

A Knowledge-Based System for Patient Image Pre-Fetching in Heterogeneous Database Environments—Modeling, Design, and Evaluation

Chih-Ping Wei, *Member, IEEE*, Paul Jen-Hwa Hu, and Olivia R. Liu Sheng

Abstract—When performing primary reading on a newly taken radiological examination, a radiologist often needs to reference relevant prior images of the same patient for confirmation or comparison purposes. Support of such image references is of clinical importance and may have significant effects on radiologists' examination reading efficiency, service quality, and work satisfaction. To effectively support such image reference needs, we proposed and developed a knowledge-based patient image pre-fetching system, addressing several challenging requirements of the application that include representation and learning of image reference heuristics and management of data-intensive knowledge inferencing. Moreover, the system demands an extensible and maintainable architecture design capable of effectively adapting to a dynamic environment characterized by heterogeneous and autonomous data source systems. In this paper, we developed a synthesized object-oriented entity-relationship model, a conceptual model appropriate for representing radiologists' prior image reference heuristics that are heuristic oriented and data intensive. We detailed the system architecture and design of the knowledge-based patient image pre-fetching system. Our architecture design is based on a client-mediator-server framework, capable of coping with a dynamic environment characterized by distributed, heterogeneous, and highly autonomous data source systems. To adapt to changes in radiologists' patient prior image reference heuristics, ID3-based multidecision-tree induction and CN2-based multidecision induction learning techniques were developed and evaluated. Experimentally, we examined effects of the pre-fetching system we created on radiologists' examination readings. Preliminary results show that the knowledge-based patient image pre-fetching system more accurately supports radiologists' patient prior image reference needs than the current practice adopted at the study site and that radiologists may become more efficient, consultatively effective, and better satisfied when supported by the pre-fetching system than when relying on the study site's pre-fetching practice.

Index Terms—Data/knowledge modeling, knowledge-based system, patient image pre-fetching, patient image retrieval, synthesized object-oriented entity-relationship model.

Manuscript received August 27, 1999; revised April 13, 2000. This work was supported in part by the Medical Systems Division, Toshiba Corporation, Japan, and by the Sun Microsystems Computer Corporation.

C.-P. Wei is with the Department of Information Management, College of Management, National Sun Yat-Sen University, Kaohsiung 804, Taiwan, R.O.C. (e-mail: cwei@mis.nsysu.edu.tw).

P. J.-H. Hu is with the Department of Accounting and Information Systems, David Eccles School of Business, University of Utah, Salt Lake City, UT 84112 USA.

O. R. L. Sheng is with the Department of Management Information Systems, College of Business and Public Administration, University of Arizona, Tucson, AZ 85721 USA.

Publisher Item Identifier S 1089-7771(01)01714-9.

I. INTRODUCTION

ADVANCED medical imaging technology has significantly expanded the role of radiology in clinical medicine, as manifested by increasing dependence of physicians and specialists on information obtained from radiological examinations for clinical decision making and patient management. At the same time, concurrent information technology (IT) advancements have initiated and greatly facilitated a progressive transformation of radiology from the traditional film-based practice into a digital practice where radiologists can perform examination reading and consultation services beyond organizational and geographical constraints. Jointly, the growing role of radiology and the observed trend toward a digital radiology practice have made patient image management a growing concern for many healthcare organizations. One critical image management issue is prior patient image reference support for radiologists' examination reading.

A fundamental role of radiologists in clinical medicine is to provide physicians or specialists reports containing timely patient information obtained from image interpretation, together with appropriate recommendations for subsequent radiological investigations [5], [9]. Typically, a radiologist's examination reading proceeds as follows. The reading radiologist first orients himself or herself to the case using the patient and examination information provided, including the referring physician's clinical questions, and then scans the current examination images to generate a set of hypotheses, based on suspicious or interesting radiographic signs detected during the image scan. Subsequently, the reading radiologist validates these hypotheses by making reference to relevant prior images of the same patient and other referential information and revises the preliminary hypotheses accordingly. A radiologist has specific patient prior image reference patterns and often uses his or her heuristic knowledge to determine the relevance of each prior patient image with respect to the current examination. This hypothesis-generation-validation process usually proceeds by iteration, terminated when the validated hypotheses have reached a subjectively determined certainty threshold. As such, the hypotheses then become examination reading findings, which will be compiled as a formal report, summarizing the radiologist's impression of and diagnostic opinion about the examination with appropriate follow-up-examination recommendations.

Support of patient prior image references is critical to a radiologist's examination reading, providing a baseline essential for confirmation, comparison, and/or evaluation of the suspicious radiographic signs detected on the current examination images. Such support is of clinical importance and may have significant effects on radiologists' examination reading efficiency, work satisfaction, and service quality [5]. To alleviate time and physical demands on radiologists and support their patient image reference needs, many healthcare organizations have adopted a pre-fetching strategy that selects a set of patient images presumably relevant to a current examination reading task and makes them available to the radiologist in advance of an examination reading.

To a large extent, the heuristic nature of radiologists' prior image reference knowledge makes a knowledge-based approach intuitively sound and appropriate for patient image reference support. Several knowledge-based approaches to supporting patient image retrieval and pre-fetching have been explored [7]–[11], [20]. However, most prior research focused on knowledge acquisition and knowledge base (KB) construction. Although a KB is essential to knowledge-based pre-fetching, the resulting system needs to address several challenging design issues. For instance, patient image reference heuristics reflect not only pertinent medical and radiological knowledge, but also individual practice styles and preferences. Thus, patient image reference heuristics may vary with individual radiologists. Such heuristics may dynamically evolve over time as a radiologist accumulates additional knowledge, training, and clinical experience. In combination, inter- and intra-radiologist variations make the construction, customization, and maintenance of a patient image reference KB challenging. Furthermore, patient image reference heuristics are highly data intensive, encompassing a sizable set of essential attributes (decision factors). Such important data attributes as patient's gender and condition need to be extracted from various data sources, including the Hospital Information System (HIS) and Radiology Information System (RIS). Considerable data intensity makes knowledge representation challenging. Data referenced by image reference heuristics are stored in multiple database systems that are often heterogeneous in schema representation, data model, and query language. As a result, a patient image pre-fetching system needs to access information from multiple heterogeneous and autonomous data sources during its knowledge inferencing process. Thus, an effective mechanism for coping with system heterogeneity and autonomy is necessary and essential.

The organization of the remainder of this paper is as follows. Section II analyzes and discusses the requirements and challenges unique to radiologists' patient prior image retrievals and pre-fetching support. Section III describes our conceptual modeling of the target application, as well as the representation of image reference heuristics and their underlying semantic knowledge, with appropriate illustrations. Section IV details our overall system architecture, including description of specific designs for the respective components of the proposed patient image pre-fetching system. Section V highlights some interesting preliminary system evaluation results and discusses their implications for both radiologists' examination reading and patient image management in healthcare organizations. A conclu-

sion is presented in Section VI, inclusive of a summary and some future research directions.

II. APPLICATION CHARACTERISTICS AND SYSTEM DESIGN REQUIREMENTS

In a typical examination reading process, the reading radiologist applies his or her heuristic knowledge to determine the relevance of each prior image of the same patient with respect to the current examination. Such heuristics may vary with individual radiologists and can dynamically evolve over time. Moreover, patient prior image reference heuristics usually require such information as patient demographics and clinical history, examination information, and the requesting physician's clinical questions, stored in various data repositories that include the HIS and RIS. These characteristics represent challenging requirements that need to be properly addressed in the design of a knowledge-based patient image pre-fetching system.

Such a system needs to support efficient and effective customization. Largely determined by pertinent medical and radiological knowledge, patient image reference heuristics, to some extent, is also dependent on individual practice styles and preferences. For example, while a post-heart-transplant patient's recovery and potential complications largely follow some defined and documented physiological processes, radiologists nevertheless develop individual heuristics that determine their specific image reference patterns. Conceivably, one radiologist may consider a patient's age an important factor and, when a patient belongs to a specific age group, will reference prior images taken at a particular time interval, say, 24 h. Another radiologist may consider the age factor irrelevant and use one interval across different age groups. Thus, efficient and effective customization of the patient prior image reference KB is essential for adequate patient image reference support. In addition, radiologists' patient image reference heuristics may subtly evolve over time, necessitating continual knowledge updates or replenishments. Inductive learning techniques appear to be an appropriate approach to the needed KB customization and maintenance. Specifically, an adequate learning system can automatically extract significant patient image reference knowledge from logged examination-reading cases, including both input attributes (decision factors) and actual image reference behaviors (decision outcomes), and update the existing KB without intervention of knowledge engineers or radiologists.

To be effective, an inductive learning technique needs to encompass the required tolerance for inconsistent decision outcomes, immunity to missing data, and flexibility in managing multiple decision outcomes. An effective learning system needs to cope with inconsistent image references resulting from the inter- and intra-radiologist variations already discussed. Values of input attributes are not always guaranteed and may be incomplete. Some attributes are applicable to some, but not all examination reading cases. The value of a specific input attribute in an examination reading case may be unavailable or difficult to derive from the patient and examination information provided. When interpreting a current examination, a reading radiologist may reference patient images from multiple prior examinations. In effect, the tendency to reference patient images from multiple

prior examinations is prevalent and often observed in radiologists at work. Jointly, these challenging characteristics demand an inductive learning technique capable of managing missing or incomplete input data, as well as potentially inconsistent and multiple decision outcomes [9], [10], [19].

From the system design perspective, an effective knowledge-based patient image pre-fetching system needs to encompass a seamless integration of knowledge inferencing and data access. Radiologists' patient image reference heuristics are highly data intensive. For example, when following up on a patient who has a benign tumor and is in satisfactory condition, a reading radiologist may reference images from the most recent prior examination, the last normal examination, and all prior examinations taken between them at three-month intervals. To identify which patient prior images are appropriate to be pre-fetched for a scheduled current examination, the patient image reference heuristics maintained by the KB need to be utilized, and relevant information, including the reason for examination for the current examination and patient condition, needs to be extracted from the respective data repositories. Hence, a knowledge-based patient image pre-fetching system must be capable of supporting knowledge inferencing and data access interaction. Furthermore, modeling of a radiologist's patient prior image reference knowledge is needed to properly represent his or her problem-solving heuristics and the underlying data semantics.

Ideally, a knowledge-based patient image pre-fetching system should shield knowledge inferencing from the underlying system heterogeneity and, at the same time, preserve the autonomy of individual systems where data reside. During the knowledge inferencing process, the pre-fetching system needs to access data from multiple sources, which often are different in schema representation, data model, or query language. Such heterogeneity needs to be transparent to the knowledge inferencing. Autonomy of local data sources is also an issue of concern. The implementation of a knowledge-based patient image pre-fetching system should not change the database systems involved or their existing applications. To achieve the desired heterogeneity transparency and autonomy preservation, an integrated schema ensuring adequate integration of the schemata of the related data sources needs to be maintained and made available to the pre-fetching system. Furthermore, during the knowledge inferencing process, the pre-fetching system needs a mechanism for decomposing a data access request into an equivalent set of sub-queries and translating each of them into a language understandable to the respective autonomous database system where the resulting sub-query will be executed.

A desirable knowledge-based patient image pre-fetching system needs to be extensible. Each underlying data source is by no means static and may unilaterally change its schema or implementation, including data model and query language, in response to local operations requirements. Effects of such changes in local data sources on the pre-fetching system should be eliminated or minimized by making the management of the discussed dynamics in data sources and local changes extensible.

In sum, a knowledge-based patient image pre-fetching system supporting radiologists' examination reading must address several challenges, including representation and learning of image

reference heuristics, and management of data-intensive knowledge inferencing. Together, these challenges and other application requirements demand an extensible and maintainable architecture design capable of effectively adapting to a dynamic environment characterized by heterogeneous and autonomous data source systems. The following section describes our conceptual model for representing the complex heuristics and data intensive knowledge of the target application.

III. CONCEPTUAL DATA/KNOWLEDGE MODELING AND REPRESENTATION

Modeling of the data-intensive image reference heuristics of radiologists requires proper representation of the heuristics as well as the underlying semantic knowledge concerning structural and behavioral aspects of the relevant data. The heuristics mimic the high-level problem-solving expertise of radiologists and are formulated using pertinent semantic knowledge. As such, an effective conceptual model for the target application needs to support the extrapolation of high-level heuristics using low-level semantic knowledge and, at the same time, provide the necessary access paths to relevant semantic knowledge.

We therefore developed a synthesized object-oriented entity-relationship (SOOER) model [18], which is appropriate for representing the required knowledge, including patient image reference heuristics and the underlying semantic knowledge. This model synthesizes and extends relevant concepts from object-oriented and entity-relationship modeling techniques, two dominant and potentially complementary modeling paradigms. To make the model more communicative, we retain the conventional notations for the respective models when possible. In a SOOER model, semantic knowledge is represented by object classes, whereas heuristic knowledge is represented with IF-THEN production rules encapsulated within pertinent object classes. The following describes the modeling constructs essential for representing semantic and heuristic knowledge.

A. Constructs for Semantic Knowledge

Semantic knowledge describes the structural and behavioral aspects of data, including entities, attributes, relationships, and behaviors. The primary constructs of a SOOER model include entity and relationship classes. Encapsulated in an entity or relationship class are attributes describing the properties of the class and methods defining its behavior. The graphical notations to represent semantic knowledge in a SOOER model are summarized in Fig. 1.

Entity Class: An entity class is an abstraction of a group of objects that share common characteristics (i.e., attributes), behavior (i.e., methods), and relationships with other objects. The identity of an entity class is implemented using a subset of its attributes.

Relationship Class: A relationship class is a structural (or logical) connection between or among entity classes. In a SOOER model, three types of relationship classes are supported: specialization, aggregation, and association. Each type of relationship classes has distinctive purpose and semantics, described as follows.

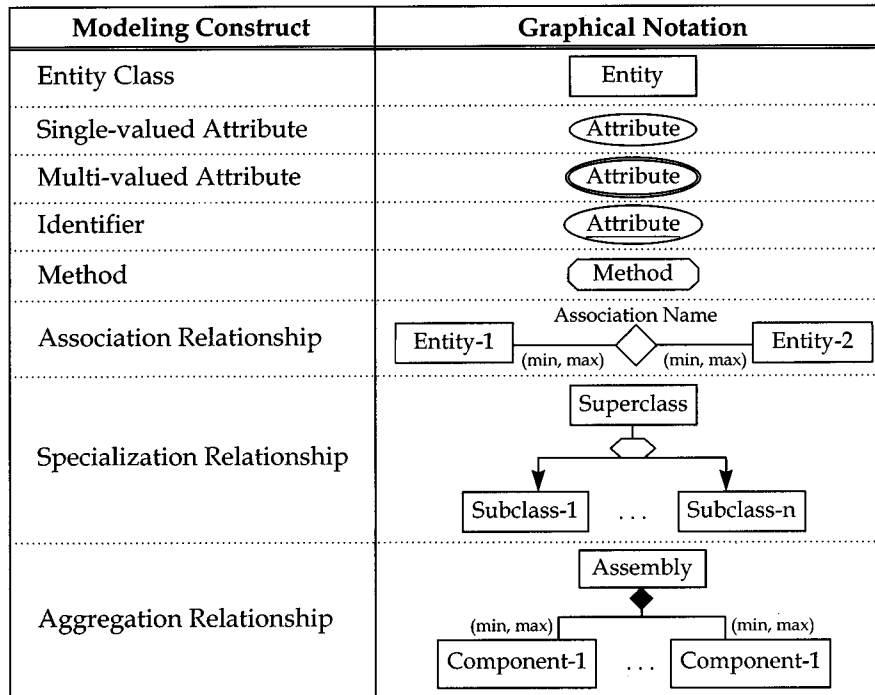


Fig. 1. Graphical notations of a SOOER model constructs for representing semantic knowledge.

Specialization Relationship Class: A specialization relationship class categorizes a general entity class (i.e., superclass) into two or more specialized entity classes (i.e., subclasses). One essential mechanism in specialization relationship classes is inheritance, through which a subclass inherits properties, including attributes, methods, and relationships from its superclass.

Aggregation Relationship Class: An aggregation relationship class can be described as an entity class that is a component of another entity class (assembly entity class); i.e., it shares an *a-part-of* relationship. An aggregation relationship class usually encompasses the *existence dependency* between the assembly entity class and its component entity classes. This dependency dictates deletions of the component objects when the object in which they participate is removed from the assembly entity class. Another important property of aggregation relationship classes is *operation propagation*, enforcing an operation performed on an assembly object to be propagated to all of its component objects. Participation of a component entity class in an aggregation relationship is jointly described by minimal and maximal cardinalities. Minimal cardinality specifies the minimum number of objects from a component class that must take part in an assembly object (contained in the assembly entity class). Similarly, maximal cardinality defines the upper bound of such participation.

Association Relationship Class: An association relationship class can be unary, binary, or *n*-ary, where *n* otherwise independent entity classes are related to one another. Each entity class participating in an association relationship class assumes a distinctive role in the relationship. The role name can be omitted in situations where the role of a participating entity class is unambiguous. Similarly to aggregation relationship classes, participation of an entity class in an association relationship class is confined by minimal and maximal cardinalities. Minimal cardi-

nality specifies the minimum number of objects from an associated entity class that must participate in association relationship instances. On the other hand, maximal cardinality defines the participation ceiling, restricting the maximum number of objects from the entity class that are allowed to take part in association relationship instances.

Attribute: An attribute describes a property of a class, entity, or relationship. An attribute can be atomic or composite. A composite attribute can be further decomposed into a set of sub-attributes, each of which is atomic or composite. An attribute can be single-valued or multivalued. A single-valued attribute can assume at most one value for the attribute, whereas a multivalued attribute can take multiple values simultaneously. Attributes uniquely identify individual objects of an entity or relationship class. Specifically, one or a combination of several attributes of a class serves as a unique identifier for individual instances of the class when the attribute value or values are unique across all the objects contained in the class.

Method: Collectively, methods describe the behavior of a class and, consequently, define the behavior of all the objects in the class. A method can be applied at two different levels, object and class, and, therefore, can be classified accordingly. An object method applies to individual objects of a class, whereas a class method is applied for the class as a whole.

The following illustrates partial modeling of the target application, as a result of integrating the schemata of multiple relevant data sources, including the HIS, RIS, and picture archiving and communication system (PACS). The specific semantic knowledge needed by a knowledge-based patient image pre-fetching system includes data about the patient, examination, image, anatomical portion, and modality. A patient may have many prior radiological examinations, each of which belongs to one and only one patient. A patient has

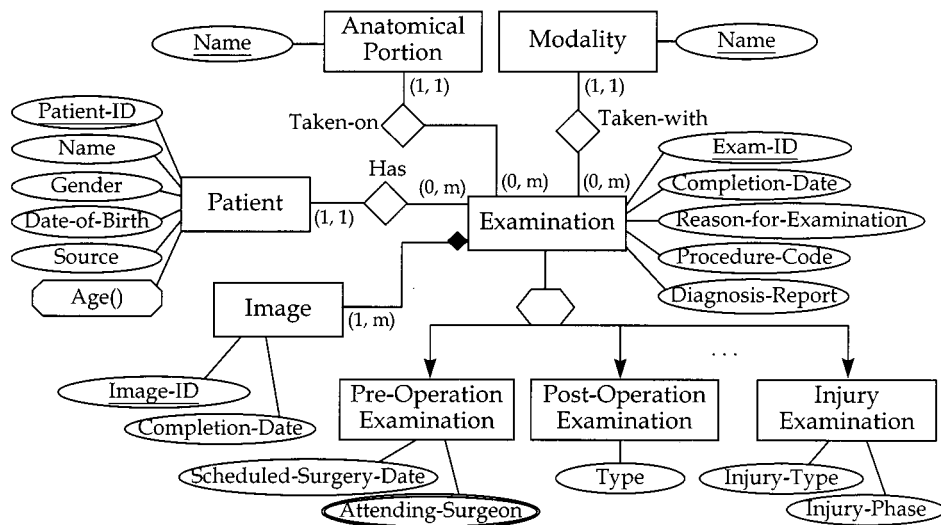


Fig. 2. Conceptual modeling of semantic knowledge in a knowledge-based patient image pre-fetching system.

a unique patient identification (ID), name, gender, date of birth, and patient source (e.g., intensive-care unit or outpatient clinics). A method is needed to calculate patient age and can be defined within patient entity class. An examination usually contains such information such as examination ID, completion time, reason-for-examination, procedure code, and diagnosis report. An examination consists of one or multiple images, each of which has a specific anatomical portion (part) examined and is taken using a particular imaging modality. Based on their reasons, examinations can be classified into different subclasses that include pre-operation, post-operation, and injuries. Fig. 2 graphically depicts conceptual modeling of the described semantic knowledge in SOOER model.

B. Constructs for Heuristic Knowledge

Due to their simplicity, communicability, and expressiveness, IF-THEN production rules were chosen to represent heuristic knowledge. Each class has a set of rules that describe the specific heuristics pertaining to the class. These rules are constructed with relevant semantic knowledge and, therefore, require the establishment of a reference mechanism in the schema represented with a SOOER model.

Definition: Instance Variable: An instance variable provides a means of referencing an object in a class, entity, or relationship. Instance variable *this* is a special instance variable, denoting the particular object that is currently being processed.

Definition: Path Expression: Let *t* be an instance variable that references an object in the class *T*. A path expression $t.A_1.A_2, \dots, A_n$, where *t* denotes the origin of the path and A_n is the terminal of the path, refers to a path in a SOOER schema and satisfies the following constraints for each $j \in \{1, \dots, n\}$.

- 1) If A_j is an attribute, then A_j is the terminal of the path or A_{j+1} is a sub-attribute of A_j .
- 2) If A_j is a method, then A_j is the terminal of the path.
- 3) If A_j is an entity class, then there must exist A_{j+1} that is either an attribute of A_j , a method of A_j , an entity class (if there exists only one relationship between A_j and A_{j+1}), or a relationship in which A_j participates.

- 4) If A_j is an association relationship class, then there must exist A_{j+1} that is an attribute of A_j , a method of A_j , or an entity class participating in A_j .
- 5) If A_j is a specialization or aggregation relationship class, then there must exist A_{j+1} that is an entity class participating in A_j .

Jointly, these constraints suggest the following.

- 1) A path starts from an instance variable.
- 2) The terminal of a path is either an attribute or a method, but cannot be a class, entity, or relationship.
- 3) When the terminal of a path is a method, the path may return a single value or a set of values. Similarly, a path may return a single value or a set of values when it is terminated with an attribute and the specific number of values returned depends on the maximal cardinality of the respective entity classes, as well as the multiplicity of the attributes on the path. If the maximal cardinality of concern is one and all of the attributes involved in the path are single valued, the path will return a single value; otherwise it will return a set of values.

In effect, the path expression $t.A_1.A_2, \dots, A_n$ is an abbreviated form in which the role name of each entity class participating in the subsequent association relationship class is omitted. To fully express an association relationship class on a path, the preceding and subsequent entity classes of the association relationship need to be specified with the respective role names, singling out how the path traverses through this association relationship class. That is, if $A_{j-1}.A_j.A_{j+1}$ is a part of the path where A_j is an association relationship class and A_{j-1} and A_{j+1} are entity classes, the full expression becomes $A_{j-1}.role_{j-1}.A_j.role_{j+1}.A_{j+1}$, where the preceding entity class A_{j-1} takes the role $role_{j-1}$ and the subsequent entity class A_{j+1} takes the role $role_{j+1}$. However, role names need not be signified in a relationship class where all the participating entity classes are distinct.

Definition: Function Expression: Path expressions can be manipulated by proper functions to derive aggregate information of interest. Each function follows a particular format, func-

tion-name(path). Examples of common functions include $\min()$ that returns the minimal value of a value set, $\max()$ that returns the maximal value of a value set, $\text{sum}()$ that returns the summation of a numeric value set, $\text{avg}()$ that returns the average of a numeric value set, $\text{count}()$ that returns the number of elements in a set, and $\text{first-}n()$ that returns the first n elements in a set.

Based on the discussed reference mechanism, heuristics formulated with intensive data can be formally defined using the following language.

```
rule := IF path rel-op [constant | constant-set | path]
      <<(AND path rel-op [constant | constant-set | path])>>
      THEN [var-assignment | data-retrieval]
      <<(AND [var-assignment | data-retrieval])>>
var-assignment := var = [constant | constant-set]
data-retrieval :=
  SELECT [path | func-exp] <<(, [path | func-exp])>>
  FROM class-name instance-variable
      <<(, class-name instance-variable)>>
  (WHERE path rel-op [constant | constant-set | path]
    <<(AND path rel-op [constant | constant-set | path])>>
    (ORDER BY path [ASC | DESC] <<(, path [ASC |
  DESC])>>
rel-op := <| ≤ | = | >| ≥ | ≠ | ∈ | ∉ | ⊂ | ⊆ | ⊃ | ⊇ | ⊄
constant-set := {constant <<(, constant)>>}
```

Annotations:

$:=$ is defined as

$[a | b]$ denotes either a or b

$()$ denotes the enclosed item repeats one or more times

$\langle \rangle$ denotes the enclosed item is optional

Comments:

func-exp: the function expression on a path

var: an intermediary variable derived during the knowledge inferencing

ASC or DESC: ascending or descending order

In the proposed representation, the IF clause specifies the condition to be evaluated and the THEN clause defines the action to be performed when the associated condition is satisfied. The THEN clause can be used to represent the derivation of an intermediary result, as well as a data retrieval operation expressed in structural query language (SQL) with an appropriate path expression extension. In the latter case, the SELECT sub-clause in the THEN clause lists the data to be retrieved, the FROM sub-clause specifies the data source(s), the WHERE sub-clause qualifies the data to be retrieved, and the ORDER BY sub-clause determines the presentation sequence of the retrieved data.

Example 1: The following hypothetical example illustrates the proposed heuristic knowledge representation. A radiologist scheduled to read images from a radiological examination of a seven-year-old patient (or younger) needs to compare current images with those from all prior examinations of the same patient. Based on the SOOER schema shown in Fig. 2, this image reference heuristic can be encapsulated in examination entity class and represented the following.

```
IF this.Patient.Age() ≤ 7
THEN SELECT e.Exam-ID
```

```
FROM Examination e
WHERE e.Patient.Patient-ID = this.Patient.Patient-ID
```

Example 2: To continue with our illustration, we further assume that the radiologist has the following additional image reference heuristics specifically related to examinations taken to follow up patient recovery from a surgery: if the current examination is for “urgent” post-operation follow-up, retrieve patient’s images from the two most recent prior examinations in which anatomical portion examined and imaging modality used were identical to those of the current examination. As described in Fig. 2, this heuristic is encapsulated in the entity class named “post-operation examination” and can be represented as follows.

```
IF this.Type = “Urgent”
THEN SELECT First-2(e.Exam-ID)
FROM Examination e
WHERE e.Patient.Patient-ID = this.Patient.Patient-ID
AND
  e.Anatomical-Portion.Name = this.Anatomical-Portion
AND
  e.Modality.Name = this.Modality.Name
ORDER BY e.Completion-Date DESC
```

IV. ARCHITECTURAL DESIGN

We designed and developed a knowledge-based patient image pre-fetching system based on the requirements described in Section II. As shown in Fig. 3, the system retrieves relevant information from multiple distributed and heterogeneous database systems and performs data-intensive knowledge deduction essential to supporting radiologists’ patient prior image reference needs. To address the challenges of heterogeneity, dynamics, and local autonomy preservation of the underlying database systems, we adopted a client–mediator–server system architecture, in which the knowledge-based patient image pre-fetching system is the client and the underlying database systems assume the role of servers. A mediator creates a desirable appearance of single data source and presents a uniform interface to the client, shielding the underlying data source distribution and heterogeneity from the patient image pre-fetching system and, at the same time, resulting in desired transparency and necessary local autonomy preservation. The adopted three-tier architecture also provides desirable independence to the patient image pre-fetching system, making it immune, or at least less susceptible to problems resulting from data source dynamics. The following describes our system architecture and the design of the mediator and the knowledge-based patient image pre-fetching system, as shown in Fig. 3.

A. Mediator

The mediator provides distribution and heterogeneity transparency to the knowledge-based patient image pre-fetching system by maintaining the semantic knowledge (i.e., global database schema) through integrating schemata of the participating local sources. We used a SOOER model to represent the semantic knowledge (stored in a catalogue) and the

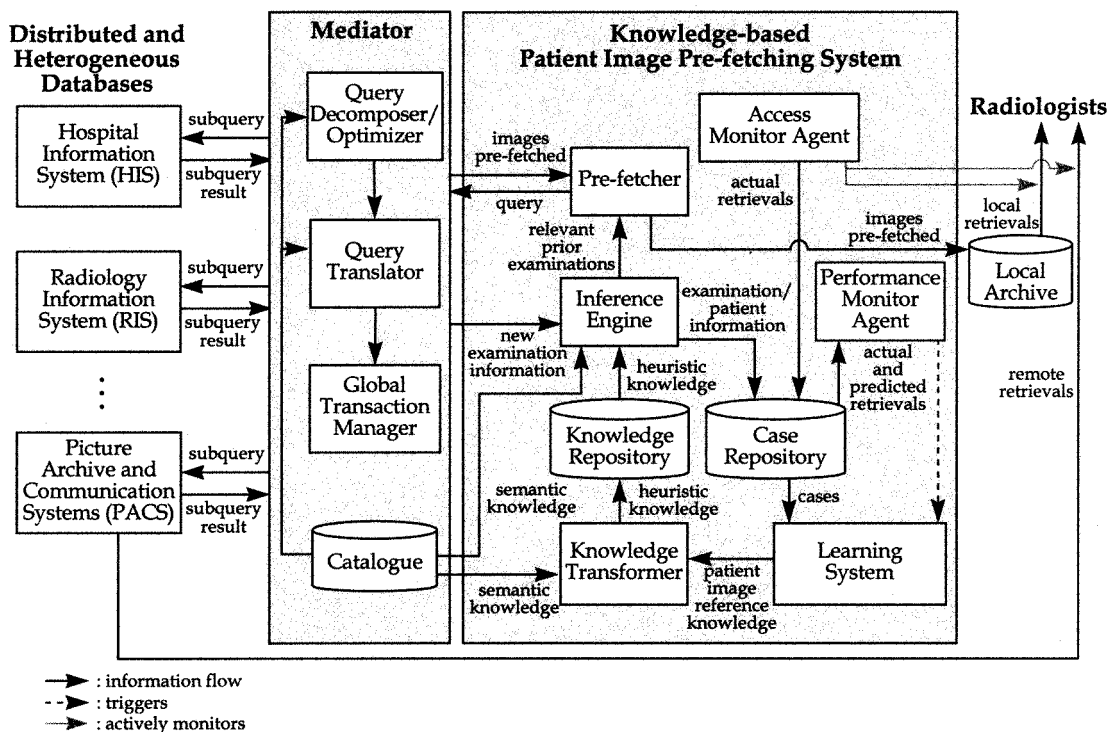


Fig. 3. Architecture of mediator and knowledge-based patient image pre-fetching system.

problem-solving heuristic knowledge (maintained in a knowledge repository). The representation of choice facilitates the construction of radiologists' patient image reference heuristics without worrying about the representations and retrieval operations of the underlying data sources. During the subsequent knowledge inferencing session performed by the pre-fetching system, the mediator uses the semantic knowledge to formulate a query submission and expresses it in a SQL-like language with a path expression extension. At this time, a query decomposer/optimizer decomposes a submitted query into a set of sub-queries intended for the respective data sources, generates and evaluates alternative execution strategies, and selects the most efficient one [3], [6], [14]. The resulting sub-queries are then submitted to a query translator, which translates each sub-query into a language understandable by the target information system or database system where the sub-query will be executed [14]. Furthermore, a global transaction manager communicates with the respective local transaction managers to oversee and maintain the necessary consistency during the execution of the set of sub-queries over different participating systems and databases [4], [12], [13].

The catalogue is essential to the operations of the mediator, as well as the knowledge-based patient image pre-fetching system. Typical information maintained by the catalogue includes the semantic knowledge (i.e., global database schema), local database schemata, mappings between the global and the local database schema, and important statistics on local databases. For example, the query decomposer/optimizer needs relevant information about the global database schema and its mappings with related local database schemata to perform query decomposition. Moreover, the query decomposer/optimizer depends on

local database statistics in selecting the most efficient execution strategy [3], [6], [14]. At the same time, the knowledge-based patient image pre-fetching system needs semantic knowledge to perform the necessary reasoning about radiologists' patient prior image reference heuristics, as discussed in the following subsection.

B. Knowledge-Based Patient Image Pre-Fetching System

To facilitate KB customization and maintenance, the knowledge-based patient image pre-fetching system needs to monitor and adapt to radiologists' patient image reference behaviors proactively. Furthermore, the pre-fetching system also needs to monitor the accuracy of its image reference heuristics continuously because radiologists' image needs may evolve over time. In situations where the existing heuristics cannot effectively support radiologists' examination reading tasks, the pre-fetching system has to relearn and induce emerging or additional heuristics from the newly observed image reference patterns. In addition, for a scheduled radiological examination reading session, the pre-fetching system must have the deductive capability to select and deliver potentially useful patient prior images that are based on the current heuristics of the system. The following describes the essential capabilities of the pre-fetching system together with their respective enabling system components, as depicted in Fig. 3.

Learning Capability: Learning or relearning is activated when the pre-fetching system fails to support an individual radiologist's image needs effectively. The desired learning capability of our pre-fetching system is jointly provided by several system components.

Case Repository: The case repository maintains a collection of historical cases, each of which consists of the radiological examination read and the associated patient prior image reference pattern of the reading radiologist. Using the case repository, the learning system induces individual radiologists' patient prior image reference heuristics. Annotated by its reading radiologist, a historical case contained in the repository is comprised of two parts: decision factors instrumental in explaining and predicting the radiologist's prior image reference pattern, and decision outcomes specifying the actual patient prior images referenced by the radiologist. Collecting historical cases is accomplished by collaboration between the inference engine and access monitor agent. When performing knowledge inferring for a newly scheduled examination, the inference engine processes the scheduled examination as a case subsequently deposited in the case repository. As the reading of the scheduled examination proceeds, the access monitor agent monitors the reading radiologist's actual patient prior image reference behavior that is documented as the decision outcome for this particular examination reading task (case) to be stored in the case repository.

Learning System: The learning system induces a radiologist's patient prior image reference heuristics from his or her image reference patterns documented in the case repository. That is, the learning system uses individual historical cases to perform desirable generalization of the respective documented decision factors and outcomes. Three different approaches from the inductive learning paradigm have been adopted: backpropagation neural network [10], [17], ID3-based multidecision-tree induction (MDTI) [9], [15], [16], and CN2-based multidecision-outcome induction (CN2-MD) [1], [2], [19]. These learning techniques were adopted primarily because of their capabilities in managing the challenging characteristics of radiologists' image references, including incomplete decision factors, as well as multiple and potentially inconsistent decision outcomes. Results from our previously conducted experimental evaluations suggested that the MDTI and CN2-MD techniques may outperform the backpropagation neural network in accuracy (as measured by precision and/or recall rates), learning, and execution efficiency, and, most importantly, explanatory utility [9], [19]. Due to their superiority in these essential learning evaluation dimensions, both MDTI and CN2-MD techniques appear to be appropriate learning algorithms for the learning system. Learning evaluation criteria and comparative results of these two candidate techniques will be discussed in the following section. Interested readers are referred to [9] and [19] for detailed discussion of their theoretical foundations, algorithms, and comparative evaluation results.

A learning system learning session is triggered by the performance monitor agent when it detects unsatisfactory image selection and pre-fetching by the pre-fetching system that fails to support individual radiologists' examination reading tasks. In a learning session, the cases read by an individual radiologist are retrieved from the case repository and randomly assigned for training or testing purposes. Using the documented decision factors and corresponding decision outcomes of the training cases, the learning system constructs multiple decision trees or multiple sets of decision rules, depending on the spe-

cific learning technique applied (i.e., MDTI or CN2-MD). The resulting learning outcomes then are validated using the testing cases. However, the KB pertinent to a specific radiologist is updated only when the validated learning results are significantly superior to those achieved by the existing heuristics.

Knowledge Transformer: The knowledge induced by the learning system is represented using either multiple decision trees or multiple sets of decision rules, neither of which is compatible with the SOOER-based representation of radiologists' heuristic knowledge. Thus, knowledge representation transformation becomes necessary and is performed by the knowledge transformer, which transforms the knowledge from the representation scheme used by the learning system to the appropriate SOOER model representation. Separation of knowledge transformation from learning provides desirable flexibility in the development of the learning system, which, as a result, is not confined by the representation scheme adopted by the underlying KBs. In addition, this separation also allows the learning system not to be concerned with data operation details when representing learning results.

The complexity of a knowledge transformation algorithm greatly depends on the similarity between the source and target representation. Knowledge constructed by MDTI or CN2-MD can easily be converted to predicate-based production rules, similarly to the heuristic rule representation used in a SOOER model. Thus, the knowledge transformation is straightforward. Based on the semantic knowledge maintained in the mediator's catalogue, knowledge transformation requires modifying each predicate in the IF clause to a path expression, changing the THEN clause to corresponding SELECT-FROM-WHERE-ORDER BY clauses with appropriate path expressions in each sub-clause, and determining the entity class in which each transformed rule should be encapsulated.

Knowledge Repository: The knowledge repository maintains a set of KBs, each of which is tailored to a particular individual radiologist and represents his or her patient prior image reference heuristics. Upon completing transformation of the learning results constructed by the learning system into the heuristic rule representation used in a SOOER model, the resulting heuristics are deposited into the KB pertaining to the target radiologist.

Monitoring Capability: The pre-fetching system monitors radiologists' image reference behaviors and its pre-fetching accuracy. The monitoring capabilities are provided jointly by the access monitor agent and performance monitor agent in the following manner.

Access Monitor Agent: The access monitor agent monitors a radiologist's patient prior image reference behavior within an examination reading task. The actual image reference pattern is stored in the case repository to support future learning. The specific patient prior images referenced by the reading radiologist can reside in local or remote image archives (i.e., PACS). Local retrievals refer to those images needed by the reading radiologist that have been correctly identified and selected by the pre-fetching system and transmitted to local archives before the examination reading takes place. On the other hand, remote retrievals are patient images retrieved by the radiologist from remote archives as the examination reading proceeds. Thus, re-

remote retrievals are necessitated by failures of the pre-fetching system to support radiologists' patient prior image reference needs effectively. Both local and remote retrievals are necessary for the performance monitor agent's decision making when initiating a relearning session to update image reference heuristics.

Performance Monitor Agent: The performance monitor agent keeps track of the accuracy of the pre-fetching system with respect to the actual image reference behavior of an individual radiologist. Pre-fetching accuracy is measured by both recall and precision rates. The recall rate measures the effectiveness of the induced image reference heuristics and is defined as "the percentage of the patient prior examinations referenced by the reading radiologist that have been correctly suggested and pre-fetched by the pre-fetching system." On the other hand, the precision rate is concerned with the efficiency of the induced image reference heuristics and is defined as "the percentage of patient prior examinations suggested by the pre-fetching system that are actually referenced by the reading radiologist during the examination reading."

Relearning becomes necessary when the average accuracy of the pre-fetching system in supporting an individual radiologist's examination reading is lower than a pre-specified threshold, as measured by the recall rate, precision rate, or both. When this happens, the performance monitor agent activates the learning system, which, in turn, performs relearning of prior image reference heuristics for the particular radiologist.

Deduction Capability: The knowledge-based patient image pre-fetching system has a built-in deduction capability to determine and pre-fetch the patient prior images presumably needed by a radiologist during reading of a newly scheduled radiological examination. The following discusses the important system components that jointly provide the pre-fetching system's deduction capability.

Inference Engine: The inference engine performs forward-chaining knowledge inferencing to identify the particular patient prior images to be pre-fetched for a newly scheduled examination. Depending on the reading radiologist, a specific KB from the knowledge repository will be applied for image pre-fetching. During a knowledge inferencing session, the IF clauses of the heuristic rules contained in the chosen KB need to be instantiated from local databases or the inference engine's working memory to decide whether or not a rule has been satisfied. When involved in rule instantiation, a local database receives a query (queries) from the inference engine through the mediator. For example, to determine whether or not the following rule can be applied on the newly scheduled examination with Exam-ID being "1123,"

```
IF this.Patient.Age() < 7
THEN SELECT e.Exam-ID
      FROM Examination e
      WHERE e.Patient.Patient-ID = this.Patient.Patient-ID
```

The inference engine formulates a query based on the predicate of the IF clause to obtain the patient's age and submits it to the appropriate local database system(s) through the mediator:

```
SELECT m.Patient.Age()
FROM Examination m
WHERE m.Exam-ID = "1123"
```

When the IF clause of a rule is satisfied by the query result(s), the inference engine subsequently executes the corresponding THEN clause of the rule. Depending on the specific THEN clause, the inference engine may simply update its working memory with the query result(s) or submit to the mediator data retrieval request(s). In the latter case, the data retrieval request(s) is the same as the THEN clause, except that all of the instance variables "this" need to be replaced by a specific data object in data sources, according to the semantic knowledge maintained in the catalogue. To continue with our example, when the above rule is satisfied, the corresponding data retrieval submitted by the inference engine is as follows:

```
SELECT e.Exam-ID
FROM Examination e, Examination m
WHERE m.Exam-ID = "1123" AND
      e.Patient.Patient-ID = m.Patient.Patient-ID
```

Once the inference engine completes the knowledge inferencing session for a newly scheduled examination and the mediator returns the query results in response to the fired rules, the inference engine identifies the specific patient prior examinations presumably relevant to the reading radiologist's reading of the examination under consideration. Information about the identified patient prior examinations is then passed to pre-fetcher, which literally executes the actual image pre-fetching.

Pre-Fetcher: The pre-fetcher receives from the inference engine the information on the relevant patient prior examinations, retrieves the corresponding images from the archives (i.e., PACS), and delivers these images to the local archive of the designated reading site. The pre-fetcher submits queries to the mediator, requesting the images from some identified patient prior examinations. Furthermore, when the local archive has insufficient space to accommodate all pre-fetched images, the pre-fetcher needs to make image replacement decisions to minimize "image faults" that would unnecessarily prolong an examination reading.

Prototype mediator and knowledge-based patient image pre-fetching systems have been implemented. Sybase, a relational database management system (DBMS), is used to develop the catalogue, knowledge repository, case repository, and local archive. The software modules of both the mediator and knowledge-based patient image pre-fetching system were developed in C/C++, except for the MDTI technique, which was implemented in Pascal. On the local data source side, experiments with the HIS, RIS, and PACS have been undertaken using Sybase and Versant, an object-oriented DBMS. The implementation will continue to include other DBMS's employed by different local data sources. Accordingly, the query translator and global transaction manager will need to be extended to manage the additional heterogeneity.

V. SYSTEM EVALUATION

The fundamental objective of the developed patient image pre-fetching system is to effectively adapt to and accurately support radiologists' image reference needs. In this connection, predictive utility is of critical importance in evaluating the system's effectiveness. Image pre-fetching accuracy, in turn, may affect a reading radiologist's decision making in his or her examination reading. Thus, impacts of the pre-fetching system on radiologists' examination reading are also an essential evaluation dimension. Pragmatically, the system has been developed to support radiologists' examination reading thus as to increase efficiency, enhance service quality, and improve work satisfaction. The following discusses some encouraging results from our preliminary evaluative studies conducted at the University Medical Center (UMC), University of Arizona, Tucson, a moderate-sized teaching/tertiary hospital that serves the population of southern Arizona. Choice of the evaluation site was made primarily based on accessibility and considerable needs for effective patient image pre-fetching support. Accordingly, we included the current patient image pre-fetching practice at the UMC to provide a desired benchmark again in which the performance achieved by the respective learning techniques can be evaluated.

A. Evaluation of Learning Accuracy

A study was conducted to evaluate the learning accuracy of the knowledge-based patient image pre-fetching system's learning system implemented using either the MDTI or the CN2-MD technique. Two hundred historical cases, radiological examinations for which clinical reading had been completed by radiologists at the UMC, were included in this study. These cases covered several radiology subspecialty areas, including chest X-ray, computed tomography/magnetic resonance imaging (CT/MRI) body, CT/MRI neural, and musculoskeletal X-ray.

Each case included was stored in the case repository and was jointly described by 14 decision factors identified by previous studies [9], [10], [19]. These decision factors can be classified into three broad categories: current-examination related (e.g., examined anatomical portion, modality used and reason for examination), patient related (e.g., patient's gender, clinical status, and source), and disease/abnormality related (e.g., disease/abnormality type, nature, and phase). The decision outcomes for each case were the reading radiologist's actual image reference behavior monitored by the access monitor agent.

Table I summarizes the aggregate accuracy of several proposed techniques for the learning system (more detailed empirical evaluations can be found in [9] and [19]). The accuracy of the UMC current pre-fetching practice and that of the patient prior image reference heuristics acquired through an intensive knowledge acquisition process (knowledge-engineer-driven approach) were used to provide desirable benchmarks. As shown, the recall rates of both MDTI and CN2-MD techniques were significantly higher than that of the UMC current pre-fetching practice, which, nevertheless, was superior in precision rate, showing a 24.42% or 5.37% advance. The observed low recall and high precision rate combination may have resulted par-

TABLE I
COMPARATIVE AGGREGATE ACCURACY OF PATIENT PRIOR IMAGE REFERENCE HEURISTICS

Source of Knowledge	Recall Rate	Precision Rate	# of Prior Examinations Suggested*
UMC Practice	34.39%	91.06%	0.87
Knowledge-Engineer-Driven	83.69%	73.05%	2.72
MDTI ^a	79.65%	66.64%	2.83
CN2-MD	66.19%	85.33%	1.84

*: Average number of prior examinations referenced by radiologists is 2.33.

^aA pre-pruning approach was taken with the MDTI technique to manage the noise in both decision factors and outcomes. The specific method adopted was critical value pruning (CVP) with which the expansion of a decision tree halts at a node where the ratio of the dominant class is greater than a prespecified threshold value (i.e., the critical value) [9]. In this experiment, the critical value was set to 0.55.

tially from an under-provision tendency of the UMC practice. On average, a radiologist needed to reference 2.33 prior examinations in each examination reading, but the current UMC practice pre-fetched only a 0.87 examination. On the other hand, MDTI and CN2-MD, on the average, delivered 2.83 and 1.84 prior examinations, respectively, close to the average number of prior examinations referenced by radiologists.

The recall and precision rates achieved by the knowledge-engineer-driven approaches were higher than that accomplished by the MDTI technique. On the other hand, the CN2-MD technique appeared to outperform the knowledge-engineer-driven approach in the precision rate, but had a lower recall rate. The better performance yielded by the knowledge-engineer-driven approach can, in part, be explained by its inherent white-box nature. Through interactions with radiologists, it can capture rich relationships between decision factors and outcomes. In contrast, as inductive learning techniques, MDTI, or CN2-MD may suffer from their data-driven nature. That is, a small data set, together with its inherent inductive bias, may put either technique at a distinct disadvantage. However, with a relatively small data set (i.e., 200 cases), the respective accuracy of the MDTI and CN2-MD techniques were largely reasonable, as manifested by the respective 79.65% and 66.19% recall rates and 66.64% and 85.33% precision rates. The resulting learning accuracy of both techniques is expected to improve as the size of data for learning grows. Thus, both techniques clearly demonstrated their applicability and desirability for the learning of radiologists' patient prior image reference heuristics.

B. Evaluation of Effects on Decision Making

If patient prior images are to support radiologists' examination reading, the pre-fetching system's impacts on radiologists' decision making need to be evaluated. Toward this end, we performed a field experiment investigating such impacts, including those on decision time requirements (i.e., examination reading time), decision quality (i.e., accuracy in diagnosis), and satisfaction with the decision making process (i.e., examination reading process) [5].

A nested-within experimental design was used in this study using the UMC's current pre-fetching practice as a benchmark. The KBs acquired by the knowledge-engineer-driven approach were adopted, primarily because of their superiority in pre-fetching accuracy over those currently achieved by KB's induced by either MDTI or CN2-MD, as measured

TABLE II
EXPERIMENTAL RESULTS—MUSCULOSKELETAL EXAMINATIONS

	Residents			Junior Radiologists			Senior Radiologists		
	KB	UMC	Δ^b	KB	UMC	Δ	KB	UMC	Δ
Examination Reading Time (in second)									
Mean	452	570	21%	261	305	14%	92	137	33%
Standard Deviation	147	170		111	73		46	91	
Objective Report Quality Score									
Mean	68	39	74%	81	64	26%	83	73	12%
Standard Deviation	22	49		22	37		20	33	
Satisfaction with Examination Reading^c									
Mean	6.9	3.7	86%	9.0	7.0	29%	9.2	5.5	67%
Standard Deviation	3.2	2.3		1.2	3.1		0.8	3.9	

^b Δ (Rate of improvement)

= (UMC - KB)/UMC for examination reading time.

= (KB - UMC)/UMC for objective report quality score and satisfaction with examination reading.

^cBased on a ten-point Likert scale with "1" being unsatisfied and "10" being very satisfied.

TABLE III
EXPERIMENTAL RESULTS—CHEST EXAMINATIONS

	Residents			Junior Radiologists			Senior Radiologists		
	KB	UMC	Δ	KB	UMC	Δ	KB	UMC	Δ
Examination Reading Time (in second)									
Mean	284	371	23%	278	132	(111%)	212	196	(8%)
Standard Deviation	172	405		120	34		116	115	
Objective Report Quality Score									
Mean	93.2	70.0	33%	87.5	73.0	20%	65.2	69.1	(6%)
Standard Deviation	13.0	36.6		25.0	18.2		35.4	43.5	
Satisfaction with Examination Reading									
Mean	8.4	6.5	40%	8.5	10.0	(15%)	9.1	5.1	78%
Standard Deviation	1.9	3.7		1.0	0.0		1.0	3.3	

by a combination of the recall and precision rates achieved. Two radiology sub-specialties were investigated: chest and musculoskeletal diagnostic radiology. These sub-specialties were chosen because they were the most common examination types at the UMC. Jointly, they accounted for approximately 70% of the radiological examinations performed there at the time of the study. In both sub-specialties, radiologist subjects were classified into three experience levels: senior, junior, and resident. The decision to control for clinical experience levels was made based on results from some previously conducted studies, suggesting that clinical experience differences may contribute to variations in individual radiologists' patient prior image reference needs [5]. Specifically, senior radiologists were those who had more than ten years of post-residency practice in a specified (or certified) radiology subspecialty area, whereas junior radiologists had 3–7 years of post-residency practice.

Two experimental sessions were conducted, one for each investigated subspecialty. Randomly drawn from a previously established historical case pool consisting of examinations collected from clinical settings, eight cases were included in each experimental session. Each radiologist subject performed read-

ings on all of the included cases in the same way they would have done for clinical purposes. At the end of each examination reading, each subject was asked to produce a verbal report for a subsequent report quality analysis. In addition, the total amount of time a subject spent reading an examination was also documented. Also, when completing an examination, each subject was asked to assess his or her satisfaction with the examination reading process. To analyze the diagnosis accuracy of individual radiologist subjects, a gold-standard report was produced for each included examination, using a majority-based method for consolidating individual examination reports rendered by a review panel consisting of three radiologists who did not participate in the experiments as subjects. These reports were then used to evaluate the quality of diagnosis reports generated by subjects.

Tables II and III summarize experimental results obtained from the musculoskeletal and chest examinations, respectively. Overall, it appeared that use of the pre-fetching system to support radiologists' examination reading may have resulted in increased reading efficiency, enhanced service quality, and improved satisfaction when compared with the UMC's current practice. However, the effects seemed stronger with the mus-

culo-skeletal than with the chest examinations. The differential effects may, in part, have resulted from the relative complexity of interpreting chest versus musculo-skeletal images. As several radiologists commented, the radiographic signs on a musculo-skeletal image may be more subtle and complex than those commonly detected on a chest image. Consequently, increased complexity or difficulty may have made radiologists increasingly dependent on prior image pre-fetching support in their examination reading. In this connection, effects of the knowledge-based patient image pre-fetching system may appear to have been stronger with musculo-skeletal than with chest examinations.

Judging from the experimental results, use of the knowledge-based patient image pre-fetching system would result in desired improvements in radiologists' efficiency, service quality, and satisfaction. On average, the knowledge-based patient image pre-fetching system might be able to reduce radiologists' examination reading time for musculo-skeletal examinations by approximately 20%, while its effect on chest examinations might be experience dependent. Use of the knowledge-based patient image pre-fetching system was also shown to be likely to result in better examination report quality, leading to 35% and 15% improvement for the musculo-skeletal and chest examinations, respectively. Furthermore, radiologists may become better satisfied with their examination readings supported by the pre-fetching system than by the UMC's pre-fetching practice. A 60% improvement for musculo-skeletal examinations and a 30% improvement for chest examinations were observed.

VI. CONCLUSIONS AND FUTURE RESEARCH

When performing primary reading on a newly taken radiological examination, a radiologist often needs to reference relevant prior images of the same patient for confirmation or comparison purposes. Support of such image references is of clinical importance and may have significant effects on radiologists' examination reading efficiency, service quality, and work satisfaction. To support such image reference needs effectively, we proposed and developed a knowledge-based patient image pre-fetching system. A conceptual model, i.e., the SOOER model, was used to represent radiologists' highly heuristics-oriented and data-intensive prior image reference heuristics. We detailed the system architecture and design for a knowledge-based patient image pre-fetching system. This architecture, based on a three-tier client-mediator-server design, is capable of coping with a dynamic environment characterized by distributed, heterogeneous, and highly autonomous database systems. To adapt to changes in radiologists' patient prior image reference heuristics, MDTI and CN2-MD learning techniques were developed and empirically evaluated. Effects of the pre-fetching system were also examined. Preliminary results showed both that the knowledge-based patient image pre-fetching system more accurately supported radiologists' prior image needs than the practice currently adopted by the study site and that radiologists may become more efficient, consultatively effective, and better satisfied professionals when supported by the knowledge-based patient image pre-fetching system rather than by the study site's current practice.

Continued theoretical and applied investigations along this line of research are necessary. Several specific topics or issues demand our immediate research attention. For instance, the identification of related or similar applications to expand our research results is desirable. Case selection for clinical education and training appears to be a promising potential application. Intelligent Intranet-based applications that distribute information to users within an organization offer another interesting future research direction. Additional evaluations of MDTI and CN2-MD learning techniques are also important. Large-scale case collection and evaluation studies in different radiology environments are appealing. Performing a longitudinal empirical evaluation of the effects of the knowledge-based patient image pre-fetching system on radiologists' examination reading efficiency, diagnosis quality, and work satisfaction is essential as well. In addition, extending the proposed architecture to incorporate advanced functionality, including image prioritization, is also interesting and would have considerable theoretical and pragmatic impacts.

ACKNOWLEDGMENT

The authors wish to thank radiologists at the UMC, University of Arizona, Tucson, for their cooperation during our requirement analysis and case collection.

REFERENCES

- [1] P. Clark and R. Boswell, "Rule induction with CN2: Some recent improvements," in *Proc. 5th European Conf.*, 1991, pp. 151-163.
- [2] P. Clark and T. Niblett, "The CN2 induction algorithm," *Mach. Learn.*, vol. 3, no. 4, pp. 261-283, 1989.
- [3] W. Du, R. Krishnamurthy, and M. C. Shan, "Query optimization in heterogeneous DBMS," in *Proc. 18th VLDB Conf.*, 1992, pp. 277-291.
- [4] D. Georgakopoulos, M. Rusiniewicz, and A. P. Sheth, "Using tickets to enforce the serializability of multidatabase transactions," *IEEE Trans. Knowledge Data Eng.*, vol. 6, pp. 166-180, Feb. 1994.
- [5] P. J. Hu, C. P. Wei, and O. R. Liu Sheng, "Evaluating a knowledge-based patient image pre-fetching system: A field experiment," Dept. Inform. Syst. Decision Sci., Univ. South Florida, Tampa, FL, working paper, 1999.
- [6] Y. E. Ioannidis, "Query optimization," *ACM Comput. Surv.*, vol. 28, no. 1, pp. 121-123, Mar. 1996.
- [7] K. Levin and R. Fielding, "Method to pre-fetch comparison images in image management and communication systems (IMCS)," in *Proc. SPIE Med. Imag. IV*, Newport Beach, CA, Feb. 1990, pp. 270-274.
- [8] O. R. Liu Sheng, P. J. Hu, C. P. Wei, T. Ovitt, and S. Akata, "Acquisition and evaluation of knowledge-based image retrieval system," in *Proc. SPIE Med. Imag. VIII*, Newport Beach, CA, Feb. 1994, pp. 714-725.
- [9] O. R. Liu Sheng, C. P. Wei, P. J. Hu, and N. Chang, "Automated learning of patient image retrieval knowledge: Neural networks versus inductive decision trees," *Decision Support Syst.*, to be published.
- [10] O. R. Liu Sheng, C. P. Wei, and P. J. Hu, "Neural network learning for intelligent patient-image retrieval," *IEEE Intell. Transport Syst.*, vol. 13, pp. 49-57, Jan./Feb. 1998.
- [11] H. Lodder, B. M. Poppel, and A. R. Bakker, "Pre-fetching: PACS image management optimization using HIS/RIS information," in *Proc. SPIE Med. Imag.*, 1991, pp. 227-233.
- [12] S. Mehrotra, R. Rastogi, Y. Breitbart, H. F. Korth, and A. Silberschatz, "The concurrency control problem in multidatabases: Characteristics and solutions," in *ACM SIGMOD Rec.*, June 1992, pp. 288-297.
- [13] S. Mehrotra, R. Rastogi, H. F. Korth, and A. Silberschatz, "Ensuring consistency in multidatabases by preserving two-level serializability," *ACM Trans. Database Syst.*, vol. 23, no. 2, pp. 199-230, June 1998.
- [14] W. Meng and C. Yu, "Query processing in multidatabase systems," in *Modern Database Systems: The Object Model, Interoperability, and Beyond*, W. Kim, Ed. Reading, MA: Addison-Wesley, 1995, ch. 27, pp. 551-572.

- [15] J. R. Quinlan, "Learning efficient classification procedures and their application to chess and games," in *Machine Learning: An Artificial Intelligence Approach*, R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, Eds. Palo Alto, CA: Tioga, 1983, vol. 1.
- [16] —, "Induction of decision trees," in *Mach. Learn.*, 1986, vol. 1, pp. 81–106.
- [17] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, D. E. Rumelhart and J. L. McClelland, Eds. Cambridge, MA: MIT Press, 1986, vol. 1, pp. 318–362.
- [18] C. P. Wei, "Schema management for large-scale multidatabase systems," Ph.D. dissertation, Dept. Manag. Inform. Syst., Univ. Arizona, Tucson, AZ, 1996.
- [19] C. P. Wei, P. J. Hu, O. R. Liu Sheng, and Y. H. Lee, "Inductive learning approach to intelligent patient image pre-fetching: Extension and evaluation of CN2 algorithm," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, Maui, HI, Jan. 2000.
- [20] D. L. Wilson, D. Smith, and B. Rice, "Intelligent pre-fetch strategies for historical images in a large PACS," in *Proc. SPIE Med. Imag.*, Newport Beach, CA, Feb. 1994, pp. 112–123.

Chih-Ping Wei (M'98) received the B.S. degree from the National Chiao-Tung University, Taiwan, R.O.C., in 1987, and the M.S. and Ph.D. degrees in management information systems from the University of Arizona, Tucson, in 1991 and 1996, respectively.

He is currently an Associate Professor in the Department of Information Management, National Sun Yat-Sen University, Taiwan, R.O.C. He has authored and co-authored papers published in the *IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT*, *IEEE Intelligent Systems*, *IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE*, *Decision Support Systems*, *Telecommunication Systems*, *Journal of Organizational Computing and Electronic Commerce*, and *Topics in Health Information Management*. His current research interests include multidatabase management and integration, knowledge discovery and data mining, information retrieval and text mining, and knowledge management.

Dr. Wei is a member of the Association for Computing Machinery (ACM).

Paul Jen-Hwa Hu received the Ph.D. degree in management information systems from the University of Arizona, Tucson, in 1998.

He is currently an Assistant Professor of Information Systems at the David Eccles School of Business, University of Utah, Salt Lake City. He has authored and co-authored papers published in *Decision Support Systems*, *Journal of Management Information Systems*, *IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT*, *IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE*, *IEEE Intelligent Systems*, *Telecommunication Systems*, *Journal of Organizational Computing and Electronic Commerce*, *Topics in Health Information Management*, and *Journal of Telemedicine and Telecare*. His current research interests include IT management and applications in health care, electronic commerce, human-computer interaction, and software project management.

Olivia R. Liu Sheng received the B.S. degree from the National Chiao-Tung University, Taiwan, R.O.C., in 1981, and the Master's and Ph.D. degrees in computers and information systems from the University of Rochester, Rochester, NY, in 1983 and 1986, respectively.

In 1985, she joined the Management Information Systems faculty, University of Arizona, Tucson, where she has been the Department Head since 1997 and a McCoy-Rogers Professor. From 1995 to 1997, while on a leave of absence from the University, she was an Associate Professor in the Department of Information and Systems, Hong Kong University of Science and Technology. Her research focuses on global IT and management, including distributed enterprise data and knowledge management, electronic commerce, telemedicine, telework, and distributed learning. The research issues she has examined include distributed data allocation, data warehouse and mining, object-oriented models and methods for data and KB and application design, performance and architectural issues in data warehouse and enterprise resource planning (ERP) applications, and IT management. She currently also directs the Enterprise and Business Intelligence System (e.bis) Laboratory where such research is housed. She has authored or co-authored in over 30 journal papers and has received funding from the U.S. Army, National Science Foundation, IBM, Toshiba Corporation, Sun Microsystems, Hong Kong Research Grants Council, Asia Productivity Organization, and SAP University Alliance. She is an Associate Editor for the *Telecommunication Systems and Journal of Information Technology and Management*. She has also guest or co-guest edited special issues for *Decision Support Systems*, *Journal of Parallel and Distributed Computing and Telecommunication Systems* on multimedia information systems and healthcare decision support.

Dr. Sheng has been an guest or co-guest editor for the *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*.