

Combining Case and Rule Based Reasoning for the Diagnosis and Therapy of Chronic Obstructive Pulmonary Disease

Souad Guessoum¹, M.Tayeb Laskri², Hayet djellali³ and M. Tarek Khadir⁴

^{1,3}LRS Laboratory, ²LRI Laboratory, ⁴LabGed Laboratory
Computer Science Department, Badji Mokhtar University, BP 12, Annaba, Algeria
souadguessoum@yahoo.fr, laskri@univ-annaba.org, hayetdjellali@yahoo.fr,
khadir@labged.net

Abstract

In this paper, we present a work consisting of conceiving and realizing a decision making support system for the clinical decision. Being based on Case-Based Reasoning (CBR) the system will highlight the impact of this technique in the clinical decision making domain. The experience has its importance in medicine, and the nature of CBR has motivated us to use it for the diagnosis and the therapy of a very grave respiratory disease bound to tobacco, for instance the Chronic Obstructive Pulmonary Disease (COPD). Our approach proposes to apply a CBR process by modeling the adaptation phase with an independent component: an expert system using forward chaining. For the retrieval phase, some heuristics have been traced for estimating the similarity on missing data and symbolic descriptors. Preliminary experimentations show that the expert system models well the adaptation task and gives satisfactory answers to the user. For the retrieval phase, our heuristic functions have proved a good impact on the results.

Keywords: Case-Based Reasoning; CBR; Rule-Based Reasoning; RBR; Diagnosis of the COPD; Medical Decision-making support System

1. Introduction

Artificial Intelligence (AI) and medicine maintain strong relations since more than three decades. Indeed medicine always constituted an excellent field of experiment to test and estimate various paradigms of AI. These experiments are very important for AI as for medicine which takes widely advantage of it by obtaining Decision-Making Support System for the diagnosis, interpretation of radiological images, training and education. These systems return an important impact on progresses of researches in AI, because they allow researchers to test the validity of their approaches [1], to reveal their gaps, to bring corrections and thus to evolve better.

The clinical reasoning is also source of inspiration in AI. It contributes widely in the evolution of techniques concerning the reproduction of the human reasoning on machines. For this reproduction multiple approaches of AI were exploited, particularly in the realization of decision-making support systems. The demand on these systems is important especially in domains where knowledge and experiences evolve quickly [2], what is the case in the medical domain.

The literature can show many medical CBR systems which have done successfully their tasks [3, 4, 5, 6, 7, 8, 9, 10, 11]. Their success is due to the fact that CBR is very similar to clinical reasoning. Indeed, during his activity, the doctor often appeals to his memory of experiences crossed in the exercise, to look for any resemblance between the former situations and the new one. Such resemblance (if it exists) can help enormously in the

resolution of the new problem, in term of making the most precise diagnosis, or proposing the most effective treatment.

CBR allows very well this recourse to these acquired experiences for solving new problems, and that, has motivated us to use it in our medical application. This work consists of conceiving, developing and evaluating a decision-making support system for the diagnosis of COPD: very dangerous disease caused by tobacco [12] and two others diagnoses bound to it.

We are cooperating with the service of respiratory diseases of Dorban Hospital (Annaba, Algeria) and we are working with specialist doctors for giving a system that can help young clinicians in the diagnosis of COPD. The system will gather experiences of many doctors and so it can give a big help for future pulmonologists.

Being based on the technique of Case-Based Reasoning, the system will allow us to estimate the impact which brings this technique in the detection of COPD. In order to give a better performance to the system, we propose an integration of Rule-Based Reasoning system for conceiving the adaptation task which is the most delicate and the most complicated phase in the CBR cycle. Rules are typically used for the adaptation in CBR systems and in the literature we can find many contributions to conceive and perform this phase by integrating other paradigms of AI [13, 14, 15].

Our approach will include a special and independent module; an Expert System (ES) with production rules based reasoning and forward chaining. Note that expert systems have proved widely their power in the modeling of the human reasoning that can be expressed by rules. This motivated us to use and evaluate it for the adaptation phase of our CBR system.

The retrieval phase is enriched by a work concerning the estimation of similarity between cases. In our system, some descriptors of the patient state are symbolic with discontinuous values. For each symbolic descriptor we have established a heuristic function to estimate similarity between its different values. For missing data we have proposed some ideas to estimate their similarity and compared their results.

This paper is organized in varied sections. Following a brief overview of CBR principle and its process is given. Section 3 is dedicated to the related works. The Proposed model is presented in section 4. After some preliminary results are presented in section 5 and the work is concluded in the last section.

2. Case-Based Reasoning

CBR is a recent technique based mainly on analogical reasoning. It was inspired by a real cognitive model observed in the human behavior, which consists of learning from past experiences to bring a solution to new problems. Indeed, it was demonstrated in psychology that in many situations, humans begins to resolve their problems by basing themselves more on their past experiences than on their knowledge gathered in the domain [16].

Thus the global idea of CBR consists of reusing the solution of former similar problem to solve a new one. All past experiences are gathered in the case base, where the case is a couple of descriptors of the problem and its solution. At these two basic elements comes to be sometimes added a third descriptor bringing an explanation, a justification or rather an estimation of the quality of the reuse of the case [17].

The main advantage of CBR is that it doesn't require an acquisition of "deep" knowledge of the considered domain. We don't need to know how the expert thinks for resolving the problem, because the knowledge consists just of establishing a description of a problem and its solution. It bases itself on a large number of problems solved in past instead of counting on explicit domain knowledge [18]. The second advantage is that the implementation of this technique is relatively simple and easy compared with other AI techniques. The third strength

of this mode of reasoning is that learning is easily realized in CBR system. It often consists simply of the insertion of new cases in the memory of the system [19].

2.1. CBR Process

CBR cycle is identified par Aamodt and Plaza in [20] as a process of four steps.

- *Retrieval phase*: It is the most important phase of the cycle. It consists of measuring the similarity of the current problem to previous problems gathered in the memory of the system [21] for retrieving one or more most similar case(s). This process is based on similarity metrics whose the choice is fundamental [22]. Indeed, for a given target case, the good selection of the most similar case depends strictly on these metrics. There are multiple measures which can be used in this phase, the most popular being the method of the k-nearest neighbors.
- *Reuse phase*: It is the most delicate step; it allows adapting the solution (if need be) of the found case to the target problem. Its difficulty is mainly due to the fact that heuristics and knowledge adaptation are strictly depending of application fields. For certain CBR systems particularly those dedicated to the diagnosis task this phase is ignored, and the process consists just of finding the most similar diagnosis.
- *Revise phase*: During which the adapted solution is presented to the user who will decide on his validity. In the affirmative case the last phase is begun.
- *Retain phase*: It consists of adding the new problem with its solution to the case base. And so the system learns of its experiences!

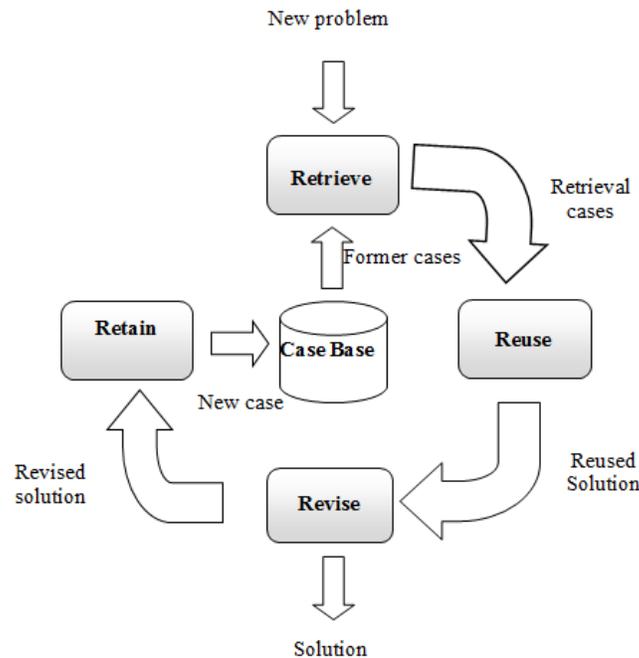


Figure 1. The CBR Cycle

2.2. CBR in Medicine

Rainer Schmidt presents in [23] an outline on medical CBR systems realized in the last twenty years. He distinguishes mainly three tendencies from it. We find at first Diagnosis Systems, such as: TeComMed [3] which supplies forecasts on influenza waves, CASEY [4] dedicated to cardiac disorders, MEDIC [5] which takes care of lung problems, FLORENCE [6] which gives at the same time diagnosis, forecast and prescription, MERSY [7] for the health of the workers of the rural region and FM-Ultranet [8] for the fetal deformations. The second tendency is drawn by Therapeutic Systems, such as: ICONS [23] which supplies advice (councils) in the antibiotic treatment, ISOR [9] is an interactive system helping doctors in the explanation and the interpretation of exceptional positions, where a theoretical therapeutic approach does not give the estimated result on a particular case. TA3-IVE [10] which supplies plans for treatments of cases of in-vitro fertilization, and CASIMIR [11] which proposes treatments for the breast cancer. Finally, we find systems specialized in the Interpretation of Medical Images, such as: ProtoISIS, MacRad [24] and SCINA [25].

3. Related Works

The Literature shows us many contributions merging CBR process with the integration of other AI techniques. Each combination has its own purpose and aims to perform a specific task in the CBR cycle. We can find for example in [26] a synthesis of some integrations of CBR with other reasoning modalities like RBR, model-based reasoning, constraints satisfaction problem solving, information retrieval and planning.

In [13] we find a work evoking the use of Genetic Algorithms (GA) during the reuse phase to formulating university timetables. GA are also used in [27] for the retrieval phase. Pedro A. et al propose in [14] the use of inference mechanisms of Description Logic to formalize an adaptation scheme based on substitutions. We find in [28] another work related to the exploitation of Fuzzy Logic to perform the revision stage of the CBR process. A hybrid system based on combination of CBR/Bayesian Networks (BN) is proposed in [29] for clinical decision support, where BN have been chosen for supporting the uncertainty in medical domain.

Wenqi and Barnden present in [15] a combination of CBR/RBR for diagnosing multiple medical disorder Cases. They propose an inductive learning method to generate diagnostic rules and to help the retrieval procedure in compositional CBR. In [30] we find another combination of CBR/RBR for designing nutritional menus where each reasoning is used for a specific task, to satisfy multiple constraints for CBR and to allow the introduction of new foods into menus for the RBR. The work [31] presents an approach of integration of rule and case based reasoning for internal bank audit. The approach includes two stages of reasoning: the first one is based on rule-based reasoning. Its result is passed to the second one which is based on case-based reasoning.

4. The Proposed Model

4.1. Field of Application: COPD

Chronic Obstructive Pulmonary Disease is an obstruction of lung branches, slowly progressive over several years. This obstruction is due to a chronic inflammation of bronchi, essentially bound to tobacco. Most of affected persons are smokers or ex-smokers, aged over 50 years. The World Organisation for the Health estimates at 2.74 millions the number of deaths due to the COPD in 2000. It is the 4th world cause of death today while it occupied only the 6th place in 1990 [32].

The COPD evolves according to four of gravity [33], it can be "little severe" where symptoms are variable, "moderate intensity" where symptoms are permanent, "severe" where the dyspnoea appears even in the repose accompanied by frequent exacerbations decreasing the quality of life or "very severe" where we observe the respiratory incapacity.

4.2. Architecture of the Proposed Model

RBR systems have proved their performance in the modeling of reasoning that can be explained and formalized by rules. In the goal of exploiting this performance, we propose CBR system which consists of integrating an RBR component in the reuse phase. This independent module will ensure the adaptation task which is usually realized by simple rules within the system. However in our approach these rules are contained in knowledge base of the RBR system. By drawing inferences from rules, this RBR module will adapt the diagnosis of retrieval case to the target case. A second task is given to this module: proposing an adequate treatment for the patient. The architecture and the cycle of the proposed model are given in Figure 2.

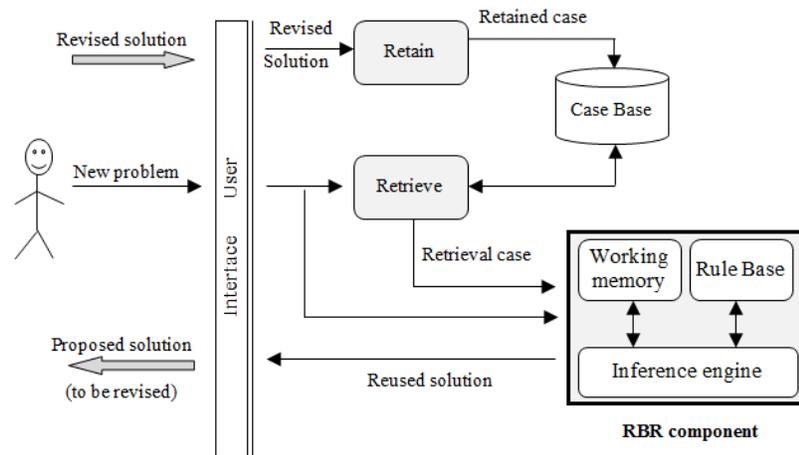


Figure 2. Architecture of the Proposed Model

4.3. Case Representation

The case base is structured in flat memory where former cases are stored with a representation <attribute-value> which seemed us the most appropriate. This representation has been established with association of several pulmonologists doctors and based on seventeen descriptors parameters representing the patient's state affected by COPD. Values of these descriptors are of heterogeneous types; numeric, symbolic or binary (yes for the appearance of the symptom and no for its absence). With the collaboration of doctors, forty real cases were collected of the archive of the pulmonology's service. These cases are jugged sufficient for initial representation of our three diagnoses.

4.4. Retrieval Phase

For retrieval phase we adopted the method of k nearest neighbors (knn) with k=1. The measure of similarity of the target case to the source case is made in two stages: first, local similarities have to be measured. These are similarities between each attribute of the target case and his source case correspondent. Thereafter, an aggregation is applied to consider all

local similarities. After discussion with doctors it turned out that descriptors considered to represent a case do not participate all with the same importance in the detection of COPD diagnosis. This importance is reflected in our system by coefficients representing weights of participation of each descriptor in the decision-making of the diagnosis.

4.4.1. Similarity Metrics for Numeric Descriptors: The distance of Manhattan is used to estimate the similarity between values of digital attributes. It is estimated by the following function,

$$sim(p_i(an), p_i(nc)) = 1 - distance(p_i(an), p_i(nc)) \quad (1)$$

In which,

an: the former case stored in the case base,

nc: the new case to be treated,

pi: descriptor i

$$distance(p_i(an), p_i(nc)) = \frac{p_i(an) - p_i(nc)}{max_dist\ i} \quad (2)$$

Where max_dist represents the distance between the superior border and lower border of the data interval of the descriptor pi.

4.4.2. Similarity Metrics for Binary Descriptors: When the descriptor has binary nature, we use the following similarity function,

$$sim(p_i(an), p_i(nc)) = \begin{cases} 0 & \text{si } p_i(an) \neq p_i(nc) \\ 1 & \text{si } p_i(an) = p_i(nc) \end{cases} \quad (3)$$

4.4.3. Similarity Metrics for Symbolic Descriptors: Generally when an attribute X is of symbolic type not binary, the similarity between two values v1, v2 of X is estimated by the binary function (2). In our work, six among the seventeen descriptors are symbolic and not binary. Each has between four and twenty possible values, and to estimate the similarity between two values of the same descriptor, the binary similarity is inadequate. In order to treat this point we have developed heuristic functions of similarity with the collaboration of doctors. These heuristics are able to give a similarity value for each pair of possible values of the same descriptor. For example, the value "dry" of cough attribute is estimated similar to the value "productive with mucous expectoration" with a rate of 20%, whereas the similarity between the values "productive with mucous expectoration" and "productive with mucopurulent expectoration" is estimated at 80%. For each of the six symbolic descriptors a function of similarity has been established.

4.4.4. Global Similarity: As indicated in the previous section, the method of knn is applied with k=1 and the function of similarity below is used for determining the similarity between two cases,

$$Similarity(ac, nc) = \frac{\sum_{i=1}^n w_i * sim(p_i(an), p_i(nc))}{\sum_{i=1}^n w_i} \quad (4)$$

Where wi: the weight attributed to the descriptor pi,

4.4.5. Missing Data: Sometimes the doctor has to decide about a diagnosis whereas he is in the lack of data. This is generally the case for example in emergency when the patient arrives

in respiratory crisis. Thus, our case base which was collected from archival files of patients can contain missing data which can be seen in a different attribute for each case. Note that this absence can be observed in the target case as well. In this situation we have thought to establish some similarity functions to apply when one of the attribute values (in the target or source case) of the same attribute is missing. We wanted to know what would be the best value to allocate to local similarity when we do not have one of the values of the attribute to compare to having local similarity. Would it be better to suppose null similarity for the attribute where value is missing for this case or to consider it with a similarity equal to 0.5 for example? The value 0.5 means that we can suppose that missing value can be similar at a rate 50% to the present value. Or rather, would it be better to ignore completely the attribute with its weight and to consider only the attributes present which may differ from one case to another?

In this work, we have implemented these possibilities for the estimation of similarity on missing values. We present a comparison of some preliminary results in Section 5.

4.5. Reuse Phase by RBR Module

This work consists of integrating an RBR component in the CBR system. This integration is realized in the most difficult phase of the CBR cycle: the reuse phase. The RBR component is an expert system with knowledge base and inference engine working in forward chaining.

At the end of the previous phase, a selection is made on the attributes of the highest coefficients of the target case. The values of these attributes will compose the facts of our RBR system and are forwarded to it. In addition to these values, the RBR component needs evidently to have the result of the retrieval phase. Its knowledge base contains rules allowing him to deduce from these data the modification to do on the retrieval diagnosis, and so to adapt it for the new case.

The three diagnoses of our work are “COPD stage X” where X can be I, II, III or IV, “Exacerbation of COPD due to infection” and “Exacerbation of COPD due to pneumothorax”. For the reuse phase we modeled these diagnoses in two parts as showed in Table 1:

Table 1. The Two Parts of Diagnoses

Part 1	Part 2
<i>COPD</i>	<i>Stage X</i>
<i>Exacerbation of COPD</i>	<i>due to infection</i>
<i>Exacerbation of COPD</i>	<i>due to pneumo-thorax</i>

Note that developing rules for the adaptation phase is a complicated task because the adaptation knowledge was difficult to codify. We have traced therefore, only the adaptation rules for the second part of diagnosis. That means that the first part of the retrieval diagnosis will stay without modification whereas the second part will be substituted by other information adequate to the new patient. For example for the diagnosis ‘COPD stage II’ we retain ‘COPD’ without changes and we substitute ‘stage II’ for example by “stage IV” according the symptoms of the new case.

In addition to the adaptation task, we have given to the RBR component a second task: to provide an adequate treatment to the new patient. Indeed, through a set of production rules and the adapted diagnosis the component can draw some inferences to deduce the treatment to propose. The output of this Reuse phase is so the couple (adapted diagnosis, proposed therapy).

4.6. Revise and Retain Phases

Once the diagnosis and the treatment are found, they are presented to the doctor who will proceed to the validation by a confirmation or a modification of the diagnosis. In the retain phase, the system proceeds to the addition of the new problem with its solutions (diagnosis) in the case base. The therapy does not make object to the addition to the case base.

5. Implementation and Results

The user interface of our system allows entering all symptoms observed at the patient. Seized values will build for us the descriptors of the new case to be treated. To simplify the entry of symptoms, for each symptom descriptor, all possible values are listed below, so the user simply clicks on the corresponding value to the patient. When the doctor ignores the value of a specific parameter, he must click on the symbol "?" mentioned at the end of possible values list of this parameter, which will cause a missing data in the new case. We conducted two experiments in the system for evaluate the impact of our ideas on similarity for missing data and of our heuristics functions for symbolic descriptors. The third experiment is done for evaluate results after adaptation given by the reuse phase which is modeled and implemented by RBR module. Results of these experiments are presented in following.

5.1. Missing Data

We have selected a set of 13 real cases presenting our three diagnoses and which will serve as sample test. To compare the three ideas on missing values, we have implemented them, and have removed some values of each test case, so as to make at least four absences in each one. In following "Diag1" means that the diagnosis is done with similarity supposed equal to 0,5 when one of the compared values is missing. "Diag2" means that similarity is supposed equal to 0 and in "Diag3" only present attributes are considered.

Table 2 shows the results of the retrieval phase for Diag1 before adaptation. Firstly, we can observe that it may be a strong similarity between two cases which have not the same diagnosis and that explain well the complication of the COPD diagnosis.

Table 2. Results of Diagnosis "Diag 1"

Case	Real Diagnosis	Diag 1	Sim
1	COPD Stage II	COPD Stage II	0,62
2	COPD Stage III	Ex. due to infection	0,81
3	COPD Stage III	COPD Stage IV	0,69
4	COPD Stage IV	Exacerbation of COPD due to infection	0,74
5	COPD Stage II	COPD Stage II	0,69
6	COPD Stage IV	COPD Stage IV	0,75
7	Exacerbation of COPD due to pneumo-thorax	Exacerbation of COPD due to infection	0,72
8	Exacerbation of COPD due to infection	Exacerbation of COPD due to infection	0,79
9	Exacerbation of COPD due to infection	Exacerbation of COPD due to infection	0,88
10	Exacerbation of COPD due to infection	Exacerbation of COPD due to infection	0,86
11	Exacerbation of COPD due to pneumo-thorax	Exacerbation of COPD due to infection	0,69
12	Exacerbation of COPD due to pneumo-thorax	Exacerbation of COPD due to infection	0,68
13	Exacerbation of COPD due to infection	Exacerbation of COPD due to infection	0,81

For each test we can have one of three possibilities for the result:

- Good: means that the result is consistent with the new case even before the adaptation, like for examples cases n °1, 6 and 13.
- Medium: means that the adaptation made from this result may lead to a diagnosis consistent with the new case. So the result of the search is accepted, like in cases n °3 and 7.
- Bad: means that adaptation rules that are available at the knowledge base will not permit to find the appropriate diagnosis to the new case from the result found in the retrieval phase, like for examples cases n °2 and 4.

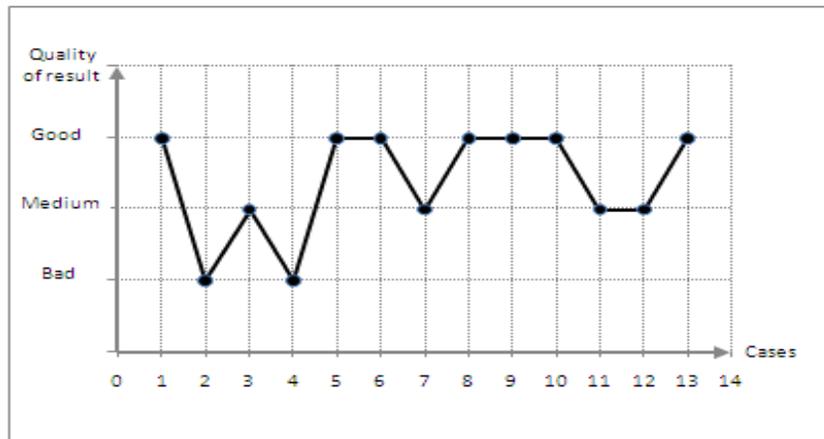


Figure 3. Evaluation of Diagnosis “Diag1”

We have done the same work for “Diag2” and “Diag3” which we present by the following figures.

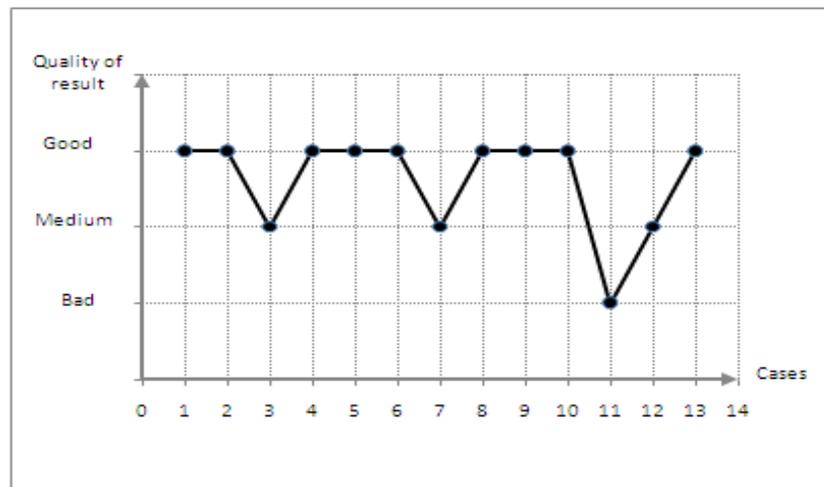


Figure 4. Evaluation of Diagnosis “Diag2”

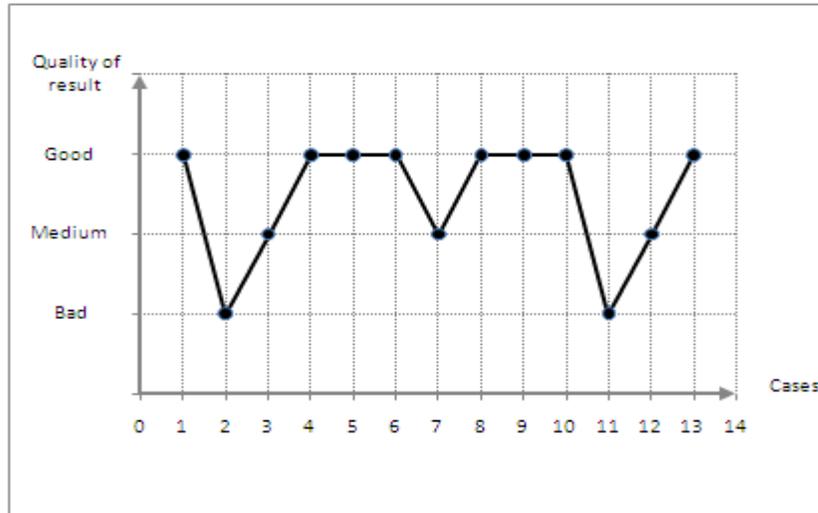


Figure 5. Evaluation of Diagnosis "Diag3"

Table 3 summarizes the results of the three ideas. We can observe that "Diag2" corresponding to null similarity for missing data with a consideration of its weight present the best results. That has encouraged us to retain it for our retrieval phase.

Table 3. Results Quality of the Three Diagnoses

	Good	Medium	Bad
Diag 1	54%	31%	15%
Diag 2	69%	23%	8%
Diag 3	62%	23%	15%

5.2. Symbolic Descriptors

As indicated in section 4, we have established six heuristics functions for estimating similarity between symbolic values of the same attribute. To show the impact of these functions on the system, we have done some tests based on eight cases.

For each case we have executed "Diag2" presented in the previous section considering null similarity for missing data with consideration of their weights and with binary similarity for symbolic attributes. Thereafter we have executed "Diag4" that consider null similarity for missing data with consideration of their weights (like "diag2") but with gradual similarity for symbolic attributes. Evidently the test is done without the adaptation.

- Cases 1, 2, 4, 7: in these cases, Diag2 and Diag4 have given a good result corresponding to the real diagnosis, but Diag4 brings an improvement in the similarity of 0.17, 0.11, 0.11, and 0.04 respectively.
- Cases 3, 8: in these cases, Diag2 and Diag4 gave the same diagnosis that can be adapted by the expert system to give the real diagnosis. But Diag4 has improved the similarity of 0.13, and 0.04.
- Case 5: Diag2 and Diag4 have given a false diagnosis that can't be adapted by the expert system. Diag4 increased the similarity so it downgraded the result but just of 0.04.

- Case 6: Diag2 has given a false result that can't be adapted by the reuse phase, whereas Diag4 found the same real diagnosis.
- These results proved that the heuristic functions established for estimating the similarity on the symbolic attributes have a good impact on the system.

In Tables 4 and 5, we used following abbreviations for diagnoses:

- "Ex. due to Pn.Thor" for "Exacerbation of COPD due to Pneumo-Thorax"
- "Ex. due to infect." for "Exacerbation of COPD due to Infection"

Table 4. Results of Diagnosis "Diag 4" Relatively to "Diag 2"

case	Real Diagnosis	Diag 2	Sim	Diag 4	Sim
1	COPD Stage II	COPD Stage II	0,51	COPD Stage II	0,68
2	COPD Stage III	COPD Stage III	0,73	COPD Stage III	0,84
3	COPD Stage III	COPD Stage II	0,53	COPD Stage II	0,66
4	COPD Stage IV	COPD Stage IV	0,67	COPD Stage IV	0,78
5	Ex. due to Pn.Thor.	COPD Stage IV	0,55	COPD Stage IV	0,59
6	Ex. due to infect.	COPD Stage III	0,67	Ex. due to infect.	0,77
7	Ex. due to infect.	Ex. due to infect.	0,86	Ex. due to infect.	0,90
8	Ex. due to Pn.Thor.	Ex. due to infect.	0,59	Ex. due to infect.	0,63

5.3. Reuse Phase by RBR Module

As indicated, the reuse phase of CBR process is modeled and implemented by an expert system whose knowledge base contains average 30 rules allowing to adapt the diagnosis of the previous phase to the new patient. These rules are established on some attributes having the highest weights in addition to the retrieval diagnosis. This component which works in forwarding chaining, plays in reality two roles in the same time: an adapter for the CBR system and a decision making system for the therapy of COPD because its set of rules can deduce not only the modification to do on the retrieval diagnosis but also the adequate therapy proposed to the doctor.

As indicated in section 4, the modification that can be done here is just on the second part of the diagnosis. In other words, if the retrieval phase gives 'COPD Stage X' where X can be I, II, III or IV, the adaptation will be done only on the stage of COPD. Whereas if the retrieval diagnosis gives 'Exacerbation of COPD due to Y' where Y can be 'infection' or 'pneumo-thorax' the adaptation will be on just the cause of Exacerbation. That means that the Expert System cannot go from for example 'COPD stage X' to 'Exacerbation of COPD due to Y' or the inverse.

Table 5 gives some results obtained after experimenting 21 cases presenting the three diagnoses. Column 2 contains the real diagnosis of the set of cases. Column 3 shows the results of the retrieval phase. Cells in gray indicate diagnoses that do not correspond to real patients diagnoses.

We can note that we have 13/21 good diagnoses, 5/21 medium and 3/21 bad results. Column 5 gives the results after the adaptation, where the good diagnoses stay without changes, the medium ones are adapted for giving the real diagnosis (in bold) and the bad ones cannot be modified for the reason explicated above.

In conclusion we can say that our adaptation modeled with an expert system brings a good impact on the system which gives 0.86% of real diagnoses.

Table 5. Results after Adaptation given by “Diag 5”

case	Real Diagnosis	Diag 4	sim	Diag 5
1	COPD Stage II	COPD Stage II	0,59	<i>COPD Stage II</i>
2	COPD Stage III	COPD Stage III	0,73	<i>COPD Stage III</i>
3	COPD Stage III	COPD Stage II	0,62	COPD Stage III
4	COPD Stage IV	COPD Stage IV	0,61	<i>COPD Stage IV</i>
5	COPD Stage II	COPD Stage II	0,66	<i>COPD Stage II</i>
6	COPD Stage IV	Ex.due to infection	0,75	Ex.due to infection
7	Ex. due to Pn.Thor.	Ex.due to infection	0,59	Ex. due to Pn.Thor.
8	Ex.due to infection	Ex.due to infection	0,61	<i>Ex.due to infection</i>
9	Ex.due to infection	Ex.due to infection	0,75	<i>Ex.due to infection</i>
10	Ex.due to infection	Ex.due to infection	0,77	<i>Ex.due to infection</i>
11	Ex. due to Pn.Thor..	Ex.due to infection	0,59	Ex. due to Pn.Thor.
12	Ex. due to Pn.Thor.	Ex.due to infection	0,59	Ex. due to Pn.Thor.
13	Ex.due to infection	Ex.due to infection	0,66	<i>Ex.due to infection</i>
14	COPD Stage II	BPCO Stade II	0,68	<i>COPD Stage II</i>
15	COPD Stage III	BPCO Stade III	0,84	<i>COPD Stage III</i>
16	COPD Stage III	Ex.due to infection	0,73	Ex.due to infection
17	COPD Stage IV	COPD Stage IV	0,78	<i>COPD Stage IV</i>
18	Ex. due to Pn.Thor.	COPD Stage IV	0,59	COPD Stage IV
19	Ex.due to infection	Ex.due to infection	0,77	<i>Ex.due to infection</i>
20	Ex.due to infection	Ex.due to infection	0,90	<i>Ex.due to infection</i>
21	Ex. due to Pn.Thor.	Ex.due to infection	0,63	Ex. due to Pn.Thor.

6. Conclusion

In this paper, we proposed clinical decision support system based on combination of Case and Rule-Based Reasoning applied to the diagnosis and therapy of COPD which is a very grave disease bound to tobacco. Two diagnoses related to this pathology are also considered in this work, namely “Exacerbation of COPD due to infection” and “Exacerbation of COPD due to pneumo-thorax”.

The combination of CBR and RBR is made in order to gather their powers within the same system. RBR has proved its performance in modeling of reasoning which can be explain by humans, that is why we adopted it to conceive the reuse phase of CBR process by an expert system. The inference engine associated to knowledge base and forwarding chaining ensure adaptation task by drawing inferences starting from the diagnosis found in the retrieval phase and basing on some attributes having highest weights whose values will compose the set of facts of our expert system.

We wanted to give to the expert system a second task in addition to adaptation. Its knowledge base is enriched by rules modeling strategies of therapy and allowing to propose an adequate treatment to the doctor. Some preliminary experimentations show that reuse phase modeled with an expert system brings good impact on results of the system. Indeed, the adaptation was successful in most cases to give the real diagnosis of the patient.

This work consists also to propose some heuristics traced for estimating the similarity on missing data and symbolic descriptors in the retrieval phase. Results show also that these heuristics functions are benefic for the system because they optimize the result of the retrieval phase.

References

- [1] I. Bichindaritz, Case-Based Reasoning in the Health Sciences, “Why It Matters for the Health Sciences and for CBR”. K.-D. Althoff et al. (Eds.): ECCBR, LNAI 5239, (2008), pp. 1–17.
- [2] U. A. Mobyen, S. Begum and E. Olsson, “Case-Based Reasoning for Medical and Industrial Decision Support Systems”, Successful Case-Based Reasoning Applications, SCI 305, (2010), pp. 7–52.
- [3] R. Schmidt and L. Gierl, “A prognostic model for temporal courses that combines temporal abstraction and case-based reasoning”.
- [4] P. Koton, “Reasoning about evidence in causal explanations”. In: J. Kolodner, (Ed.): Proceedings of the Case-Based Reasoning Workshop, Clearwater Beach, Florida, (1988), pp. 260–270.
- [5] R. Turner, “Organizing and using schematic knowledge for medical diagnosis” in: J. Kolodner. (Ed.): Proceedings of the Workshop on Case-Based Reasoning, Florida, (1988), pp. 435–446.
- [6] C. Bradburn and J. Zeleznikow, “The application of case-based reasoning to the tasks of health care planning”, in: Wess S., Althoff K.-D., Richter M.M. (Eds.): Topics in case-based reasoning. Proceedings of the European Workshop on Case-Based Reasoning, EWCBR-93. Lecture Notes in Artificial Intelligence, vol. 837, (1993), pp. 365–378.
- [7] E. T. O. Opiyo, “Case-based reasoning for expertise relocation in support of rural health workers in developing countries”, in A. Aamodt and M. Veloso (Eds.): Case-based reasoning research and development. Proceedings of the International Conference on Case-Based Reasoning, ICCBR-95. Lecture Notes in Artificial Intelligence, vol. 1010, (1995), pp. 77–87.
- [8] Z. El Balaa and al, “FM-Ultranet: A decision support system using case-based reasoning applied to ultrasonography”, in L. McGinty, (Ed.): Workshop Proceedings of the International Conference on Case-Based Reasoning, (2003), pp. 37–44.
- [9] R. Schmidt and O. Vorobieva, “Combining Statistics and Case-Based Reasoning for Medical Research”, C.L. Mumford and L.C. Jain (Eds.): Computational Intelligence, ISRL 1, (2009), pp. 673–696.
- [10] I. Jurisica and al, “Case-based reasoning in IVF: Prediction and knowledge mining”.
- [11] J. Lieber and B. Bresson, “Case-based reasoning for breast cancer treatment decision helping”, in E. Blanzieri and L. Portinale, (Eds.): Advances in case-based reasoning. Proceedings of the European Workshop on Case-Based Reasoning, EWCBR. Lecture Notes in Artificial Intelligence, vol. 1898, (2000), pp. 173–185.
- [12] L. M. Fabbri, L. Fabrizio, B. Beghe and K. F. Rabe, “The Multiple Components of COPD. COPD: A Guide to Diagnosis and Clinical Management, Respiratory Medicine”, DOI 10.1007/978-1-59745-357-8_1, LLC (2011).
- [13] A. Grech and J. Main, “A Case-Based Reasoning Approach to formulating University Timetables Using Genetic Algorithms”, R. Khosla and al. (Eds): KES 2005. LNAI 3681, (2005), pp. 76-83.
- [14] A. Pedro, G. Calero, M. G. Albarran and D. A. Belen, “A Substitution-based Adaptation Model”, A work supported by the Spanish Committee of Science & Technology (CICYT TIC98-0733), (1998).
- [15] S. Wenqi and J. A. Barnden, “How to combine CBR and RBR for Diagnosing Multiple Medical Disorder Cases”, H. Munoz-Avila and F. Ricci (Eds): ICCBR, LNAI 3620, (2005), pp. 477-491.
- [16] C. Maurente, E. O. Edye, S. H. Delgado and D. R. Garcia, “Evaluation of Case Based Reasoning for Clinical Decision Support Systems applied to Acute Meningitis Diagnose”.
- [17] B. B. Spörl, M. Lenz and A. Hubner, “Case-Based Reasoning – Survey and future directions”.
- [18] I. Bichindaritz and C. Marling, “Case-based reasoning in the health sciences: What’s next?” Artificial Intelligence in Medicine, vol. 36, (2006), pp. 127-135.
- [19] I. Watson and F. Marir, “Case-Based Reasoning: A Review, AI-CBR”, Dept. of Computer Science, University of Auckland, New Zealand, <http://www.aicbr.org/classroom/cbr-review.html>, (2009).
- [20] A. Aamodt, “Case-based reasoning: Foundational issues, methodological variations, and system approaches”, in: AICOM, (1994), pp. 39–58.
- [21] R. Lopez de Mantaras, D. Mcsherry, D. Bridge, D. Leake, B. Smit, S. Craw, B. Faltings, M. Lou Maher, M. T. Cox, K. Forbus, M. Keane, A. Aamodt and I. Watson, “Retrieval, Reuse, Revision and Retention in Case-Based Reasoning”.
- [22] H. Papadopoulos, A.S. Andreou and M. Bramer, “A Fuzzy Non-linear Similarity Measure for Case-Based Reasoning Systems for Radiotherapy Treatment Planning” (Eds.): AIAI, IFIP AICT 339, (2010), pp. 112–119.
- [23] R. Schmidt, “Case-Based Reasoning in Medicine Especially an Obituary on Lothar Gierl”, Studies in Computational Intelligence (SCI), vol. 48, (2007), pp. 63–87.

- [24] L. Gierl, M. Bull and R. Schmidt, "CBR in Medicine".
- [25] M. Haddad, K. P. Adlassnig and G. Porenta, "Feasibility analysis of a case-based reasoning system for automated detection of coronary heart disease from myocardial scintigrams", *Artificial Intelligence in Medicine*, vol. 9, (1997), pp. 61–78.
- [26] C. Marling, E. Rissland and A. Aamodt, "Integrations with case-based reasoning". *Journal of The Knowledge Engineering Review*, vol. 00:0, 1–4. c, Cambridge University, (2005).
- [27] J. Jarmulak, S. Craw and R. Rowe, "Genetic Algorithms to Optimise CBR Retrieval", E. Blanzieri and L. Portinale (Eds.): *EWCBR 2000, LNAI 1898*, (2000), pp. 136–147.
- [28] F. –R. Florentino and J. M. Corchado. "Employing TSK Fuzzy Models to Automate the Revision Stage of a CBR System", R. Conejo and al. (Eds.): *CAEPIA-TTIA, LNAI 3040*, (2004), pp. 302–311.
- [29] T. Bruland, A. Aamodt and H. Langseth, "Architectures integrating Case-Based Reasoning and Bayesian networks for Clinical Support System", Z. Shi and al. (Eds.): *IIP, IFIP AICT 340*, (2010), pp. 82–91.
- [30] C. Marling, G. Petot and L. Sterling, "A CBR/RBR Hybrid for Designing Nutritional Menus", *AAAI Technical Report SS-98-04*, (1998), *AAAI* (www.aaai.org).
- [31] G. H. Lee, "Rule-based and case-based reasoning approach for internal audit of bank", *Journal of Knowledge-Based Systems*, vol. 21, (2008), pp. 140–147.
- [32] www.capitalsouffle.fr/COPD.htm.
- [33] T. Anane, M. Baghriche, R. BenAli, R. Boukari, Y. Berrabah, P. Chaulet, A. Fissah, D. Larbaoui, S. Lellou, S. Nafti and N. Zidouni, "Guide pour la prise en charge des maladies respiratoires de l'enfant et de l'adulte dans les unités sanitaires de base en Algérie", *Institut National de Santé Publique*, (2004).

Authors



GUESSOUM Souad

Guessoum Souad was born on the 15th of April 1970 at Khenchela (Algeria). She graduated from the University of Badji Mokhtar Annaba, Algeria, with a state Engineering degree in Computer Science, in June 1996. She is a doctoral student since 2006, and joined the department of computer science in her original university of Badji Mokhtar Annaba in 2006 as an assistant teacher. Her main area of expertise is Artificial Intelligence applied in medical domain.



LASKRI Mohamed Tayeb

Mohamed Tayeb Laskri was born at Annaba (Algeria) in 1958. He obtained a 3rd cycle doctorate on computer science (France, 1987). He obtained his PhD degree from the Annaba University (Algeria, 1995). He is chief of the Research Group in Artificial Intelligence within laboratory LRI. Its current research takes the reasoning in artificial intelligence like field of application privileged in particular in the image processing, the multi-agents systems, the engineering of the Human-machine interfaces and the automatic treatment of the natural language. He was President of the University of Annaba and the Scientific Council of the University of Annaba on 2002-2010. He was President of the Consortium CEMUR «collaboration Europe-Maghreb of Universities in Network» in 2009. He is Founder member of Arab Computer Society (ACS).



DJELLALI Hayet

DJELLALI Hayet was born on 16th of December 1969 in Annaba from Algeria; she graduated from the University of Badji Mokhtar Annaba, Algeria, with a state Engineering degree in Computer Science, in June 1994. She is a doctoral student since 2008, and joined the department of computer science in her original university of Badji Mokhtar Annaba in 2009 as an assistant teacher. Her main area of expertise is speaker recognition.



KHADIR Mohamed Tarek

Pr. Khadir Med Tarek was born on the 5th of July 1972 in Annaba Algeria. After succeeding in the baccalaureat, majoring in Maths and Technology in 1989, he graduated from the university of Badji Mokhtar Annaba, Algeria, with a state Engineering degree in Electronics Majoring in Control, in 1995. After two years' work in the computer industry, he undertook an M.Eng. at Dublin City University, Ireland Graduating with First class honors in 1998. Tarek Khadir received a Ph.D. degree from National University of Ireland, Maynooth, Ireland in 2002. He then continued with this institution as a post-doctoral researcher until September 2003 when he joined the department of computer science at my university of origin, university Badji Mokhtar Annaba, Algeria, as a senior lecturer. He succeeded in obtaining the HDR (Habilitation to Direct Research) in January 2005, and promoted full Professor in 2010.

