The SAFRS System: A Case-Based Reasoning Training Tool for Capturing and Re-Using Knowledge

Souad Demigha

***Abstract*—**The paper aims to specify and build a system, a learning support in radiology-senology (breast radiology) dedicated to help assist junior radiologists-senologists in their radiology-senology-related activity based on experience of expert radiologists-senologists. This system is named SAFRS (i.e. system supporting the training of radiologists-senologists). It is based on the exploitation of radiologic-senologic images (primarily mammograms but also echographic images or MRI) and their related clinical files. The aim of such a system is to help breast cancer screening in education. In order to acquire this expert radiologist-senologist knowledge, we have used the CBR (case-based reasoning) approach. The SAFRS system will promote the evolution of teaching in radiology-senology by offering the "junior radiologist" trainees an advanced pedagogical product. It will permit a strengthening of knowledge together with a very elaborate presentation of results. At last, the know-how will derive from all these factors.

***Keywords*—**Learning support, radiology-senology, training, education, CBR, accumulated experience.

I. Introduction

EARLY detection of breast cancer is considered as a major public health issue. Breast cancer incidence is the highest among female cancers and the second cause of mortality in Europe [1]. To address this problem, it is necessary to create the adequate conditions allowing for the installation of mass detection campaigns, i.e. involving the maximum number of women at risk. Detection is carried out starting from the analysis of breast images, primarily mammograms but also echographic images or MRI (*Magnetic Resonance Imaging*), coupled with the exploitation of information derived from the patient's history, from punctures, etc.

Therefore, the radiologist-senologist (i.e. the clinician in charge of breast cancer detection) grounds his diagnosis on the result of image analysis procedures and on the synthesis of various types of information. This process requires a significant amount of knowledge and know-how, which can be acquired only through a long time practice [2].

It is thus critical, in order to meet the requirements of mass detection, to have tools that support the initial training of radiologists-senologists, as well as their continuous training in this fast evolving domain.

In this context, the aim of our work is to specify and build a system dedicated to training in radiology-senology based on the exploitation of radiologic-senologic images and their related clinical information. The system is called SAFRS (*Système d'Aide à la Formation des Radiologues-Sénologues*, i.e. system supporting the training of radiologists-senologists). It is aimed at junior radiologists-senologists from the Department of Radiology of the Necker Hospital (Paris, France). The basic idea is to enable junior radiologists-senologists to have access to and learn from the experience of senior, expert colleagues.

The most commonly used mode in medical education consists in teaching with experiments, called *clinical cases*. These cases learned individually or in groups are examples resulting from real situations. The case-based reasoning (CBR) approach represents expert knowledge as a set of cases [3]. This set of cases (experience) may then be reused when solving new problems, e.g. when making new diagnoses. The CBR approach is totally suited to the aim of our SAFRS system, namely to enable junior radiologists-senologists to learn from the experience of their senior colleagues.

Therefore, we have adopted this approach and defined a model to represent the experience of expert radiologists-senologists as cases. This model has been empirically validated on about forty real cases.

The paper is organized as follows:
- section 2 positions our work with respect to existing case-based reasoning systems in radiology and training;
- section 3 describes knowledge capitalizing in the medical imaging domain;
- section 4 presents the case-based reasoning approach applied to the SAFRS system;
- section 5 presents the architecture and working principle of the SAFRS system;
- section 6 presents the case representation model with the UML language;
- section 7 details the retrieval process in the SAFRS system and describes similarities used to compare cases;

Manuscript received September 30, 2007. This work was supported in part by the C.R.I (Centre de Recherche en Informatique) Department of Computer Science at the Sorbonne University (Paris 1), Paris (France) and in partnership with the Department of Radiology of the Necker Hospital in Paris (France).

S. Demigha is a membership of the C.R.I (Centre de Recherche en Informatique) Department of Computer Science at the Sorbonne University (Paris 1), Paris (France), 90, rue de Tolbiac 75634 Paris Cedex 13 France (corresponding author to provide phone: (33 1) 44 07 86 04; fax: (33 1) 44 07 89 54; e-mail: souad_demigha@ yahoo.fr).

- section 8 presents parts of implementation of the SAFRS system;
- section 9 provides a discussion on the strong and weak points encountered during the realization of the SAFRS system and points to further research work ;
- Finally, section 10 concludes this research work.

II. State of the Art

In radiology, there exist several systems based on case-based reasoning (CBR). Among those which approach more our field, we should mention the system IDEM (*Images et Diagnostic par l'Exemple en Médecine*) in the anatomo-pathology domain [4], the system ImageCreek for scanner images interpretation [5], the two systems CASIMIR/RBR and CASIMIR/CBR for decision making in senology [6] and finally the system MacRad, a case-based retrieval system in radiology [7]. These systems are rather dedicated to decision making and diagnosis and are generally limited to one process (anatomo-pathology, mammography or scanners). No application of these systems to the training of radiologists is mentioned.

Case-based reasoning was applied in the training area but paradoxically very little [8]. We have found some training systems in senology [9], [10], [11] and [12], but they don't use case-based reasoning. Contrarily to the senology domain, we have found some case-based training systems in medical imaging more generally. Let us mention for example the CADI (Cardiac Auscultation Diagnosis Instruction) system which is an intelligent tutoring environment to teach cardiac auscultation. It allows students to "learn by doing", while being guided by an expert cardiologist [13]. The design of the environment is grounded in two established theories: *goal-based scenarios* and *case-based reasoning*. The principles of *goal-based scenarios* guide the framing of the learning task. In a typical CADI scenario, the student plays the role of an internist who is supervising his/her residents on rounds. *Case-based reasoning* suggests a pedagogical approach of teaching through the use of case presentation.

Our SAFRS system is a training system in radiology-senology grounded on the case-based reasoning approach.

The system covers all processes of the radiology-senology domain (*clinical examination, image reading, radiological interpretation, and anatomo-pathologycal examination).*

The SAFRS system aims at capitalizing and re-using both *product knowledge (mammographies and associated diagnoses…)* and *process knowledge (heuristics)*. While *the product* is the result to be achieved, *the process* is the way which leads to the achievement of this result.

Section 3 presents knowledge capitalizing in the medical imaging domain.

III. Knowledge Capitalizing in the Medical Imaging Domain

In medicine, current methods are unable to capitalize and to re-use knowledge acquired from experience. Re-use is employed with an *ad hoc* manner. It is a traditional technique based on experience acquired during developments of different systems in a specific domain. The *ad hoc* manner is inadequate because it does not allow to share accumulated experience, contrarily to what medical experts wish to obtain. In order to achieve investigations on a vast number of cases, experts only use their own experience, even though this is a vast amount of experience [14].

It would be advisable to gather the experience of the numerous experts for a shared utilization. Besides the obvious advantages that could result from this shared knowledge, it would allow to homogenize the knowledge on the same topics by standardizing the vocabulary and definitions. Identical notions should be labelled using the same terminology so as to compare them [15].

It is a well-know fact that in medicine, the development of specialized ontologies is a mandatory step for elaboration and maintenance of increasing a thesaurus and not an ambiguous one, for the sake of communication between terminologists [16].

We have built an ontology based on the standard BI-RADS (Breast Imaging Reporting and Data System) [17], on the scientific reports of the EBM (Evidence-based on medicine) and on the reports and experience of radiologists-senologists of the Necker Hospital in Paris (France) in view of representation of radiologic-senologic knowledge and associated clinical reports [15].

This ontology is fitted to the description of the radiologic-senologic knowledge shared by the scientific community of technicians, practitioners, gynecologists, radiologists, surgeons and anatomo-pathologists. It represents a unifying scope for reducing and eliminating ambiguities as well as conceptual and terminological confusions. This ontology allowed us to obtain the conceptual model of the domain in radiology-senology, which is structured as cases using the case-based reasoning approach [14].

We have analyzed requirements of radiologists-senologists with the Department of Radiology of the Necker Hospital in Paris using the Crews-l'Ecritoire approach (Cooperative Requirements Engineering With Scenarios [18]. Radiologic-senologic knowledge is made of both text and images. We have only considered textual knowledge; images are just associated to patients' reports for the sake of information. Analysis performing has allowed to structure the radiologic-senologic knowledge according to stringent rules. It is an original approach to solve the issue consisting in considering the ontology definition as an engineering issue requirement.

Section 4 presents the applying of the case-based reasoning approach in the SAFRS system.

IV. The Case-Based Reasoning in the SAFRS System

Case-based reasoning (CBR) is an Artificial Intelligence approach to learning and problem solving based on past experience. A past experience is stored under the form of solved problems ("cases") in a so-called "*case base*". A new problem is solved grounded on adapting solutions to similar problems (see Fig. 1) to this new problem.

In the SAFRS system, the case–based reasoning is a cyclic, five-phased process [15]:

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:10, 2007

1) Elaborate: It consists in describing a new case (new case to be diagnosed). From the new problem, the case is elaborated from the general knowledge, while keeping only relevant information.
2) REtrieve: the aim of this phase is the selection of one (or several) case(s) which solve(s) a problem similar to that of the new case (also called the target).
3) REuse (adaptation): the target and the retrieval case (source) are combined to reach a solution. The solution of the source is adapted to account for the differences between the target and the source.
4) REvise: the purpose of this phase is to make sure that the proposed solution is correct and shall lead to success if applied.
5) REtain: the new case and its solution are stored into the case base. Thanks to this learning phase, the system requires new knowledge at each reasoning cycle.

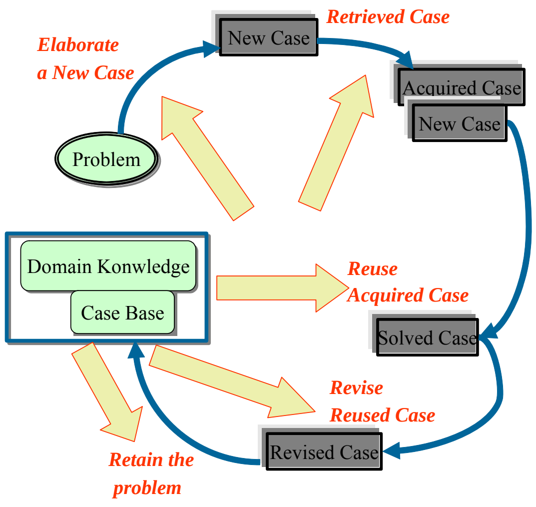

Fig. 1 The CBR reasoning in the SAFRS system

Section 5 presents in details the SAFRS system.

V. SAFRS: A Case-Based Training System in Radiology-Senology Domain

This section describes the architecture of the SAFRS system and its principle functioning.

A. Architecture of the SAFRS system

Fig. 2 represents the architecture of the SAFRS system.

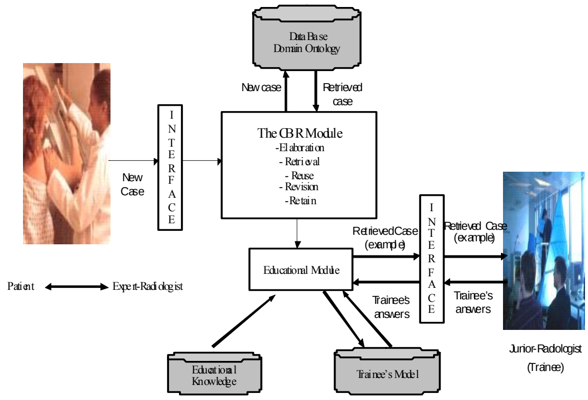

Fig. 2 Architecture of the SAFRS system

B. Working Principles of the SAFRS System

Carrying out the SAFRS system relies on 5 modules: *the domain ontology (data base module)*, *the CBR module*, *the educational module*, *the trainee's model (trainee's module)* and the *user's interface*.

- The domain ontology

The *domain ontology (data base module)* contains, on the one hand, the domain ontology built from the BI-RADS dictionary. This ontology includes case knowledge, concepts used to describe cases, taxonomy of concepts, relationships between concepts and constraints. It represents the contents of teaching methods validated by expert radiologists-senologists. On the other hand, it includes knowledge of different matchings between cases.

- The CBR module

Once the new case (target case) is described by the junior radiologist-senologist (the trainee), it is stored by the system. Following this, the new case is matched with all the other cases of the case base (old cases). This matching allows to find one or several relevant cases that help resolve the new problem of the trainee. A case includes one or several educational objectives. An educational objective is associated to a strategy that the expert radiologist-senologist suggests in the training of the junior-radiologist. It is one among the strategies of the MAP allowing to guide the junior-radiologist in a flexible and a supple manner [19]. For example, this strategy (problem-solving strategy) may allow to respect the course of the senologic process: "*performing the clinical examination*", "*image reading*", "*radiological interpretation*" and *finally* "*the anatomo-pathological examination*"**.** The matching is carried out via a matching algorithm.

- The educational module

Using domain knowledge (expertise of confirmed radiologists-senologists capitalized in the form of cases), the educational module develops a reasoning that allows to evaluate the trainee and to guide him/her using an educational strategy adapted to the trainee's model.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:10, 2007

- The trainee's model

In the SAFRS system, the trainee's model (module) takes into account capabilities of the trainee by proposing him/her several levels of exercises; it generates a feedback suited to each type of error. It allows understanding the origin of the error and proposes a remediation strategy; it builds a diagnosis of errors (misdiagnosis) and evaluates the trainee's work. Misdiagnoses allow more relevant and more effective interventions of the system. It is aimed at helping the trainee to use the necessary knowledge and to neglect the non-relevant one.

- The interface

The dialogue *'trainee-system'* is carried out by two adapted interfaces: the *author's interface* and the *trainee's interface*. The *author's interface* serves as a communication tool between the expert-radiologist and the system. It allows him/her to describe a new case and to store it into the case base. The *trainee's interface* serves as dialogue between the trainee and the system.

Section 6 presents in details the case representation model.

VI. The Case Representation Model

The case representation model is structured according to the four phases [20] of the radiologic-senologic process:

1) *clinical examination*, which includes data about health patient history: screening history, current health status, and previous clinical.
2) *image reading*, which consists in searching and extracting relevant information (imaging data and textual data).
3) *radiological interpretation*, which is based both on clinical data (patient's history screening, current health status, information on previous clinical examination) and radiological data (information such as defined by the BI-RADS standard).
4) *anatomo-pathological examination,* which depends on the result of radiological interpretation. It grasps information about anatomo-pathological examination such as type of procedure, reporting source, laterality, histopathology, staging and therapy.

Fig. 3 presents an outline of the case representation model, by highlighting the links between the phases of the radiologic-senologic process. The case representation model uses the UML formalism [21].

We have chosen to represent the cases using an object-oriented approach (and hence the UML formalism) for the following reasons: on one hand, radiologists-senologists manage complex data (images, sounds, temporal data,…).

To represent this complexity, an object model is perfectly suited. On the other hand, the object representation combines with case-based reasoning in a natural way [22].

We divide our case representation model into 3 hierarchic levels:

1) *The first level (the case).* We consider that a case is a patient at different intervals of treatment (time). A case may comprise several successive radiologic-senologic episodes.
2) *The second level (the sub-case).* It is one radiologic-senologic episode *(clinical examination, image reading, radiological interpretation* and *anatomo-pathology)* for a given patient.
3) *The third level (the sub-sub-case).* It represents one phase of a radiologic-senologic episode for a given patient (clinical examination OR image reading,…).

The problem part of a *sub-sub-case* generally refers to solutions or more generally data produced in the previous phases of the same episode. For example, the problem part of anatomo-pathological phase contains data of radiological interpretation solution.

Experts' experience is represented as knowledge; both *product knowledge (mammographies and associated diagnoses…)* and *process knowledge (heuristics)* are considered. The case representation model complies with the standards defined for digital mammography and CAD mammography [23]. In particular, we use the DICOM (Digital Imaging and COmmunication in Medicine) [24] and BI-RADS dictionaries to describe and index data.

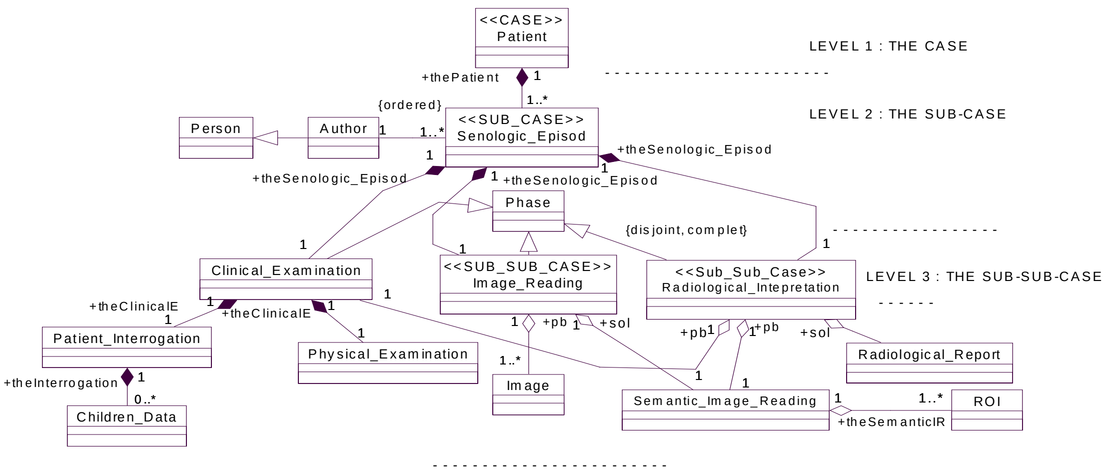

Fig. 3 The case representation model: general view

Section 7 details the retrieval process in the SAFRS system.

VII. The Retrieval Process of the SAFRS System

In the SAFRS system, the retrieval process is modelled by a MAP called the retrieval MAP (see Fig. 4), [19].

A. The Model of the MAP

The model of the MAP is used to represent parts of processes included in the product model (case representation model). The model of the MAP is an intentional representation system. It is based on concepts of *intention* and *strategy*. It includes one or several *sections*. A *section* is based on two concepts: *intention (or goal)* and *strategy*. The concept of *intention* aims to capture the objective to be achieved at one time of a process. A *strategy* is the manner to achieve an intention. A *section* is an aggregation of two types of intentions: a *source intention*, a *target intention* and a *strategy* as well. Each section corresponds to a strategy which can be used in order to achieve one target intention, once a source intention is achieved.

A MAP is represented by a graph oriented and labelled.

Intentions represent nodes and *strategies* represent the arcs. A section is then represented by two nodes linked by an arrow.

A section must be selected when it is initialized. The selection of sections is based on two types of directives: *Intention Selection Directives* (ISD) and *Strategy Selection Directives* (SSD). ISD guides the selection of the next intention to be achieved. SSD plays the role of selecting the best strategy adapted to the encountered situation. A directive includes a *signature* and a *body*. A *signature* represents the visible part of the directive. It characterizes the conditions where it is applied and the result obtained as well.

A *signature* is defined by the couple *<(intention), situation>*. Each directive is applied in a particular situation to satisfy a particular intention.

A *body* defines the followed step in order to satisfy the intention captured in the signature.

A directive includes two types of directives: *strategic directive* and *tactical ones*: (1) the *strategic directive* represents a strategic view of the multi-step development based on a set of intentions and strategies. It is represented by a MAP and a set of associated directives and (2) the *tactical directive* has a three-structure. It is composed of three other directives (*context*: a *context* represents the development of a process by a hierarchy of contexts): *plan, selection* (the selection of several alternative sub-directives) and *executable*.

- *A plan directive* corresponds to a complex problem decomposed into a set of sub-problems. The execution of the composed directives is defined by a graph. The nodes of the graph are directives (components of the plan). Arcs (previous links) represent arranged or parallel transitions between directives.
- *A selection directive* corresponds to a situation that necessitates the exploration of different possibilities.
- *An executable directive* corresponds to an intention which can be characterized by an action of the product transformation or an action of selection of an other directive. Both actions are named: engineering action (atomic or complex) and delegation action that delegates the realization of an intention to an other directive, respectively.

B. Similarities

Object-oriented case representation requires approaches for similarity assessment that allow to compare two differently structured objects, in particular, objects belonging to different object classes. In this section, we briefly illustrate the definition of similarities based on the case representation model.

In the radiology-senology case representation model, cases are collections of *objects*, each of which it is described by a set of *attribute-value pairs*. The structure of an object is described by an *object class* that defines the set of *attributes* together with a *type* (set of possible *values* or *sub-objects*) for each *attribute. Object classes* are arranged in a *class hierarchy*, that is, a tree in which *subclasses* inherit *attributes* as well as their definition from the current class.

We define a hierarchy of attribute types. New types are defined by building subtypes of the existing elementary types shown in Table I. They differ in their usability: a type may be used as an *immediate* or *derived* type. While *immediate types* cover the whole range of possible values of a type, *derived types* get restricted in their range by defining an enumeration of elements of their elementary types or, in case of numeric types, by specifying an interval [25].

TABLE I
Elementary Types in the SAFRS System

Type	Usability
Integer	Immediate and derived
Float	Immediate and derived
Date	Immediate and derived
Boolean	Immediate only
String	Immediate and derived
Enumeration	Immediate and derived
Ordered Enumeration	Derived only
Text	Derived only

The approach we have chosen to determine similarities is to establish a comparison between attributes (attribute by attribute), then to each attribute corresponds a comparison measure, it is a *local similarity measure*. It determines a similarity between two attribute values, and for each object we determine a *global similarity measure* which determines the similarity between two objects (or between the case and the query) based on the local similarity of the belonging attributes.

The *local similarity measure* allows to compare any two type values. It returns a numeric value from the interval [0..1]. This value is further used in the computation of a global similarity.

C. The Retrieval Process

The retrieval process is, with the MAP, a multi-step/multi-algorithm process, which permits to retrieve similar cases in various modes.

The retrieval MAP of the SAFRS system represented on the graph (Fig. 4) defines, besides the two intentions ‘to start’ and

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:10, 2007

‘to stop’, two major intentions for the retrieval process achievement ‘to elaborate the new case’ and ‘to retrieve similar cases’.

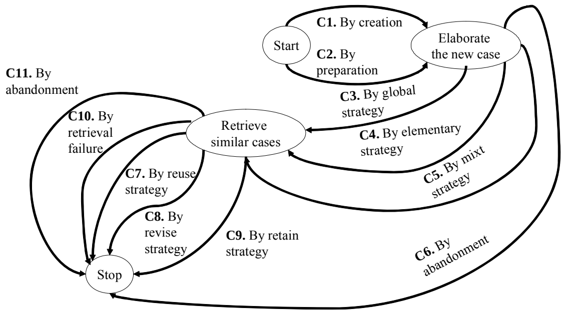

Fig. 4 The retrieval MAP

The intention ‘to elaborate the new case’ is achieved according to two strategies: ‘by preparation’ and ‘by creation’, contained in sections C1 and C2, respectively. It consists in describing a new case (new case to be diagnosed). From the new problem, the case is elaborated from the general knowledge, while keeping only relevant information. If the case description is complete, then the intention ‘to elaborate the new case’ has the same meaning as the creation (strategy ‘by creation’) or the preparation (strategy ‘by preparation’) of the case, else the creation phase is simplified.

The intention ‘to retrieve similar cases’ is achieved according to three complex strategies: ‘global strategy’ (or global retrieval strategy), ‘elementary strategy’ (or elementary retrieval strategy) and ‘mixed strategy’ (or mixed retrieval strategy) that are contained in sections C3, C4 and C5, respectively. The ‘*global strategy*’ included in section C3 allows for retrieval at the global level, i.e. the case. The retrieval process starts at the sub-sub-case level, then we go to the intermediate level, the sub-case, and finally it ends to aggregate at the case level. ‘*Elementary* strategy’ included in section C4 allows to combine one to three phases, i.e. the sub-sub-case (‘image-reading’, ‘radiological interpretation’ and ‘anatomo-pathological examination’) of the senological process.

The ‘*mixed strategy’* included in section C5 allows to combine the first two strategies (global and elementary ones). It aims to go back and start from the elementary level (sub-case), until it finds cases of interest in the treatment of the new case (target case). The ‘abandonment strategy’ included in section C6 allows the ‘case expert’ to abandon his/her retrieval process for the new case, before starting the retrieval when he/she makes mistakes in his/her reasoning, thus allowing him/her to start again the retrieval process, without starting from the very beginning, i.e. from the source intention ‘to start’ of the MAP.

Once the ‘case expert’ has carried out the retrieval process, i.e. he/she succeeded or failed in searching an interesting case for solving the new problem, he/she has got four possibilities to treat this new case: the ‘reuse strategy’ included in section C7 allows to revise the validity of retrieved solution, which is retained for the goal problem (new problem to solve). The ‘revise strategy’ included in section C8 allows to revise the case according to three steps: to revise it ‘by test’, ‘by correction’ and finally ‘by validation’. The ‘retained strategy’ included in section C9 allows to integrate to the case base the new solved problem, if the latter confers novel abilities to the system. The strategy ‘by retrieval failure’ included in section C10 allows to send back a negative result from the case base to the ‘case expert’ when no case could be identified as similar enough to the target case (new case). Finally, the last strategy ‘by abandonment strategy’ included in section C11 allows the case expert to abandon the retrieval of similar cases if he/she deems it necessary, even after the overall process is achieved.

The retrieval MAP (see Fig. 4) proposes two strategies, from ‘to progress from the source intention’ to ‘to elaborate the new case’ since the target intention ‘to research the similar cases’, and ‘to progress from "to elaborate the new case" to ‘to stop’.

1) SSD$_1$: ‘to progress’ since ‘to research the similar cases’

The SSD$_1$ guides the selection of one of three strategies allowing ‘to progress’ since ‘to research the similar cases’ from ‘to elaborate the new case’. It is a selection directive proposing three possibilities:

- To select (**DRI$_{1.1}$**: <(elaborated case), to research the similar cases by the global strategy>)>;
- To select (**DRI$_{1.2}$**: <(elaborated case), to research the similar cases by the elementary strategy>)>;
- To select (**DRI$_{1.3}$**: <(elaborated case), to research the similar cases by the mixed strategy>)>.

2) SSD$_2$: “to progress” since “to stop”

It is a plan directive proposing an *executable plan context*: **DRI$_{2.1}$**: <(new case), (to stop by abandonment)> ;

The Global Strategy

The global retrieval strategy consists in retrieving the case in its totality (see Fig. 5).

The Fig. 5 models this strategy. Indeed, it is a *plan directive*: <(new case), to research the similar cases by global strategy> composed of a hierarchy of plans which contains three different contexts: plan, selection and executable.

The plan directive DRI$_3$ proposes three basic sub-directives:

- DRI$_{3.1}$: <(new case), to calculate similarities at the sub-sub-case level>* ; (* means an iterative form).
- DRI$_{3.2}$: <(sub-sub-cases selected), to calculate similarities at the sub-case level>* ;
- DRI$_{3.3}$: <(sub-cases selected), to calculate similarities at the case level>*.

As shown on Fig. 5, the plan directive DRI$_{3.1}$: <(new case), to calculate similarities between sub-sub cases>* proposes two plan contexts for the realization of its intention:

- DRI$_{3.1.1}$: <(new case), to retrieve similar cases by subsumption>* ;
- DRI$_{3.1.2}$: <(new case), to retrieve similar cases by similarity>*.

The subsumption is a mechanism of discrimination. The directive DRI$_{3.1.1}$ is performed by the execution of two plan contexts:

- DRI$_{3.1.1}$: <(new case), to retrieve similar cases by subsumption>*. The intention ‘to research by

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:10, 2007

subsumption' is performed via two executable contexts: $DRI_{3.1.1.1}$: <(index new case), to match the new case index with the abstract case>* and $DRI_{3.1.1.2}$: <(set of indices), to evaluate the subsumption>*.

To evaluate the subsumption consists of browsing a net of indices where, at each node, cases are selected by taking into account the subsumption criterion.

For facilitating the retrieval process, the case is abstract in order to extract indices. The abstraction is aimed to divide the problem descriptors of the input into two classes: the relevant descriptors (useful) and the non-relevant descriptors (not useful) or noises. The abstraction consists in eliminating noises [26].

- $DRI_{3.1.2}$: <(new case), to select a sub-set of relevant cases>*.

The intention 'to select a sub-set of relevant cases' eliminates the very distant cases and selects a set of cases that are suitable for the target problem. It implies that cases are organized in a classification hierarchy according to relevant characteristics. The selection of these characteristics determines the capability to retrieve the 'best' cases.

After restricting the research space, the 'case author" performs a more specific comparison between the target problem and each source case previously selected by discrimination 'by subsumption' with the plan directive 'by similarity': $DRI_{3.1.2}$: <(New case), to retrieve similar cases by similarity>* .

The directive $DRI_{3.1.2}$: <(new case), to retrieve similar cases by similarity>* is performed by two plan contexts: $DRI_{3.1.2.1}$: <(set of index), to research by similarity>* and $DRI_{3.12.2}$: <(set of similar cases), to select the most similar case>*.

- 'To research by similarity' (to research similar cases) performs a comparison more specific between the target problem and the source case previously selected by discrimination. This comparison necessitates a two by two comparison of cases, attribute by attribute. This directive proposes two plan directives for the realization of its intention $DRI_{3.1.2.1.1}$: <(selected cases), to match selected cases and the new case>* and $DRI_{3.1.2.1.2}$: <(matched cases), to evaluate the similarity>*.

- The intention 'to match selected cases and the new case': the matching process compares two by two characteristics of cases. In most systems, the matching is performed on characteristics of cases: it is a global matching (global similarity by attribute weighting at a local similarity level).

- The intention 'to evaluate the similarity': a similarity measure is used in order to arrange source cases by decreasing the similarity with the target case. The evaluation is performed by considering common characteristics; each one has a significant importance level (weight) of the role that each element of a problem plays in the reuse of elements of the solution. The similarity evaluation is assumed to depict the facility of the reuse of a source case.

- 'To select the most similar case': the solution of cases having the best 'score' is selected for the target problem.

The directive plan $DRI_{3.1.2.1.1}$: <(selected cases), to match selected cases and the new case>* proposes two selection alternatives to complete the retrieval process: $DRI_{3.1.2.1.1.1}$ <(selected cases), to calculate similarities between attributes>* and $DRI_{3.1.2.1.1.2}$ <(selected cases), to calculate similarities between objects>*. These directives allow the computation of similarity measures between attribute-values *(a local similarity measure)* and objects (*global similarity measure*) (sub-sub-case, sub-case and case).

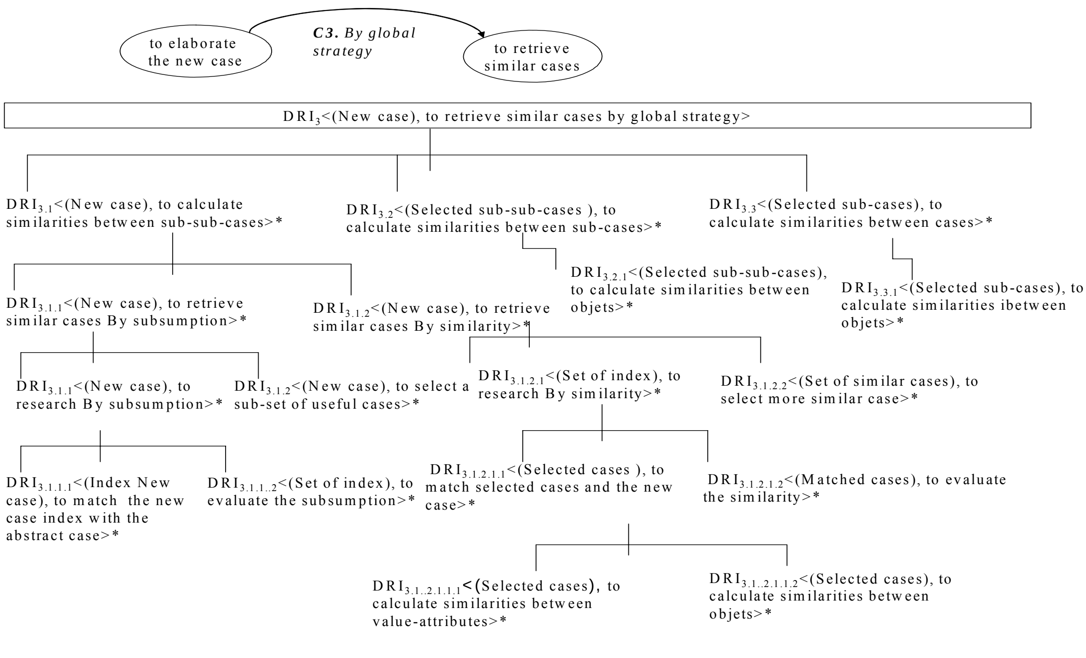

Fig. 5 The global strategy (a plan directive: hierarchy of plan contexts)

The Elementary Strategy

The objective of elementary strategy (or elementary retrieval strategy) is to offer the 'case author' various possibilities to resolve his/her problem. The phases of the radiologic-senologic process have a common importance in making the diagnosis. In the absence of complete information on the new case, the case author only considers into the case base the knowledge that resembles new knowledge. The case author can start the process with some knowledge of one phase of different phases of a patient x; for instance, this knowledge is compared with the knowledge of the new case. The case author selects knowledge of another phase concerning another patient y. He/she combines these knowledge and reiterates the process whenever required to make a diagnosis. All these fragments, coming from different phases of different patients or even from a same patient, combined together (in the case that the patient had previous reports), make up one solution of the new problem to solve. The assessment of the similarity (attributes and objects) is performed in the same manner as the global strategy and the mixed strategy.

A *local similarity measure* allows comparing any pair of value types (we have defined a hierarchy of UML attribute types). It returns a numeric value from the interval [0..1]. This value is further used in the computation of a global similarity. A *global similarity measure* determines the similarity between two objects (or between the case and the query), based on *the local similarity* of the belonging attributes.

The intention 'to calculate similarities between attribute-values' of the directive DRI$_{3.1.2.1.1.1}$ allows to use the hierarchy of UML types [20]. Indeed, according to various types of attributes, a similarity measure is selected.

The two other main sub-directives of the DRI$_3$:

- DRI$_{3.2:}$ <(selected sub-sub-cases), to calculate similarities between sub-cases;
- DRI$_{3.3}$: <(selected sub-cases), to calculate similarities between cases>* are executable contexts and thus are not factorized.

- The second sub-directive of the directive DRI$_{3.2}$: <(selected sub-sub-cases), to calculate similarities between sub-cases>* is a plan directive including one context plan: DRI$_{3.2.1:}$ <(selected sub-sub-cases), to calculate similarities between objects>*.
- The third sub-directive of the directive DRI$_{3.3}$: <(selected sub-cases), to calculate similarities between cases >* is a plan directive including one context plan: DRI$_{3.3.1}$: <(selected sub-cases), to calculate similarities between objects>*.

Fig. 6 models this strategy. Indeed, it is a plan directive: <(new case), to research similar cases by elementary strategy> composed of a hierarchy of plans which include three contexts: plan, selection and executable. The plan directive DRI$_4$ proposes three principal sub-directives:

- DRI$_{4.1}$: <(new sub-sub-case image reading phase), to calculate similarities at the image reading phase>* ;
- DRI$_{4..2}$: <(solution part of the image reading phase), to calculate similarities at the radiological interpretation level (RI)>* ;
- DRI$_{4..3}$: <(solution part of the RI phase), to calculate similarities at the anatomo-pathological examination phase (AE)>*.

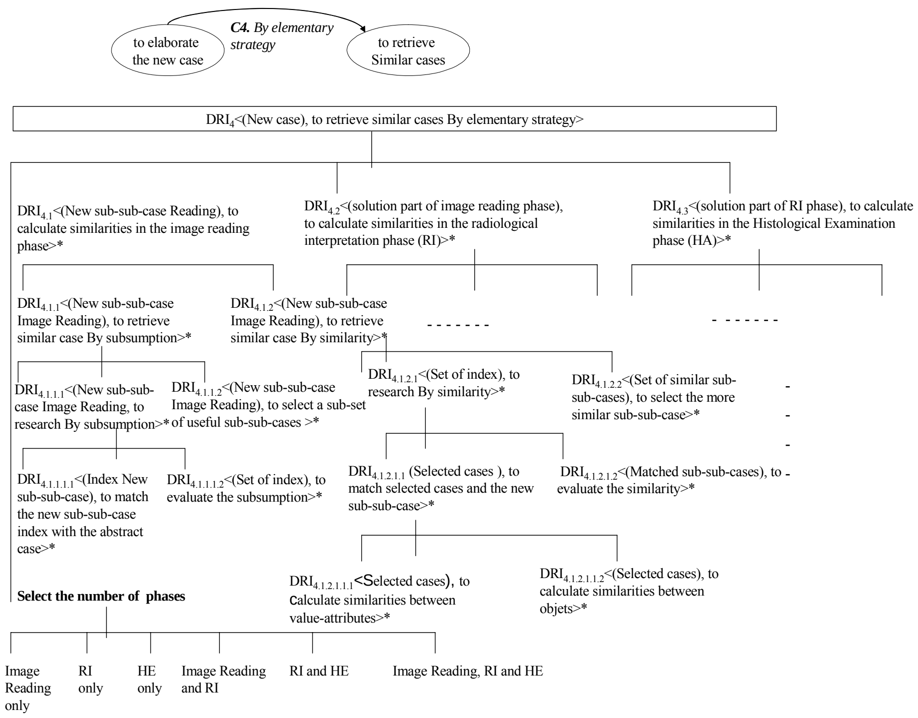

Fig. 6 The elementary strategy (a plan directive: hierarchy of plan contexts)

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:10, 2007

The Mixed Strategy

The mixed strategy allows to combine the first two strategies, the global strategy and the elementary strategy. For a radiologist (in the case that he/she combines several knowledge from various sources), the interest of this strategy lies in picking up knowledge at the intermediate level, to find again archives of previous examinations, and thus to obtain a full knowledge.

Fig. 7 models this strategy. Indeed, it is a plan directive: <(new case), to research similar cases by mixed strategy> composed with a hierarchy of plans including three contexts: plan, selection and executable.

The directive plan DRI_5 proposes five main sub-directives:

- $DRI_{5.1}$: <(new sub-sub-case image reading phase), to calculate similarities at the image reading phase>*;
- $DRI_{5.2}$: <(solution part of the image reading phase), to calculate similarities at the radiological interpretation phase (RI)>*;- $DRI_{5.3}$: <(part solution of the RI phase), to calculate similarities at the anatomo-pathological examination phase (AE);
- $DRI_{5.4}$: <(selected sub-sub-cases), to calculate similarities at the sub-case level>*;
- $DRI_{5.5}$: <(selected sub-cases), to calculate similarities at the case level)>.

As shown on Fig. 7, the directive $DRI_{5.1}$ is a hierarchy of directives of plan contexts, selection and executable. This hierarchy has the same course as the directive $DRI_{3.1}$ of the global strategy. We do not provide details of the steps of calculation of similarities.

In Section 8, we describe the implementation of certain parts of the SAFRS system.

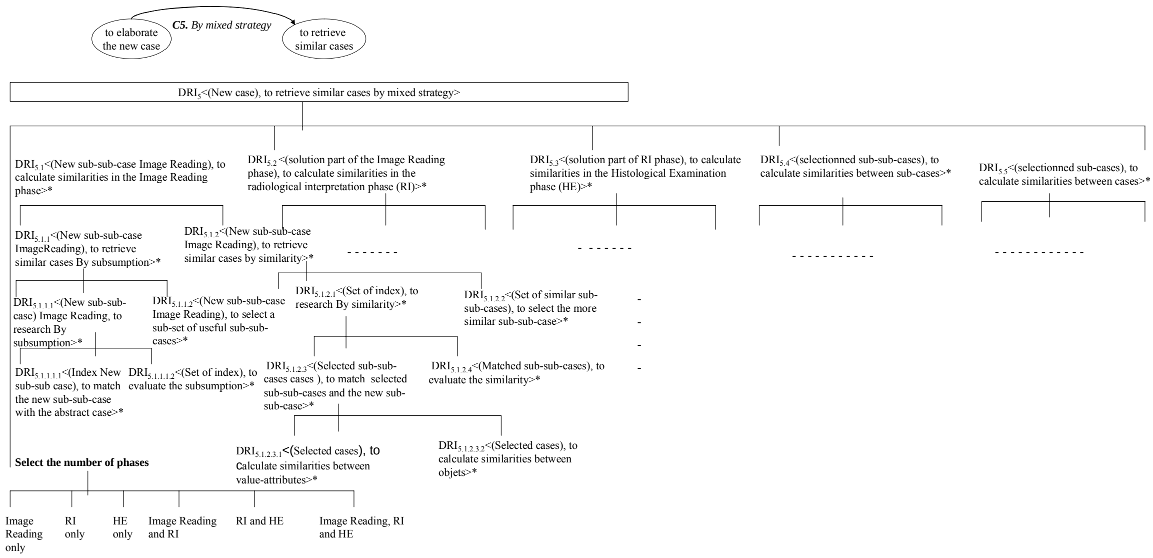

Fig. 7 The mixed strategy (a plan directive: hierarchy of plan contexts)

VIII. IMPLEMENTATION

This section presents the development of the prototype of the SAFRS system which aims to validate our approaches of capitalisation and re-use of experience of the experts in radiology-senology of the Department of Radiology of Necker Hospital of Paris, in order to develop the subsequent training of junior radiologists-senologists. The prototype was validated on forty real cases associated with 90 *mammographic* and *echographic* images of patients of the Department of Radiology of the Necker Hospital.

The final aim is to lead to a system of CBR in radiology-senology which allows for a training of the radiologists-senologists in its initial training phase. This training will have a deep interest for a junior radiologist insofar as it can provide the experience of the expert radiologists in the form of clinical cases. It will provide a tool which will enable the radiologists to capitalize and to re-use accumulated experience in radiology-senology and also the know-how of the experts.

The prototype development deals with the two following modules: *case representation model (or description module)* and part of the *similarity model (the retrieval process).* We have taken an interest in the strategy at the global level (the case). The other strategies (elementary and mixed strategies) will be developed subsequently.

The *description module* allows to store knowledge in radiology-senology and to retrieve them. The interfaces allowing for knowledge retrieving are implanted in their totality in accordance with the selected design. The similarity module led to the development of certain algorithms of the retrieval process. This module makes it possible to reason by similarity according to several selected strategies using the model of the MAP developed in Section 7.

The section is organized in four sub-sections as follows: the first sub-section describes our choices in terms of tools (implementation and validation tools). The second sub-section discusses the Database Management Systems (DBMS) to ensure the persistence of radiologic-senologic data. The third sub-section introduces the implementation of the representation model of UML case objects as displayed in Section 6 in the programming language Java which has been selected. Lastly, the fourth and last sub-section presents the implementation of the retrieval process.

A. The Programming Language

In the previous section the fact that *the case representation model* was an object model was emphasized. Therefore, we deemed it obvious to use an oriented-object programming language for the development of the prototype. Besides, we want to transfer in the long-run the prototype on the Internet to give to the junior radiologists-senologists the opportunity to acquire training in Internet technologies on various platforms. The Java programming language has the advantage of being portable.

For the development of our application in Java, we used an integrated development environment (IDE). It is software which makes it possible to write applications quickly and effectively. There is a great number of environments integrated of Java development on the market. We used *JCreator* by *Wendel de Witte (http://www.jcreator.com)*. It is a powerful IDE for *Java.technologies*.

B. The Persistence of Data

The knowledge of the radiologists-senologists structured in the form of cases must be stored in the 'case base'. Thus, we need to make persistent the cases created by the expert radiologists-senologists. A Database Management System (DBMS) makes it possible to ensure this persistence.

Among the *Database Management Systems (DBMS)* commercially available, two DBMSs are opposed for the implementation of a 'knowledge base' of medical images: relational DBMS (RDBMS) and the object-oriented DBMS (OODBMS).

Moreover, our modeling is UML object-oriented; thus, to establish the persistence of data conceptually "object" via a relational DBMS amounts to carrying out a stage of transformation of the object model into a relational model, called the *"mapping".* This stage is time-consuming and requires a double competence on behalf of the object developer and relational.

To cure these problems object DBMSs were carried out. At the end of this argument, we thus choose an Object *Data Bases Management System (ODBMS)* to ensure the persistence of the objects in the SAFRS system.

There are many commercial object databases, such as ObjectStore *(http://objectstore.com),* FastObjects (http://www.FastObjects.net) starting from PÖET *(http://www.poet.com)* which have specific connections (bindings) for the use of programming languages to objects, like C++ and Java (in the case of Java that means that the table of Java chopping or Vector of Vectors can persist in one of these databases).

We have selected an ODBMS which allows an interfacing with the Java language named FastObjects starting from PÖET with standard JDO *(Java Data Objects).* The commercial system PÖET implements for its object-oriented ODBMS an extension of the languages Java and C++, and it functions on the PC, Macintosh and Unix platforms. A customer has access to a database located on any platform. In addition, with the help of an explicit and official authorization, it is possible to use a free version of PÖET for the sake of evaluation procedures.

Fast Objects stores the objects in a transparent way. By using APIs of an ordinary database, it eliminates the mapping and gives transparent persistence. It helps to carry out applications more quickly and gives the highest performances.

Java Data Objects (JDO) provides an open and standard API to store the objects of the Java applications in any type of databases. For the object databases**,** JDO provides additionally the first and single standard API to store Java objects. It leaves to the teams of Java development the choice of the best database for their application as well during the design step as during the deployment step.

C. The Case Representation Model Implementation

In this section, we introduce the implementation of the UML case objects representation model as displayed in Section 6 in the programming language Java which has been selected. Our focus is on the main data entry screens.

These interfaces illustrate the *case representation model implementation.*

Prior to describing the information that constitutes the case, the 'case author' must authenticate himself/herself using a dialog box (see Fig. 8). He/she fills in the headings to provide the health professional ID number, name, first name and rank (in particular is he/she an expert radiologist-senologist or a junior radiologist-senologist).

The 'case author' collects information on the radiologic-senologic process and the construction or the description of the new case. A 'case author' apprehends the description of a new case in the case base according to a specific step.

A thorough or partial information on the new case can be obtained. In this case, the 'case author' will be able to fill in up to four phases for the description of this new case. The description of a new case starts with the writing of general information on the case. The case author begins the description process by the information collection of the *"Clinical Examination"* phase which includes two steps: *"the interrogation"* and *"the physical examination".* The collected information is recorded in the case base. The process is then reiterated for another new case (see Fig. 9).

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:10, 2007

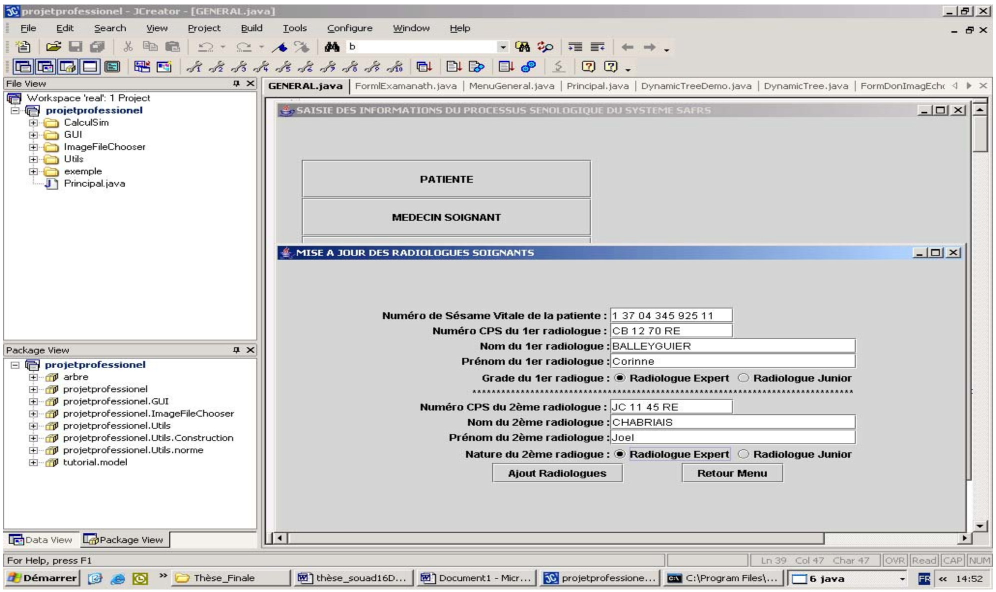

Fig. 8 Dialog box for the ‘case author’ identification

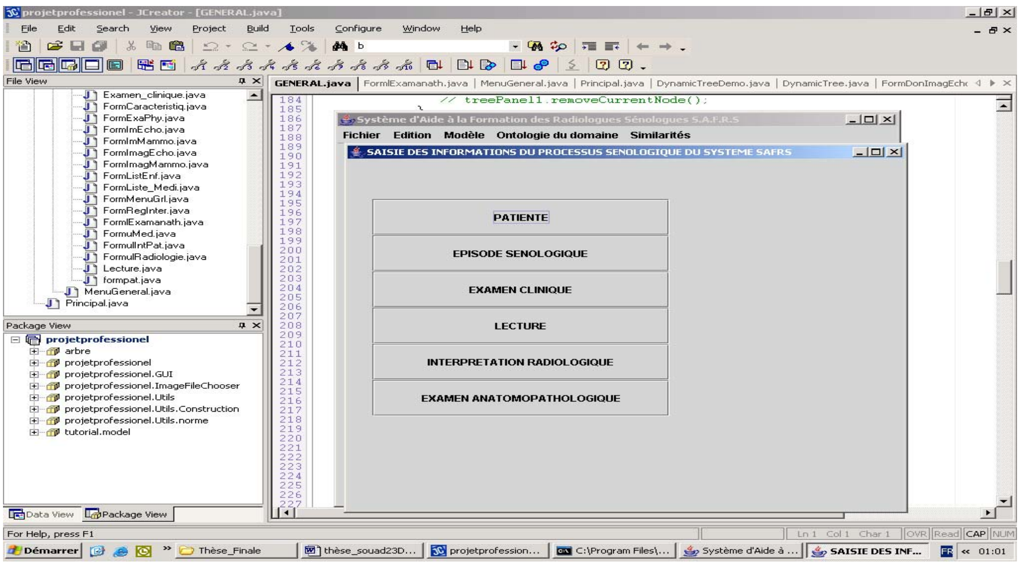

Fig. 9 Collection of information on the radiologic-senologic process

After having recorded the demographic and clinical data of the patient *("Clinical Examination" phase)*, the ‘case author’ imports the significant images of the case (mammographies and echographies) and visualizes them thanks to a graphic interface where the images under the DICOM format are stored: (*"Image Reading"* phase). The ‘case author’ then proceeds by observation to the extraction of relevant data on the image. This information is interpreted according to the author's expertise and the domain's glossary by using the ascensor (see Figure 10), which allows to keep the useful information that will subsequently be used in radiological interpretation *("radiological interpretation" phase).* If during the ("Radiological Interpretation" phase), the detected abnormality is of the BI-RADS 4 or BI-RADS 5 type, the ‘case author’ goes on and makes the description by taking into account the knowledge from the next phase: *("Anatomo-pathological Examination” phase)* and validates them. At the end of this last phase of the radiologic-senologic process, the ‘case author’ achieves the diagnosis and can finally make a statement on the patient's condition.

Once the case images are visualized, the radiologist-senologist uses his expertise and the glossary of the BI-RADS domain to extract the information on the images in the form of ROIs (Regions of Interest) that are formalized in the glossary using MCs (morphological characteristics).

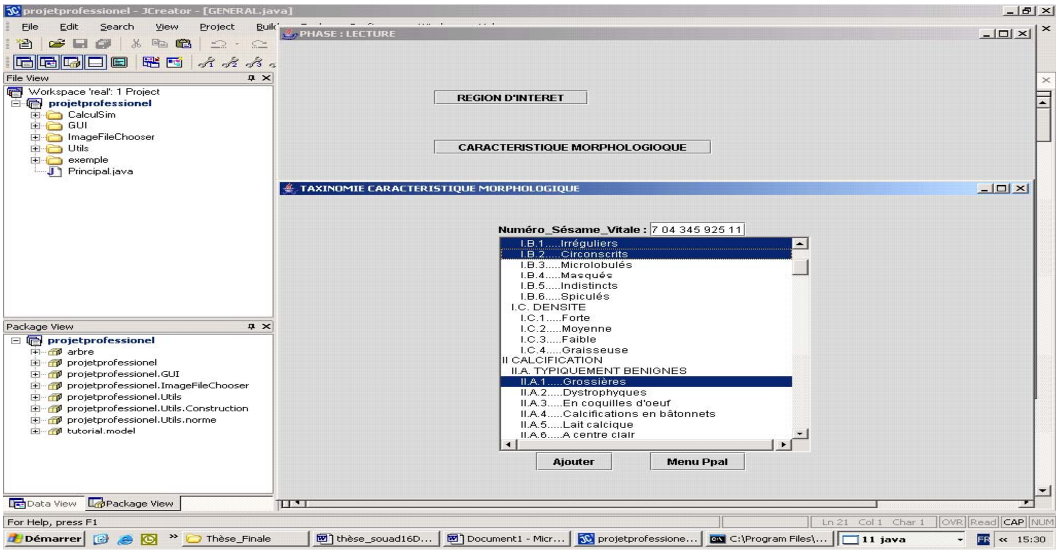

Fig. 10 Data entry screen showing the instantiation of the "morphological characteristics" (Reading Phase)

D. Implementation of the Description Module

The ‘*case description module’* of the prototype allowed to carry out a ‘knowledge base’ including 40 cases.

The *mammographic* and *echographic* files used for the achievement of the ‘knowledge base’ were provided by the Department of Radiology of the Necker Hospital. Out of 150 files, 40 were retained and digitized; then they were illustrated by 90 mammographic and echographic images. The films were digitized on the totality of their useful surface area, with a resolution of 42μm/pixel and a dynamics of 12 bits/pixel, respectively. We were given the collected information by Doctor Corinne Balleyguier, a radiologist-senologist.

E. The Retrieval Process Implementation

In this section we present the implementation of similarity measure algorithms of *the global strategy* and the instantiation of these algorithms with a real-use scenario.

General presentation of a session

The trainee starts a session by the description of the new case (target case). The system interrogates the case base. It retrieves data from the case base, thus, it chooses the case in the ‘case base’ and presents it to the trainee. The system asks the trainee to describe his/her case. The trainee answers the questions of the system. The system compares the ‘trainee’s answer’ with the ‘case base’s answer’. IF the ‘trainee’s answer’ = ‘case base’s answer’ THEN (cases are similar). It calculates similarity measures ELSE (cases are not similar). The system determines errors. The system determines errors either step by step**,** OR it asks the trainee to find the error.

The system checks after the evolution of the training. If the training progresses, then the system give to the trainee the possibility to follow the learning ELSE it changes the training strategy according to various strategies offered by the MAP.

The retrieval process

The interface of Fig. 11 illustrates the global retrieval process according to the *“global strategy”***.** The research of similar cases is performed according to 3 levels: *the sub-sub-case (elementary level), the sub-case ((the intermediate level) and the case (the global level).*

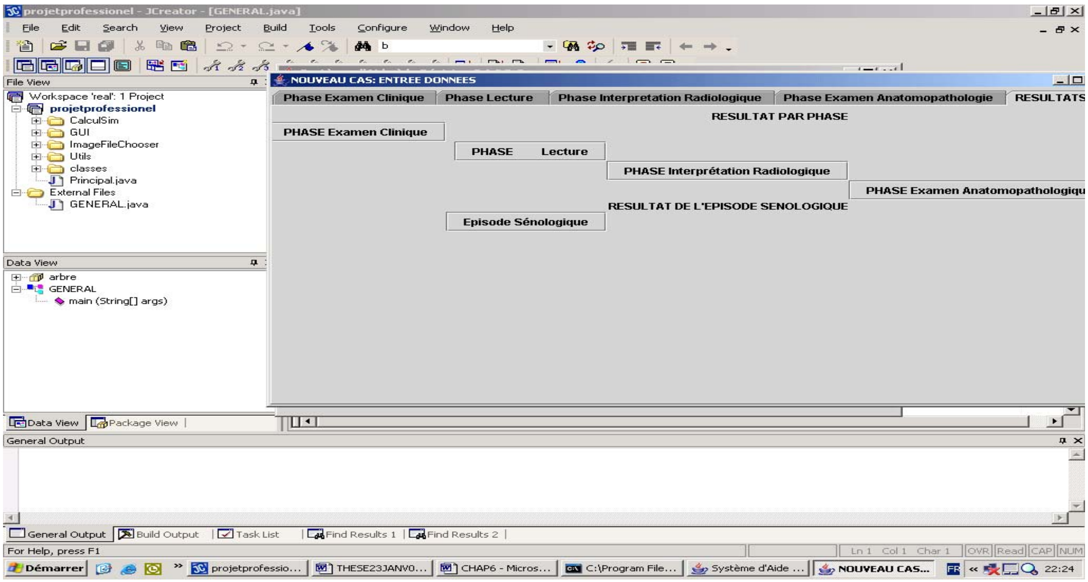

Fig. 11 The global retrieval process in the SAFRS system

The ‘CBR module’ starts the retrieval process by the instantiation of the new case (report of a new patient). The structure of the new case has the same structure as the old case already stored into the case base and the indexation is performed progressively. We store a new case into the ‘case base’. The “matching” is performed via a comparison of each new case with stored cases (initial case base). This step is performed apart from any interaction with the user (case author) and its implementation consists in the programming of “the matching algorithm”, according to the global strategy.

The retrieval process at the elementary level (research of elementary sub-sub-cases

The interface of Fig. 12 shows the starting of the course of the CBR process. We point out that the radiologic-senologic process includes 4 phases: *“Clinical Examination”, “Image Reading”, “Radiological Interpretation” and “Anatomo-pathological Examination”.*

The CBR process starts at the (‘image reading’ phase) (we don’t apply the CBR in the (‘clinical examination’ phase). Actually, this information is a useful tool for a radiologist-senologist in making the diagnosis, but comparisons with old cases are only based on data acquired from data provided by the radiological domain. According to the results obtained and the expertise acquired by the radiologist-senologist, the latter makes a decision.

After having stored demographic and clinical data concerning the patient, the radiologist-senologist acquires significant images of cases (mammographic and echographic ones) (see Fig. 13) and displays them with an icon of images where they are stored with the DICOM format (see Fig. 14). The radiologist-senologist retrieves relevant characteristics and compares them in pairs with the icon ‘‘COMPARER DONNEES IMAGE MEDICALE’’ (TO COMPARE MEDICAL IMAGING DATA) (comparison of data from problem parts of the *image reading phase (local similarity measure)* (see Fig. 15).

After having compared the knowledge provided by the “Image Reading Phase” and displayed the results of the comparison, the radiologist-senologist proceeds to the (“Radiological Interpretation” Phase). We point out that the problem of one phase is a solution for the previous phase. Data comparisons (problem parts) provided from the *(“radiological interpretation” phase)* concern the solution (extraction of relevant characteristics on images) of the previous phase, i.e. the “Image Reading Phase” (see Figure 13).

Here, the image of the “Image Reading Phase” is a mammographic type. The comparison between indexed physical data with the DICOM format consists in comparing different image profiles.

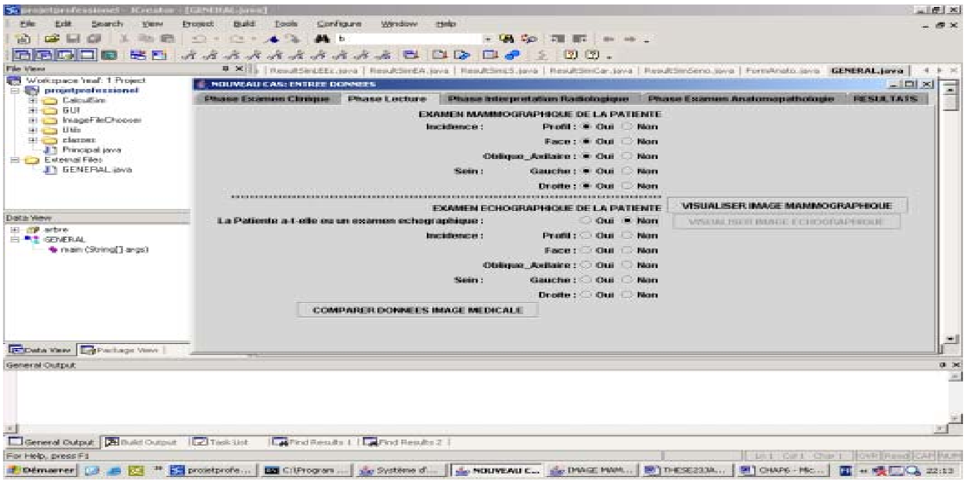

Fig. 12 Results of the comparison between the new and an old cases of the ('image reading' phase) (mammographic image)

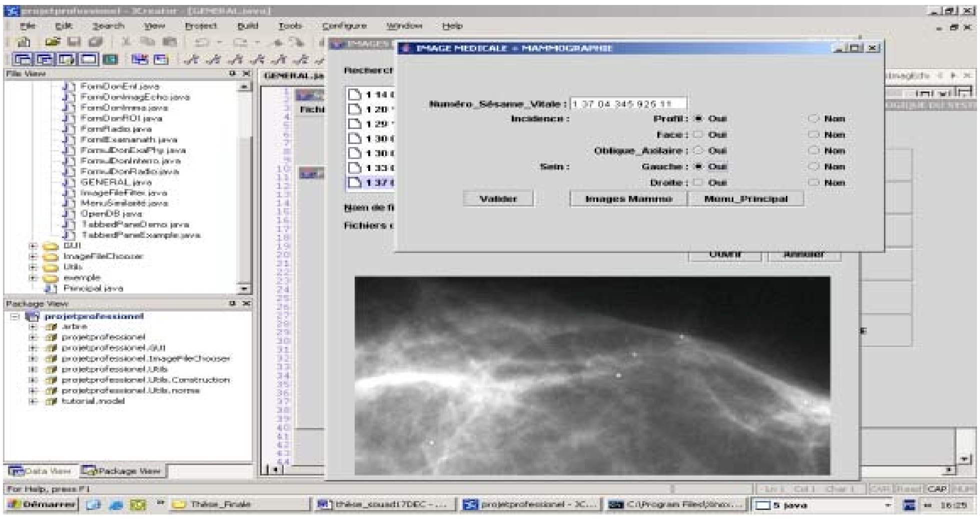

Fig. 13 Displaying mammographic images

This interface (see Fig. 14) shows values of comparisons of different similarities between attribute-values of the class image (mammographic type) (*local similarity measure*). Thereafter, *a global similarity measure* is computed according to the *local similarity measure* founded before. Here, the figure shows a similarity measure of 0.85 (a resemblance of 0.85 with the old case).

In this phase, the retrieval consists in browsing the domain glossary (hierarchy of radiological data) and selecting information chosen by the radiologist-senologist according to observations and acquired expertise. Thereafter they are validated and the radiologist-senologist selects again for this phase the icon "COMPARER DONNEES LECTURE SEMANTIQUE" (TO COMPARE SEMANTIC READING DATA), thus performing a matching. The results of this comparison are displayed via the icon "RESULTATS" (RESULTS). Finally, the solution of the comparison with the following interface (see Fig. 15) is displayed.

The results provided by the radiological interpretation phase have the type BI-RDAS2, the radiologist-senologist recommends a follow-up. The *("anatomo-pathological examination" phase)* is not treated in this case.

The interface (see Fig. 15) shows values of comparisons of different similarities between attribute-values of the class image (mammographic type) (*local similarity measure*).

After having compared data from different ROIs located on images *(local similarity measure and global similarity measure).* Afterwards, the process considers attribute-values of morphological characteristics (MC) by clicking on the button « Caractéristique Morphologique ».

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:10, 2007

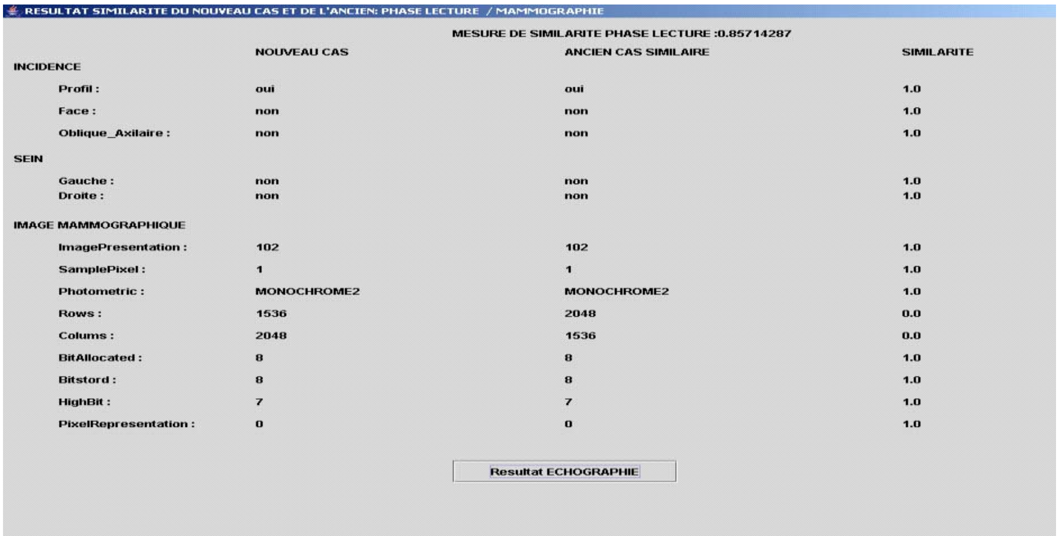

Fig. 14 Results of comparison of mammographic image data

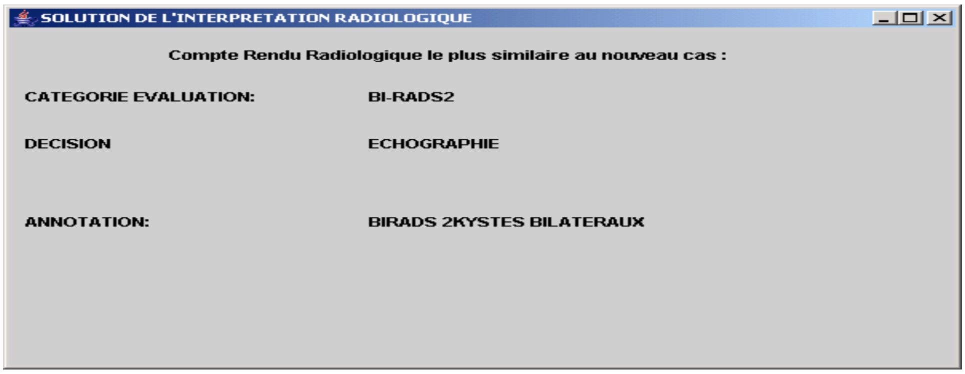

Fig. 15 Results of a comparison in the ("Radiological Interpretation" Phase): BI-RADS2 type

Fig. 16 Results of a comparison of a new case and an old case of ("the radiological interpretation" phase): MC

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:10, 2007

The retrieval process at the intermediate level (research of similar sub-cases)

Fig. 17 shows the comparison of two sub-cases of the radiologic-senologic process (at this level we consider previous reports of the patient).

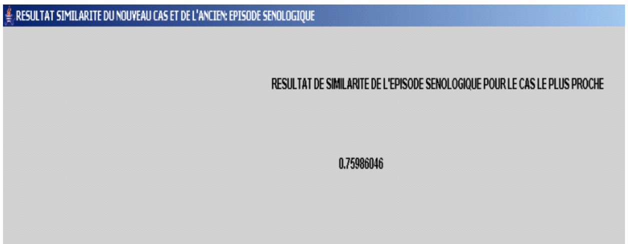

Fig. 17 Research of similar sub-cases

The retrieval process at the global level: the case (Research of similar cases)

According to the general algorithm of the SAFRS system, if two sub-cases are similar then the cases are similar. The similarity measure between the two sub-cases is in the order of

0.75. Then we can conclude that the two cases illustrated in the retrieval process with the global strategy are similar.

We have presented a part of a prototype implementation that will be considered as an evidence of the concepts put forward in this work. It displays a scenario of use of the overall system which helps understanding the concrete role that such a system can play in a real medical environment.

Section 9 provides a discussion on the strong and weak points encountered during the realization of the SAFRS system.

IX. Discussion and Future Prospects

In this paper we focused on the choice of an exclusively object-oriented implementation. That allows for a better evolution potential of the diagram of the case base that will be used as the grounding of the full system with connection to a PACS (Picture Archiving and Communication System) and to a hospital information system. Such a system will require regular modifications of the base diagram.

As regards the images, we choose the characterization of the contents of the *mammographic* and *echographic* images by textual data resulting from the interpretation of the images in the form of ROIs by the expert radiologists-senologists with the BI-RADS standard, which has the advantage of being an international standard of the radiologic-senologic activity. The image data are formatted with the DICOM standard. This makes it possible to benefit soon from the expected functionalities without requiring solutions much more intricate to implement and which are not yet routinely available (automatic extraction from the image of all types of abnormalities with their characterization). It implies that the 'case base' will be more suitable to its use by expert radiologists-senologists than by junior radiologists-senologists in view of teaching.

Currently, the development of the experimental prototype of the SAFRS is only partially achieved. It concerned the entire module of representation of cases; on the other hand, at the level of the CBR module, we have only achieved the first strategy developed with the model of the MAP (global strategy). Our first choice of development involved this strategy because it allows to compare the overall cases at once. The other strategies (elementary strategy and mixed strategy) are much more complex. At the same time they are rich and flexible, thus offering much more satisfaction to the user in the choice of the desired strategy to find the good cases, thus the best "score". The second phase shall enable us to validate the attainability of the model by expert radiologists-senologists of the Department of Radiology of the Necker Hospital.

First of all, our prospects are stated in terms of evolution of the prototype.

First, the next steps to be contemplated concern the software prototype. It deals with the design and development of adaptation and memorizing modules, in order to achieve the process of the CBR cycle, and to design and implement the module of the senologic knowledge management. Lastly, as a final prospect, we want to transpose the prototype on the Internet to be able to perform remote-training of radiologists-senologists on various platforms.

However, the *FastObject database* which has the advantage of being simple, effective and light has the drawback of managing only one customer at a time. We want to allow several radiologists to work in parallel. To eliminate this problem it is necessary to improve the prototype by integrating into *FastObjects* versant FastObject *(http://www.versant.net/fr_fr/).*

FastObjects brings transparent persistence in platform NET. With FastObjects.NET, Versant transposes the model of transparent management of data to the Microsoft.NET platform. A perfect integration with the development environment and an interface of convivial programming facilitate the direct storage of the objets.NET. Slope provides the first and only solution of transparent persistence for platform NET.

Moreover, Versant can be used to manage medical image data. Images from various sources (mammographies and echographies in our case) are transposed to an application for data acquisition. The application of data acquisition stores the meta-data associated with the images as well as the image versions. It is possible to store the images directly in Versant without passing by transition courses, as we did in the first version of the prototype.

The image data Server carries out two principal functions: first, it distributes the images in display stations (Viewing Stations) which question them and, second, it manages the repository, it migrates the old images from the hard disk to the *"tertiary media"* in the form of DLT *(Digital Linear Tape)* for the sake of filing and thus to update Versant with the new image localization. If the display station questions an image which was previously filed, the image data server controls the image retrieving mechanism starting from the previous DLT, updates Versant and in doing so it transfers the image to the display station.

Section 10 concludes the work presented in this paper.

X. Conclusion

The outcome of the research effort is both conceptual and practical. It has also a methodological dimension. Conceptually, this research contributes to the breast cancer diagnosis domain by defining a conceptual model for representing cases which are generic solutions reusable in many different settings. It contributes to the case reasoning field in two ways (a) by defining a case representation model which is not flat but organized at three different levels and (b) a search algorithm which exploits this layered structure to find similar cases by aggregating similar sub and sub-sub cases. At last, but not least, the research has a methodological contribution by which the retrieval algorithm is embedded in a broader process perspective including the capture of the actual case, the 'case base' reasoning related to this case and the support to decision making by adapting the retrieved case. An interesting aspect of the process model is its intentional dimension which makes possible the representation of different ways to achieve the result.

The last point is dealing with the implementation of a prototype that will be considered as evidence concepts put forward in this work. It displays a scenario of use of the overall system which helps understanding the concrete role that such a system can play in a real medical environment.

There are some limitations of the work: the composition of similar sub and sub-sub cases could have been more formally defined, the prototype needs improvements and to be embedded in the planned learning system.

Acknowledgment

I gratefully acknowledge the contributions of Professor Colette Rolland (Paris 1 University) and Doctor Corinne Balleyguier (Necker Hospital).

References

[1] T. P. Baum, "Aide à la détection des foyers de microcalcifications dans le cadre des campagnes de dépistage du cancer du sein," *thèse de Doctorat de l'Université Pierre et Marie Curie*, Paris VI, 1997.

[2] S. Demigha and C. Rolland, "Training-Aided System in Senology: Methodologies and Techniques," *conference Proceedings - SPIE Medical Imaging, PACS and Integrated Medical Information: Design and Evaluation*, vol 5033, pp. 339-349, San Diego, California, USA, 2003.

[3] A. Aamodt and E. Plaza, "Case-Based Reasoning: foundational issues, methodological variations and system approaches," *AI Communications*, vol 7, number 1, march 1994. C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.

[4] C. Le Bozec, "Gestion de connaissances multi-expertes en imagerie médicale (IDEM): Images et Diagnostics par l'Exemple en Médecine," *de doctorat en Informatique Médicale*, Université Pierre et Marie Curie, Paris VI, Juillet 2001.

[5] M. Grimnes and A. Aamodt, "A two Layer case-based reasoning architecture for medical image understanding," in *Advances in Case-Based Reasoning – Third European Workshop* (EWCBR-96), I.Smith, B. Faltings (eds), Lausanne, pp. 164-178. Springer Verlag, Berlin, 1996.

[6] J. Lieber and B. Bresson, "Case-Based Reasoning for Breast Cancer Treatment Decision Helping," in ***E. Blanzieri and L.Portinale, editors,*** *Advances in Case-Based Reasoning – proceedings of the fifth European Workshop On Case-Based Reasoning*, Lecture Notes in Artificial Intelligence vol 898, pages 173-185. Springer, 2000.

[7] R. T. Macura and K. J. Macura, "MacRad Radiology image resource with a case-based retrieval system," in *First International Conference on Case-based reasoning approach: Research and Development, M.Veloso, A.Aamodt*, (Eds), *Berlin: Springer Verlag:* 1995; 43-54.

[8] D. Leake, "Case-based Reasoning- Experiences, *Lessons and Future Directions*," *AAAI Press*/ *The MIT* Press, 1996.

[9] F. Cao, E. Sickles and H. K. Huang, "Interactive Digital Breast Imaging Teaching File," *Department of Radiology, University of California, School of Medicine,* San Francisco, CA, http://www.ej.rsna.org/EJ_0_96/003797.fin/mammo.html

[10] S. Caulkin, "Computer-Aided Training and Performance Evaluation in Mammography," *PHD Project*, UK, 1999.

[11] C. Lehaman, L. Miller, C. Rutter, "Effect of Training With the American College of Radiology Breast Imaging Reporting and Data System Lexicon on Mammographic Interpretation Skills in Developing Countries," *the Department of Radiology, University of Washington, Seattle, Cancer Care Alliance, WA,* University of California, San Diego, April 2001.

[12] M. Lanyi, "Morphologic Analysis of microcalcifications: a valuable differential diagnostic system for early detection of breast carcinomas and reduction of suspicious exploratory excisions," *Early Breast Cancer, Heidelberg: Spring Verlag*, 1995, 113-135.

[13] K. D. Fenstermacher, "An Application of Case-based Instruction in Medical Domains," in *Proceedings of the Spring Symposium on Artificial Intelligence in Medicine. Cambridge*, 1996, MA: AAAI Press/The MIT Press.

[14] S. Demigha, "An Ontology Supporting the Daily Practice Requirements of Radiologists-Senologists with the Standard BI-RADS," ICEIS 2007, *Ninth International On Enterprise Information Systems, Proceedings of Information Systems Analysis and Specification*, Vol ISAS, pp 243 249, Funchal, Madeira, Portugal, June, 12-16, 2007.

[15] S. Demigha, "Un système à base de connaissances en radiologie-sénologie dédié à la formation des radiologues-sénologues juniors," *Phd thesis in Informatic*, Université de Paris 1 Panthéon-Sorbonne, juillet 2005.

[16] A. L. Rector,,"Terminology and concept representation languages: where are we?," *in Artificial Intelligence in Medicine*, vol 15 (1): pages 1-4, 1999.

[17] J. Chabriais, C. Balleyguier, K. Kinkel, L. B Levey J. Stinès, and F. Thibault, "BI-RADS Illustré," *Première Edition Française basée sur la troisième Édition américaine,* 200 pages, 1998.

[18] C. Ben Achour, "Extraction des besoins par analyse de scénarios,*"Phd thesis*, Université de Paris 6, Pierre et Marie Curie, 1999.

[19] S. Demigha, "The retrieval Process in the SAFRS System with the Case-Based Reasoning Approach," ICEIS 2007, *Ninth International On Enterprise Information Systems, Proceedings of Artificial Intelligence and Decision Support Systems*, Vol AIDSS, pp 468-473, Funchal, Madeira, Portugal, June, 12-16, 2007.

[20] S. Demigha and N. Prat, "A case-based training system in radiology-senology," in *Information and Communication Technologies: from Theory to Applications (ICTTA) 04'*, Damascus, Syria, 2004.

[21] OMG (2003), OMG, *Unified Modeling Language Specification*, March 2003, version 1.5. http://www.omg.org/technology/documents/spec_catalog.htm.

[22] J. Lieber and A. Napoli, "Représentation par objets et classification pour le raisonnement à partir de cas," in ***M. Chein and F. Schmitt, editors,*** *Actes du 11ème congrès AFCET Reconnaissance des formes et intelligence artificielle*, Clermont-Ferrand, pages 345-354 (Tome III). AFCET –AFIA- Université de Clermont-Ferrand, 1998.

[23] CAD, *"Computer Aided Detection for Mammography,"* Medical Systems: www.cadxmed.com.

[24] DICOM, "Digital Imaging and COmmunication in Medicine," http://www.nema.org/press/020801.html.

[25] R. Bergmann and K. D. Althoff., "Methodology for Building CBR Applications," in: ***M. Lenz, B. Bartsh-Sport, H.-D. Burkhard, & S. Wess (Eds). Case-Based Reasoning from Foundations to Applications,*** Spring-Verlag, 1998.

[26] F. Fuchs, "Représentation des connaissances pour le raisonnement à partir de cas: le système Rocade," *thèse de doctorat en informatique*, *Université de Saint-Etienne*, 1997.

S. Demigha is a researcher at the C.R.I (Centre de Recherche en Informatique) at the Sorbonne-University (Paris 1), in Paris (France) where she has worked since 1999. She received her diploma of engineer in Informatics (Information Systems) from the Computer Science Institute of the University of Constantine in Algeria. In 1999, she joined the research team of Pr Colette Rolland, Professor at the Sorbonne University. From 2000 to 2005, she received her master and PhD in Informatics from the Sorbonne-University (Paris 1). Her research interests lie in the area of information modeling, databases, object-oriented analysis and design, requirements engineering, educational systems, medical imaging, Artificial Intelligence (the case-based reasoning approach) and data warehouses systems.