Waiting Time Screening in Healthcare

José Neves[1(✉)], Henrique Vicente[1,2], Marisa Esteves[3], Filipa Ferraz[3], António Abelha[1], José Machado[1], Joana Machado[4], and João Neves[5]

[1] Centro Algoritmi, Universidade do Minho, Braga, Portugal
{jneves,abelha,jmac}@di.uminho.pt
[2] Departamento de Química, Escola de Ciências e Tecnologia, Centro de Química de Évora, Universidade de Évora, Évora, Portugal
hvicente@uevora.pt
[3] Departamento de Informática, Universidade do Minho, Braga, Portugal
marisa.araujo.esteves@gmail.com, filipatferraz@gmail.com
[4] Farmácia de Lamaçães, Braga, Portugal
joana.mmachado@gmail.com
[5] Mediclinic Arabian Ranches, PO Box 282602, Dubai, United Arab Emirates
joaocpneves@gmail.com

Abstract. In Medical Imaging (MI), various technologies can be used to monitor the human body for diagnosing, monitoring or treating disease. Each type of technology provides different information about the body area that is being investigated or treated for a possible illness, injury or effectiveness of a medical treatment. Routine screening has identified malfunction detection in many otherwise asymptomatic patient images such as computed tomography or magnetic resonance. Studies have shown that, compared to patients whose disease was symptomatic (i.e., self-recognizing), screen-detected diseases may have more favorable clinicopathological features, leading to better prognosis and better outcome. This paper aims to assess the issue of health care wait screening. It deviates from a decision support system that evaluates the waiting times in diagnostic MI based on operational data from various information systems. Last but not least, one's assumptions may have an important impact in determining the usefulness of routine laboratory testing at admission.

Keywords: Waiting time screening · Logic programming
Case-based reasoning

1 Introduction

The characterization of health activities in terms of time-screening theory is a very recent trend in the field of research, i.e., a progressive learning experience, compensated by the occasional satisfaction of discovery. In fact, time-screening has often been defined, with meanings ranging from "not easy" to "persistent." On the other hand, technological advances are rapidly increasing interoperability, i.e., the ability to communicate and integrate information from heterogeneous sources or services. In fact,

J. J. Jung et al. (Eds.): BDTA 2017, LNICST 248, pp. 124–131, 2018.
https://doi.org/10.1007/978-3-319-98752-1_14

a variety of imaging techniques can be used to diagnose or treat diseases such as X-rays, Computed Tomography (CT), Magnetic Resonance Imaging, Positron Emission Tomography, Nuclear Medicine. Thus, a large data set extracted from various information systems such as the Radiology Information System, the Image Archiving and Communication System and the Electronic Medical Record is acquired and processed [1–3]. It is undeniable that a proactive strategy is needed to solve such a problem delay, which must take all of these factors into account. It is from this point of view that the problem is addressed. The basis of integrated care is to be understood as a patient interacting with a prepared, proactive and multidisciplinary setting. In this work, the approach focuses on estimating the waiting time in a Case Based Reasoning (CBR) approach to problem solving [4–6].

2 Knowledge Representation and Reasoning (KRR)

One aims at the understanding of the information's complexity and the associated inference mechanisms. Indeed, automated reasoning capabilities enables a system to fill in the blanks when one is dealing with incomplete information, where data gaps are common. In this study, a data item is to be understood as find something smaller inside when taking anything apart, i.e., it is mostly formed from different elements, namely the *Interval Ends* where their values may be situated, the *Quality-of-Information* (*QoI*) they carry, and the *Degree-of-Confidence* (*DoC*) put on the fact that their values are inside the intervals just referred to above. These are just three of over an endless element's number. Undeniably, one can make virtually anything one may think of by joining different elements together or, in other words, viz.

- What happens when one splits a data item? The broken **pieces** become data item for another element, a process that may be endless; and
- Can a data item be broken down? Basically, it is the smallest possible part of an element that still remains the element.

Therefore, the proposed approach to this issue, put in terms of the logical programs that elicit the universe of discourse, will be set as productions of the type, viz

$$predicate_{1 \le i \le n} - \bigcap_{1 \le j \le m} clause_j(([A_{x_1}, B_{x_1}](QoI_{x_1}, DoC_{x_1})), \cdots$$
$$\cdots, ([A_{x_m}, B_{x_m}](QoI_{x_m}, DoC_{x_m}))) :: QoI_j :: DoC_j$$

where n, $\cap$, m and A_{x_m}, B_{x_m} stand for the cardinality of the predicates' set, conjunction, predicate's extension, and the interval ends where the predicates attributes values may be situated, respectively. The metrics $[A_{x_m}, B_{x_m}]$, QoI and DoC show the way to data item dissection, i.e., a data item is to be understood as the data's atomic structure. It consists of identifying not only all the sub items that are thought to make up an data item, but also to investigate the rules that oversee them, i.e., how $[A_{x_m}, B_{x_m}]$, QoI_{x_m}, and DoC_{x_m} are kept together and how much added value is created [7–14].

3 Case Study

A database was set to create an intelligent system for the planning process of *Waiting Time Screening in Healthcare*. The knowledge database is composed by a set of predicate's extensions (Fig. 1). Some incomplete or default data are present under this scenario (for instance, the *type* in case 1 is unknown, and symbolized as ⊥).

	Waiting Time Screening								
Attributes of the Feature Vector:	#	Age	Gender	Date (days)	Modality	Type	Priority	Ordering Speciality	Description
Feature Vector Attributes' Values:	1	37	0	17	1	⊥	0	27	Description 1
	2	77	1	45	1	55	1	92	Description 2
	…	…	…	…	…	…	…	…	…
	149	68	0	21	0	83	0	43	Description 149
Feature Vector Domains:		[17, 95]	[0, 1]	[1, 366]	[0, 1]	[0, 113]	[0, 1]	[0, 192]	

Fig. 1. Healthcare table or relation.

The table has several columns such as *Gender*, *Modality* and *Priority* of *Waiting Time Screening*. The table rows have been inserted with one (1) or 0 (zero) standing for, respectively, *Male/Female*, *CT/MRI* and *Urgent/Routine*.

It is now possible to define the predicate ***waiting time screening*** (*wts*) whose extension stands for the objective function with respect to the problem under analysis, viz.

$$wts : Age, G_{ender}, Date, Mod_{ality}, T_{ype}, Prio_{rity}, O_{rdering}S_{pecialty} \rightarrow \{0, 1\}$$

in which the truth values *true* and *false* are expressed by 1 (one) and 0 (zero), respectively. Considering the feature vector (Age = 44, G_{ender} = 1, $Date$ = [30, 45], Mod_{ality} = 0, T_{ype} = 69, $Prio_{rity}$ = 1, $O_{rdering}$ $S_{pecialty}$ = ⊥), one may have, viz.

$$\{$$

$$\neg wts\left(\left((A_{Age}, B_{Age})(QoI_{Age}, DoC_{Age})\right), \cdots, \left((A_{OS}, B_{OS})(QoI_{OS}, DoC_{OS})\right)\right)$$

$$\leftarrow wts\left(\left((A_{Age}, B_{Age})(QoI_{Age}, DoC_{Age})\right), \cdots, \left((A_{OS}, B_{OS})(QoI_{OS}, DoC_{OS})\right)\right)$$

$$wts\underbrace{\left(\left((0.35, 0.35)(1, 1)\right), \cdots, \left((0, 1)(1, 0)\right)\right)}_{\substack{attribute`s\ values\ ranges\ once\ normalized\ and \\ respective\ QoI\ and\ DoC\ values}} :: 1 :: 0.86$$

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

$$\} :: 1$$

Program 1. The *Logic Program* with respect to the feature vector referred to above.

4 Case Based Reasoning (CBR)

The *CBR* cycle used in this work was proposed by Neves *et al*., [12, 15] (Fig. 2), with the ability to deal with incomplete or unknown information. Artificial Neural Networks (ANNs) [16] were used in the optimization stage. The value ranges boundaries of the attribute, the *DoCs* and *QoIs*, are the inputs of the *ANN*. The output not only provides the case assessment, but also a confidence measure that deals with such a categorization (Fig. 3).

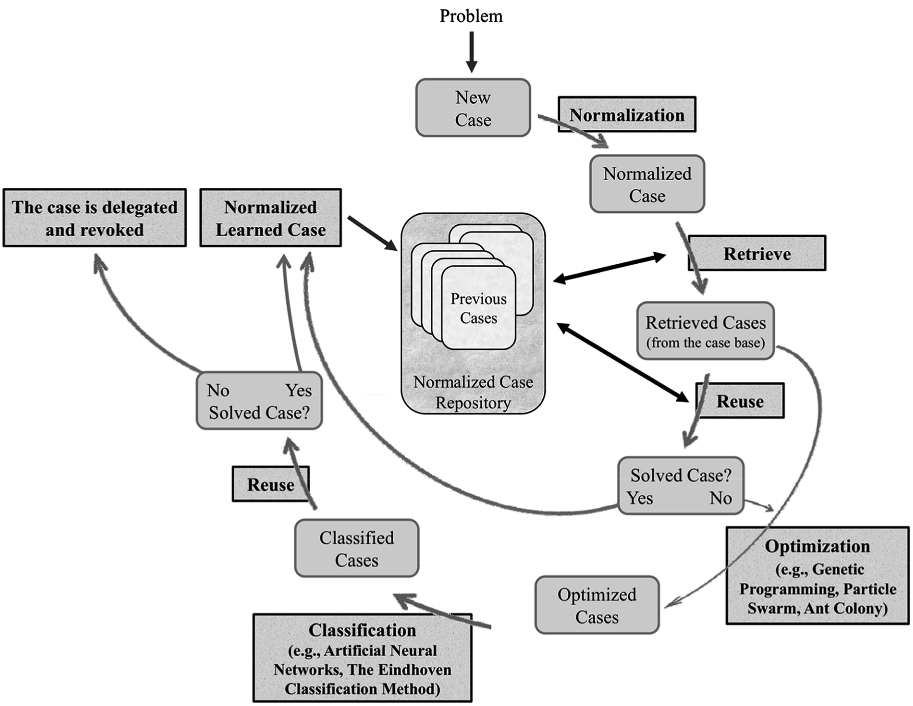

Fig. 2. An extended view of the canonical *CBR* cycle [12, 15].

When faced with a *new case*, for example the one that presents feature vector $Age = 56$, $Gender = 0$, $Date = 21$, $Mod_{ality} = \perp$, $Type = 42$, $Prio_{rity} = 0$, $O_{rdering}\ S_{pecialty} = 112$, one may have, viz.

$$wts_{newcase}\ \underbrace{(((0.5, 0.5)(1, 1)), \cdots, ((0.58, 0.58)(1, 1)))}_{\substack{attribute's\ values\ ranges\ once\ normalized \\ and\ respective\ QoI\ and\ DoC\ values}} :: 1 :: 0.85$$

leading to a retrieving of 42 cases [17], viz.

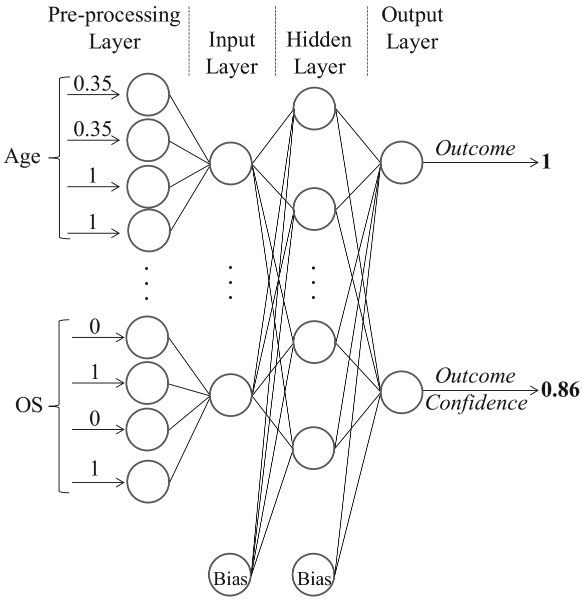

Fig. 3. The *ANNs* approach to case optimization.

$$retrieved_{case_1}(((0.43, 0.43)(1,1)), \cdots, ((0.86, 0.86)(1,1))) :: 1 :: 0.83$$
$$\vdots$$
$$retrieved_{case_{42}} \underbrace{((((0.51, 0.51)(1,1)), \cdots, ((0.57, 0.57)(1,1)))}_{normalized\ cases\ that\ make\ the\ retrieved\ cluster} :: 1 :: 0.84$$

The *new case* and the *retrieved ones* are compared using a similarity function, *sim*. This function is set as follows, viz.

$$sim^{DoC}_{newcase \to 1} = 1 - \frac{\|1-1\| + \cdots + \|1-1\|}{7} = 1 - 0.18 = 0.82$$

where $sim^{DoC}_{newcase \to 1}$ stands for the *similarity* with respect to *DoC*, between the *new case* and the retrieved ones (in this example *retrieved case$_1$*). A similar process was considered in order to evaluate the *similarity*, in terms of *QoI*, between the *new case* and the *retrieved case$_1$*, returning $sim^{QoI}_{newcase \to 1} = 1$. The *general similarity*, $sim^{QoI,DoC}_{newcase \to 1}$, is the product of the above metrics above, viz.

$$sim^{QoI,DoC}_{newcase \to 1} = 1 \times 0.82 = 0.82$$

This method was extended to all the remaining cases leading to the most similar case, i.e., the potential problem solutions. The coincidence matrix for the *CBR* model is shown in Table 1. It shows that the *CBR* model classifies properly 133 of a total of 149 cases, being the model accuracy 89.2%. In terms of the well known statistical metrics

Table 1. The coincidence matrix

Output	Model output	
	True (1)	False (0)
True (1)	92	10
False (0)	6	41

such as sensitivity and specificity, the results were 90.2%, and 87.2% respectively. The *ROC* curve is shown in Fig. 4. The area under the curve is 0.89. The performance metrics [18, 19] are close to 90% and suggest that the model has a good performance in predicting the waiting time in healthcare.

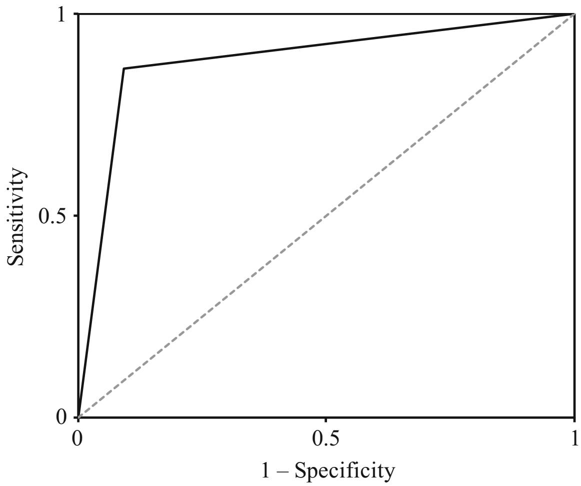

Fig. 4. The *ROC* curve.

5 Conclusions

This work begins with the development of an intelligent system for assessing the latency in providing diagnostic medical services, based on a formal framework based on Logic Programming for knowledge representation and reclaiming the CBR approach to problem solving. The knowledge presentation and enforcement apparatus presented above is very versatile and able to cover all possible data types, namely incomplete, unknown or even self-contradictory data or information. Future work should include data from various healthcare facilities (public, semi-public and private) from different regions of Portugal. On the other hand, different string similarity strategies will be considered, and their analysis complexity enumerated. On the other hand, given the cost of this relatively low score, these results have important implications for the doctor's office and cost-benefit analyzes that will be further evaluated to better determine the current benefit of routine laboratory testing on admission.

Acknowledgments. 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

1. Nuti, S., Vainieri, M.: Managing waiting times in diagnostic medical imaging. BMJ Open **2**, e001255 (2012)
2. McEnery, K.W.: Radiology information systems and electronic medical records. In: IT Reference Guide for the Practicing Radiologist, pp. 1–14. American College of Radiology, USA (2013)
3. Fotiadou, A.: Choosing and visualizing waiting time indicators in diagnostic medical imaging department for different purposes and audiences. Master's thesis in Health Informatics, Karolinska Institutet, Sweden (2013)
4. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Commun. **7**, 39–59 (1994)
5. Richter, M.M., Weber, R.O.: Case-Based Reasoning: A Textbook. Springer, Berlin (2013)
6. 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)
7. 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)
8. Neves, J., Machado, J., Analide, C., Abelha, A., Brito, L.: The halt condition in genetic programming. In: Neves, J., Santos, M.F., Machado, J. (eds.) Progress in Artificial Intelligence. LNAI, vol. 4874, pp. 160–169. Springer, Berlin (2007)
9. 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)
10. Pereira, L., Anh, H.: Evolution prospection. In: Nakamatsu, K. (ed.) New Advances in Intelligent Decision Technologies, Studies in Computational Intelligence, vol. 199, pp. 51–64. Springer, Berlin (2009)
11. 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)
12. Silva, A., Vicente, H., Abelha, A., Santos, M.F., Machado, J., Neves, J., Neves, J.: Length of stay in intensive care units – a case base evaluation. In: Fujita, H., Papadopoulos, G.A. (eds.) New Trends in Software Methodologies, Tools and Techniques, Frontiers in Artificial Intelligence and Applications, vol. 286, pp. 191–202. IOS Press, Amsterdam (2016)
13. 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)

14. Turner, M., Fauconnier, G.: Conceptual integration and formal expression. J. Metaphor Symbolic Act. **10**, 183–204 (1995)
15. Vilhena, J., Vicente, H., Martins, M.R., Grañeda, J., Caldeira, F., Gusmão, R., Neves, J., Neves, J.: A case-based reasoning view of thrombophilia risk. J. Biomed. Inf. **62**, 265–275 (2016)
16. Haykin, S.: Neural Networks and Learning Machines. Pearson Education, New Jersey (2009)
17. Figueiredo, M., Esteves, L., Neves, J., Vicente, H.: 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. Chem. Educ. Res. Pract. **17**, 156–171 (2016)
18. Florkowski, C.: Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin. Biochem. Rev. **29**(Suppl 1), S83–S87 (2008)
19. Hajian-Tilaki, K.: Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J. Intern. Med. **4**, 627–635 (2013)