDL)	Triglycerides	Lifestyle Habits	Risk Factors	Description
new	58	1	115	102	67	149	6	[0, 2]	*new description*

Fig. 4. The *new case* characteristics and description.

Descriptions will be compared using String Similarity Algorithms in order to get a similarity measure between them. It is then necessary to compare the description of the new case with the descriptions of the cases stored in the repository (in this study the strategy used was the Dice Coefficient one [26]):

$$stroke_{new\to 1}^{Description} = 0.78$$

With these values we are able to get the final similarity function, *sim*:

$$sim\_stroke_{new\to 1} = \frac{0.86 + 1 + 0.78}{3} = 0.88$$

These procedures should be applied to the remaining cases of the retrieved cluster in order to obtain the most similar ones, which may stand for the possible solutions to the problem.

5 Conclusions

In order to target the CBR cycle theoretically and practically, the Decision Support System presented in this work to assess stroke predisposing risk, is centred on a formal framework based on Logic Programming for Knowledge Representation and Reasoning, complemented with a CBR approach to computing that caters for the handling of incomplete, unknown, or even self-contradictory data or knowledge. Under this approach the cases' retrieval and optimization phases were heightened and the time spent on those tasks shortened in 18.7 %, when compared with existing systems, being its accuracy around 89 %. The proposed method also allows for the analysis of free text attributes using *String Similarities Algorithms*, which fulfils a gap that is present in almost all *CBR* software tools. Additionally, under this approach, the user may define the weights of the cases' attributes on-the-fly, letting him/her to choose the most appropriate strategy to address the problem (i.e., it gives the user the possibility to narrow the search space for similar cases at runtime).

Acknowledgments. This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013.

References

1. Go, A.S., Mozaffarian, D., Roger, V.L., Benjamin, E.J., Berry, J.D., Blaha, M.J., Dai, S., Ford, E.S., Fox, C.S., Franco, S., Fullerton, H.J., Gillespie, C., Hailpern, S.M., Heit, J.A., Howard, V.J., Huffman, M.D., Judd, S.E., Kissela, B.M., Kittner, S.J., Lackland, D.T., Lichtman, J.H., Lisabeth, L.D., Mackey, R.H., Magid, D.J., Marcus, G.M., Marelli, A., Matchar, D.B., McGuire, D.K., Mohler 3rd, E.R., Moy, C.S., Mussolino, M.E., Neumar, R.W., Nichol, G., Pandey, D.K., Paynter, N.P., Reeves, M.J., Sorlie, P.D., Stein, J., Towfighi, A., Turan, T.N., Virani, S.S., Wong, N.D., Woo, D., Turner, M.B.: On behalf of the american heart association statistics committee and stroke statistics subcommittee: heart disease and stroke statistics — 2014 update: a report from the american heart association. Circulation **129**, e28–e292 (2014)
2. Lindgren, A.: Risk factors. In: Norrving, B. (ed.) Oxford Textbook of Stroke and Cerebrovascular Disease, pp. 9–18. Oxford University Press, Oxford (2014)

3. Shah, R.S., Cole, J.W.: Smoking and stroke: the more you smoke the more you stroke. Expert Rev. Cardiovasc. Ther. **8**, 917–932 (2010)
4. Hopper, I., Billah, B., Skiba, M., Krum, H.: Prevention of diabetes and reduction in major cardiovascular events in studies of subjects with prediabetes: meta-analysis of randomised controlled clinical trials. Eur. J. Cardiovasc. Prev. Rehabil. **18**, 813–823 (2011)
5. Khoury, J.C., Kleindorfer, D., Alwell, K., Moomaw, C.J., Woo, D., Adeoye, O., Flaherty, M.L., Khatri, P., Ferioli, S., Broderick, J.P., Kissela, B.M.: Diabetes mellitus: a risk factor for ischemic stroke in a large biracial population. Stroke **44**, 1500–1504 (2013)
6. Amarenco, P., Labreuche, J., Touboul, P.: High-density lipoprotein-cholesterol and risk of stroke and carotid atherosclerosis: a systematic review. Atherosclerosis **196**, 489–496 (2008)
7. Zhang, Y., Tuomilehto, J., Jousilahti, P., Wang, Y., Antikainen, R., Hu, G.: Total and high-density lipoprotein cholesterol and stroke risk. Stroke **43**, 1768–1774 (2012)
8. Grau, A.J., Barth, C., Geletneky, B., Ling, P., Palm, F., Lichy, C., Becher, H., Buggle, F.: Association between recent sports activity, sports activity in young adulthood, and stroke. Stroke **40**, 426–431 (2009)
9. McDonnell, M.N., Hillier, S.L., Hooker, S.P., Le, A., Judd, S.E., Howard, V.J.: Physical activity frequency and risk of incident stroke in a national US study of blacks and whites. Stroke **44**, 2519–2524 (2013)
10. Sealy-Jefferson, S., Wing, J.J., Sánchez, B.N., Brown, D.L., Meurer, W.J., Smith, M.A., Morgenstern, L.B., Lisabeth, L.D.: Age- and ethnic-specific sex differences in stroke risk. Gend. Med. **9**, 121–128 (2012)
11. Kissela, B.M., Khoury, J.C., Alwell, K., Moomaw, C.J., Woo, D., Adeoye, O., Flaherty, M. L., Khatri, P., Ferioli, S., De Los Rios La Rosa, F., Broderick, J.P., Kleindorfer, D.O.: Age at stroke: temporal trends in stroke incidence in a large, biracial population. Neurology **79**, 1781–1787 (2012)
12. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Communications **7**, 39–59 (1994)
13. Balke, T., Novais, P., Andrade, F., Eymann, T.: From real-world regulations to concrete norms for software agents – a case-based reasoning approach. In: Poblet, M., Schild, U., Zeleznikow, J. (eds.) Proceedings of the Workshop on Legal and Negotiation Decision Support Systems (LDSS 2009), pp. 13–28. Huygens Editorial, Barcelona (2009)
14. Carneiro, D., Novais, P., Andrade, F., Zeleznikow, J., Neves, J.: Using case-based reasoning to support alternative dispute resolution. In: de Leon F. de Carvalho, A.P., Rodríguez-González, S., De Paz Santana, J.F., Corchado Rodríguez, J.M. (eds.) Distributed Computing and Artificial Intelligence. AISC, vol. 79, pp. 123–130. Springer, Heidelberg (2010)
15. Carneiro, D., Novais, P., Andrade, F., Zeleznikow, J., Neves, J.: Using case-based reasoning and principled negotiation to provide decision support for dispute resolution. Knowl. Inf. Syst. **36**, 789–826 (2013)
16. Guessoum, S., Laskri, M.T., Lieber, J.: Respidiag: a case-based reasoning system for the diagnosis of chronic obstructive pulmonary disease. Expert Syst. Appl. **41**, 267–273 (2014)
17. Ping, X.-O., Tseng, Y.-J., Lin, Y.-P., Chiu, H.-J., Feipei Lai, F., Liang, J.-D., Huang, G.-T., Yang, P.-M.: A multiple measurements case-based reasoning method for predicting recurrent status of liver cancer patients. Comput. Ind. **69**, 12–21 (2015)
18. Kakas, A., Kowalski, R., Toni, F.: The role of abduction in logic programming. In: Gabbay, D., Hogger, C., Robinson, I. (eds.) Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 5, pp. 235–324. Oxford University Press, Oxford (1998)
19. Pereira, L.M., Anh, H.T.: Evolution prospection. In: Nakamatsu, K., P-W, G., Jain, L.C., Howlett, R.J. (eds.) New Advances in Intelligent Decision Technologies. SCI, vol. 199, pp. 51–63. Springer, Heidelberg (2009)

20. Neves, J.: A logic interpreter to handle time and negation in logic databases. In: Muller, R., Pottmyer, J. (eds.) Proceedings of the 1984 annual conference of the ACM on the 5th Generation Challenge, pp. 50–54. Association for Computing Machinery, New York (1984)
21. Neves, J., Machado, J., Analide, C., Abelha, A., Brito, L.: The halt condition in genetic programming. In: Neves, J., Santos, M.F., Machado, J.M. (eds.) EPIA 2007. LNCS (LNAI), vol. 4874, pp. 160–169. Springer, Heidelberg (2007)
22. Lucas, P.: Quality checking of medical guidelines through logical abduction. In: Coenen, F., Preece, A., Mackintosh, A. (eds.) Research and Developments in Intelligent Systems XX, pp. 309–321. Springer, London (2003)
23. Machado J., Abelha A., Novais P., Neves J., Neves J.: Quality of service in healthcare units. In Bertelle, C., Ayesh, A. (eds.) Proceedings of the ESM 2008, pp. 291–298. Eurosis – ETI Publication, Ghent (2008)
24. Fernandes, F., Vicente, H., Abelha, A., Machado, J., Novais, P., Neves J.: Artificial neural networks in diabetes control. In: Proceedings of the 2015 Science and Information Conference (SAI 2015), pp. 362–370. IEEE Edition (2015)
25. Neves J., Vicente, H.: A quantum approach to case-based reasoning (in Preparation)
26. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**, 297–302 (1945)