Semantic Case Based Reasoning on Medical Knowledge

Dr. Passent ElKafrawy[1]
Associate Professor,
Faculty of Science, Menofia University
Shebin Elkom Menofia, Egypt

Rania. A. Mohamed[2]
Assistant Lecture,
Modern University for Technology & Information
Cairo, Egypt

Abstract: This paper presents a new approach to Case-Based Reasoning (CBR) using Semantic knowledge (SCBR) for the representation of cases, case structure, and case based ontologies in biology and medicine. The approach could be extended to other application domains of CBR. The major advantage of such approach is that Semantic data systems are designed to understand the content of the real world as accurately as possible within the data set. This paper also makes a comparison between traditional CBR and SCBR where there are some problems in traditional CBR such as adaptation may be difficult; cases may need to be created by hand; large processing time to find similar cases; and CBR systems generally give good or reasonable solutions this is because the retrieved case often requires adaptation. SCBR framework can handle these problems.

Keywords: Case-Based Reasoning, Ontology, Semantic Knowledge.

I. INTRODUCTION

The importance of the medical field in today's life cannot be underestimated simply because there seems to be a continuing advancement in the complexity and severity of many diagnosed medical maladies. The medical field is the scientific discipline that deals with finding cure for every conceivable type of illness and disease so this paper use case based reasoning in medical field to help doctors diagnose diseases, to find the appropriate treatment for the patient and to analyze causes and/or treatments.

Case-based Reasoning is an emerging field in Artificial intelligence. The common application areas of CBR includes help-desk and customer service, recommender systems in electronic commerce, knowledge and experience management, medical applications and applications in image processing, applications in law, technical diagnosis, design, planning and applications in the computer games and music domain. Case-based reasoning is an approach, which utilizes the experience gained from past solved problems [1]. This approach maintains all information of past problems solved (i.e. experience) that is called the case. The collection of all these past experiences is stored in a form of case based. There are various factors which define the efficiency of this approach [2]. The major factor is that a solution to a new problem is projected from the number of past experiences stored in the case based. A new problem should be matched to the closest problem of past experiences faced. The new upcoming problem is considered as a new case. The strategies of finding a similar case for the new case regarding past cases stored in the case based is another major factor of defining the efficiency of the case-based reasoning approach.

Case-based reasoning systems have some drawbacks such as: occupies a large storage space for all the cases, take large processing time to find similar cases in case-based, cases may need to be created by hand, adaptation may be difficult, requires a case-based, case selection algorithm, and possibly case-adaptation algorithm. When required best solution or optimum solution, then CBR may not be able to handle such solutions. Hence, we propose a case based reasoning mechanism with semantic knowledge to handle these problems. Semantic data is the information that allows machines to understand the meaning of information. It describes the technologies and methods that convey the meaning of information.

Using semantics, data can be accessed more intelligently as it contains automated agents [3] to understand information. Basically, it breaks down the information into its simplest form so that it is quickly and deeply understood by the machine.

Semantic data is not formally defined and it incorporates the following: Resource Description Framework, data interchange formats and notations, and the Web Ontology Language which all give a defined answer for concepts, terms, and relationships in a specific domain. The concept of semantic data as a whole has remained unclear and is generally speculated upon as not being a workable service.

The purpose of semantic data would be to allow computers to understand and figure out information without the help of a human user. In order to handle knowledge each piece of information must be programmed in details, however, using semantic technology, knowledge is self-defined and easily handled. The computer would be able to find information on its own, combine it with other information as needed, and act upon the information it received in an appropriate way. Semantic data plays an important part in this because the way that the data is lined up allows for the rest of the data in a sequence to be automatically figured out. The data is interpreted according to its relationship to each other and the result knows what the next data would be in sequence based on these relationships. All of the data and parts of a sequence are in an ordered hierarchy so that there is only one choice for the next part in a sequence. It's sort of like solving a math equation, there is only one correct answer almost all of the time. This relationship of data is what allows machines to work alone without human intervention.

In many domains Case-based Reasoning (CBR) has become a successful technique for knowledge-based systems and especially the medical domain. In medical domains, attempts to apply the complete CBR cycle are rather exceptional. Some systems have recently been developed [4], which on the one hand use only parts of the CBR method, mainly the retrieval, and on the other hand enrich the method by a generalization step to fill the knowledge gap between the specificity of single cases and general rules [5]. So we discuss the appropriateness of CBR for medical knowledge-based systems, point out problems, limitations and possibilities how they can partly be overcome.

This paper is organized in the following after this introduction; a theoretical background is illustrated in section 2.Section 3 gives an outline on related work. Section 4 introduces a brief description about semantic ontology. The description of the proposed architecture is given in section 5. Section 6 presents a comparative study after testing and comparing CBR applications. The conclusion and future work is presented in section 7.

II. THEORETICAL BACKGROUND

A. Case Based Reasoning

Case-Based Reasoning (CBR) is a problem solving paradigm that solves a new problem by remembering a previous similar situation and by reusing information and knowledge of that situation [6]. More specifically, CBR uses a databased of problems to resolve new problems. The databased can be built through the knowledge engineering (KE) process or it can be collected from previous cases.

In a problem-solving system, each case would describe a problem and a solution to that problem. The reasoning engine solves new problems by adapting relevant cases from the library [7]. Moreover, CBR can learn from previous experiences. When a problem is solved, case-based reasoning can add the problem description and the solution to the case library. A new case generally represented as a pair <problem, solution> is immediately available and can be considered as a new piece of knowledge.

According to Doyle et al. [8], Case-Based Reasoning is different from other Artificial Intelligence approaches in the following ways:

- Traditional AI approaches rely on general knowledge of a problem domain and tend to solve problems on a first-principle while CBR systems solve new problems by utilizing specific knowledge of past experiences.

- CBR supports incremental, sustained learning. CBR solves a problem then it will make the problem available for future problems.

The CBR Cycle can be represented by a schematic cycle, as shown in Figure 1. The first phase is the retrieve phases, which identify features via noticing the feature values of a case, initially match a list of possible candidates and select the best match from the cases.

Second phase is the reuse phase where the difference between the new and the old case is determined by copying from the old case and adapting by transforming or reusing the old solution.

The third phase is the revise phase, in this phase, if the solution from the last phase is incorrect, then this solution must be evaluated in a real environment setting and finds the errors/flaws of the solution if the solution was evaluated badly.

Finally, the retain phase which is the fourth phase, in this phase incorporate the lesson learned from the problem-solving experience into the existing knowledge by extracting or indexing. By extracting we mean if the problem was solved using an old case, the system can build a new case or generalize an old case. By indexing we mean via deciding what types of indexes to use for future retrieval and integrate by modifying the indexing of existing cases after the experience.
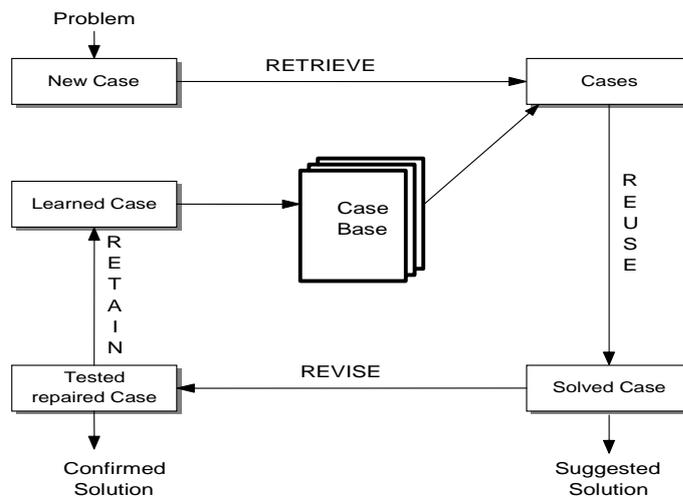

Fig. 1 Case-based reasoning

There are three main types of CBR that differ significantly from one another concerning case representation and reasoning. The first one is Structural [9] in which cases are stored using a common structured vocabulary, i.e. ontology. The second is Textual [10] in such way cases are represented as free text, i.e. strings. The third is the Conversational CBR [11] in which a case is represented through a list of questions that varies from one case to another; knowledge is contained in customer / agent conversations.

During the past twenty years, many CBR applications have been developed, ranging from prototypical applications build in research labs to large-scale fielded applications [12] developed by commercial companies.

There are some disadvantages of CBR such as: can take large storage space for all the cases, can take large processing time to find similar cases in case-based, cases may need to be created by hand, adaptation may be difficult, needs case-based, case selection algorithm, and possibly case-adaptation algorithm. Optimum solution or best solution cannot be achieved using CBR.

CBR systems generally give good or reasonable solutions this is because the retrieved case often requires adaptation.

B. Semantic ontology

Semantic web is actually an extension of the current web in that it represents information more meaningfully for humans and computers alike. It enables the description of contents and services in machine-readable form, and enables annotating, discovering, publishing, advertising and composing services to be automated. It was developed based on Ontology, which is considered as the backbone of the Semantic Web. In other words, the current Web is transformed from being machine-readable to machine-understandable. In fact, Ontology is a key technique with which to annotate semantics and provide a common, comprehensible foundation for resources on the Semantic Web. Moreover, Ontology can provide a common vocabulary, a

grammar for publishing data, and can supply a semantic description of data, which can be used to preserve the Ontologies and keep them ready for inference [13, 14].

Ontologies [15], which are used in order to support interoperability and common understanding between the different parties, are a key component in solving the problem of semantic heterogeneity, thus enabling semantic interoperability between different web applications and services.

Recently, ontologies have become a popular research topic in many communities, including knowledge engineering, electronic commerce, knowledge management and natural language processing. Ontologies provide a common understanding of a domain that can be communicated between people, and of heterogeneous and widely spread application systems. In fact, they have been developed in Artificial Intelligence (AI) research communities to facilitate knowledge sharing and reuse.

The goal of ontology is to achieve a common and shared knowledge that can be transmitted between people and between application systems. Thus, ontologies [16] play an important role in achieving interoperability across organizations and on the Semantic Web [17], because they aim to capture domain knowledge and their role is to create semantics explicitly in a generic way, providing the basis for agreement within a domain. Ontology is used to enable interoperation between Web applications from different areas or from different views on one area. For that reason, it is necessary to establish mappings among concepts of different ontologies to capture the semantic correspondence between them. However, establishing such a correspondence is not an easy task.

The primary use of the word "ontology" is in the discipline of philosophy, where it means "the study or theory of the explanation of being" [18]; it thus defines an entity or being and its relationship with an activity in its environment. In other disciplines, such as software engineering and AI, it is defined as "a formal explicit specification of a shared conceptualization" [18]. The foundations of this definition are:

- All knowledge (e.g. the type of concepts used and the constraints on their use) in ontology must have an explicit specification.

- An ontology is a conceptualization, which means it has a universally comprehensible concept.

III. RELATED WORK

Case based reasoning (CBR) is a known problem solving technique based on reutilizing specific knowledge of previously experienced problems stored as cases. The CBR cycle consists of four major stages: Retrieve, Reuse, Revise and Retain (as shown in figure 1). In the Retrieve stage, the system selects a subset of cases from the case based that are relevant to the current problem. The Reuse stage adapts the solution of the cases selected in the retrieve stage to the current problem. In the Revise stage, the obtained solution is verified (either by testing it in the real world or by examination by an expert), which provides feedback about the correctness of the predicted solution. Finally, in the Retain stage, the system decides whether or not to store the new solved case into the case based.

Fuchs and Mille [19] have proposed a modeling of the CBR at the knowledge level. They have distinguished four knowledge models: the conceptual model of the domain describing the concepts use to describe the domain ontology independently of the reasoning; the case model which separates the case in 'problem, solution', and track of reasoning; the tasks reasoning models which include a model of specification and other one of tasks decomposition and; reasoning supports model.

D'Aquin [20] worked on the integration of the CBR in semantic Web. For that purpose, they have proposed an extension of OWL (Ontology Web Language) allowing representing the adaptation knowledge of the CBR. The expression of domain and cases knowledge in OWL allowed them to add to the CBR system the appropriate reasoning capacities of OWL by exploiting, for example, the subsumption and the instantiation.

Bichindaritz has demonstrated the use of ontologies for facilitating case structuring and acquisition [21]. Diaz-Agudo and Gonzalez Calero [22] proposed an architecture independent from the domain which helps to integrate ontologies in CBR applications. Their approach consists in building integrated systems which combine cases specific knowledge with generic models of the domain knowledge. They presented CBROnto [23], as task / method ontology which supplies the necessary vocabulary to describe implied elements in the CBR processes

Case-based reasoning generally takes large storage space for all the cases, and also take large processing time to find similar cases in case-based. Moreover, cases may need to be created by hand which is another overhead when using case based reasoning and adaptation may be difficult. This paper introduces Case based Reasoning using semantic (SCBR) knowledge where it defines the cases semantically. The cases are semantically defined before being stored in the CBR system. New cases are semantically represented before being matched to the stored experiences so the case won't take large storage space, consequently, it can handle the drawbacks of CBR as shown in the next section.

IV. PROPOSED SCBR SYSTEM

The new system mainly depends on defining the cases semantically. The cases are semantically defined before being stored in the CBR system; hence cases are better understood and thus represented with higher level of understanding. Consequently, new cases are initially also semantically represented before being matched to the stored experiences, for easier matching and verification. This has been achieved by dividing the system into three main layers. Layer 1 is the GUI interface, Layer 2 is responsible for semantic knowledge representation, and finally layer 3 is concerned about the CBR process (semantic case storage and retrieve cases).
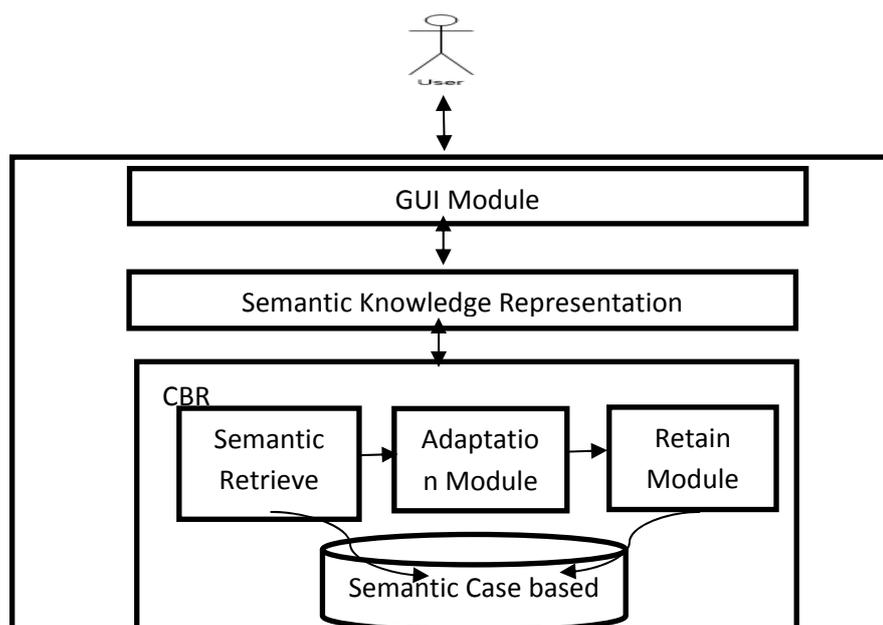

Fig. 2 SCBR System modules

Layer 1: GUI Module

This module is the first layer in the system and is responsible for the graphical user interface, in other words, the interaction between the system and the user of the system. The user (patient) will insert Symptoms of the disease and all his precautions then apply the SCBR system to diagnose the patient's condition and find the right medicine.

Dr. Passent et al.,

International Journal of Advance Research in Computer Science and Management Studies
Volume 3, Issue 11, November 2015 pg. 203-216

Layer 2: Semantic Knowledge Representation

This module is responsible for understanding the input case and provides a semantic definition to the description of the case. We extract Metadata information from the input case to store our semantic understanding of the case.

Metadata is textual data, which contains a description of the concepts of the case. The module extracts a list of feature and their values from the input that will be used as extra attribute values in the retrieval phase (not included as attributes in the case). This extra information provides a deep understanding of the content of the case that helps later in the CBR processes. It guaranties a higher accuracy in matching or searching past experiences.

The paragraph is divided into a set of sentences and each sentence contains a set of tokens as shown in Figure 3.
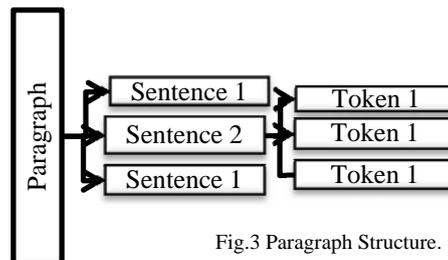

Fig.3 Paragraph Structure.

Each word in the text is represented as a token stored as an object. These objects store information like [24, 25]:

- **Stop word** (word without sense) does not contain important significance to be used in Search Queries. Usually these words are filtered out from search queries because they return vast amount of unnecessary information (i.e. a, about, before, above, after, again, the, that …)

- **Main name inside the sentence**: it is the direct word, which is related to the medical field such as Kennedy disease, pregnant, headache, etc.

- **The stemmed word**: is to reduce the word to its origin. The term doesn't have to reduce the word to its root, as some times it gives a completely different meaning. A stem may consist of just a root. However, it may also be analyzed into a root plus derivational morphemes for example, the words "argue", "argued", "argues", "arguing", and " argus" reduce to the stem "argu" (illustrating the case where the stem is not itself a word or root) but "argument" and "arguments" reduce to the stem "argument".

- **The Part-Of-Speech tag of the token:** also called grammatical tagging or word-category disambiguation, is the process of marking up a word in a text (corpus) as corresponding to a particular part of speech, based on both its definition, as well as its context. It is simple form to the identification of words as nouns, verbs, adjectives, adverbs, etc.

- **A list of relations with other similar tokens**: it means the relation between the tokens (worlds) extracted from the input text.

The organization in paragraphs, sentences and tokens is performed by NLP methods depending on the chosen implementation. The information extracted from the text is stored in the IEtext object. There are several types of information that will be obtained:

- Phrases identified in the text.

- Features: identifier-value pairs extracted from the text.

- Topics: combining phrases and features of a topic that can be associated to a text. A topic is a classification of the text.

Phrases and Features are stored using the objects implemented in the jcolibri.extensions. Textual.IE.representation.info sub-package. That package store three objects that aid in the representation of the extracted information:

- Phrase Info: stores extracted phrases.

- Feature Info: stores extracted features.

- Weighted Relation: represents a weighted relation between two tokens.   These relations are found by the glossary and thesaurus methods.

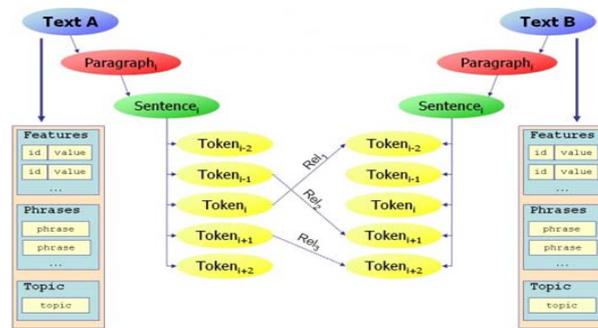Figure 4 illustrates the complete organization:

Fig.4 Global view of the representation of texts for IE.

Finally, the case will be stored in the case based with both the description of the case and the solution of the case as shown in figure [5].

The description part includes stored attributers with their values and an extra metadata that describes the case and will be used to extract an extra attribute that are not stored directly.
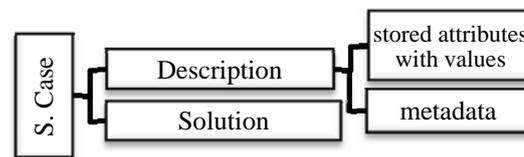

Fig. 5 Semantic Case Structure

Layer 3: Semantic CBR Module

This module is divided into three stages that represent the main CBR processes. These are the retrieval module, adaptation module and the retain module.

A.   Semantic Retrieve Module

Measure the similarity of the cases and retrieve most N similar cases [26].

Computing Similarity

The OpenNLP[X] library is a machine learning based toolkit for the processing of natural language text [27]. It supports the most common NLP tasks, such as tokenization, sentence segmentation, part-of-speech tagging, named entity extraction, chunking, parsing, and reference resolution. These tasks are usually required to build more advanced text processing services. OpenNLP also includes maximum entropy and perceptron based machine learning [28]. The methods of the IE extension extract information from texts and store it into the other attributes of the case. These attributes can be compared using normal similarity functions as shown in Figure 6.
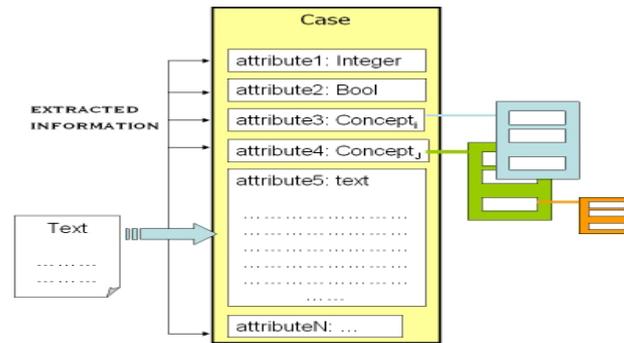
Fig. 6 Common organization of textual cases

The generated entities (words) from OPenNLP[X] will be sent in the background request to the DBpediaSpotlight[X] to annotate text and get the most important features [29].

Due to defining the cases semantically by using ontology, it solves some of the main problems in the traditional case based reasoning. Large processing time to find similar cases is solved in semantic retrieval while providing a target solution in little time. This is because it depends on complex semantic case structures.

B. Adaptation Modules

Adaptation, as one of the most difficult tasks (especially in a complex problem domain) in traditional CBR, relies on both the retrieval of proper cases that need less adaptations and the utilization of appropriate domain knowledge. In this step the Adapting module transforms or reuses the old solution because in many situations the case returned is not the exact solution needed.

Semantic representation of cases reduced the need for adaptation. Traditional CBR requires case adaptation when given solution is not as required or far from the real solution. This is reduced as case understanding is enhanced and deeper understanding of case knowledge is achieved. The new case can be adapted easily through the semantic relations of its knowledge.

C. Retain Module

In the retain step useful new cases are stored in the Semantic case for future reuse. This way the SCBR system has learned a new experience (knowledge based learning). SCBR is intuitive - it's how we work, no knowledge elicitation to create rules or methods this makes development easier and systems learn by acquiring new cases through use.

V. IMPLEMENTATION OF SCBR FRAMEWORK

A. Semantic Ontology

The ontologies are useful for designing SCBR applications because they allow the knowledge engineer to use knowledge already acquired, conceptualized and implemented in a formal language, like DLs based languages, reducing considerably the knowledge acquisition bottleneck. Ontologies used to build models of general domain knowledge. Although in a SCBR system the main source of knowledge is the set of previous experiences, the approach is to CBR is towards integrated applications that combine case specific knowledge with models of general domain knowledge. The more knowledge is embedded into the system, the more effective is expected to be. Semantic CBR processes can take advantage of this domain knowledge and obtain more accurate results.

As an example appeared in jCOLIBRI figure7 shows an example shows how to map a case into ontology.
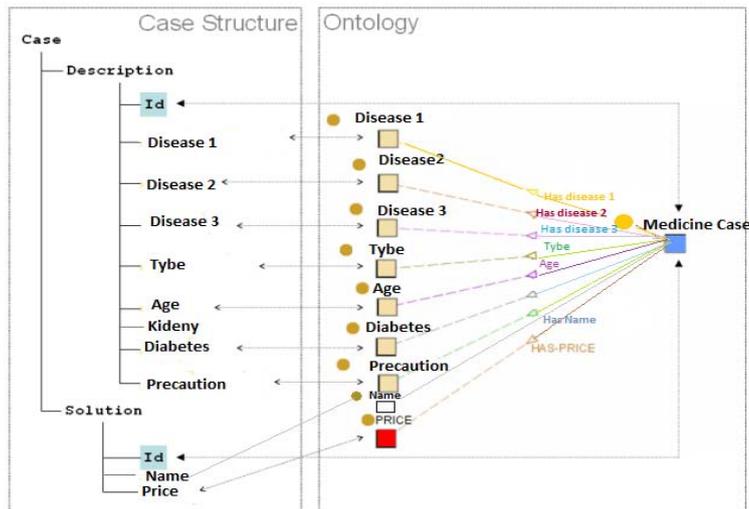
Fig. 7 Case representation ontology

B.   The algorithm to use case based reasoning with semantic knowledge is as follows:

1.   Acquire input text (Symptoms of the disease and precautions)

2.   Use open NLP to extract the precautions from the input text

3.   Use dbpedia-spotlight (semantic representation) to understand and pick the important precaution and symptoms and represent them as a query

4.   Apply the 4 steps of Case based reasoning to the extracted knowledge (as evaluated and represented semantically), cases are stored to determine the best medicine

    i.   **Retrieve:**

- Identify features: noticing the feature values of a case

- Initially match a list of possible candidates by using ontology

- Select the best match from the cases

    ii.   **Reuse:**

Here, we try to find the difference between the new and the old case by the following:

- Copying: the solution is simply copied from the old case.

- Adapting: transforming or reusing the old solution.

    iii.   **Revise:**

If the solution from the last phase is incorrect, then we must:

- Evaluate this solution in a real environment setting.

- Find the errors/flaws of the solution if the solution was evaluated badly.

    iv.   **Retain:**

Incorporate the lesson learned from the problem-solving experience into the existing knowledge by:

- Extracting: the problem was solved using an old case; the system can build a new case or generalize an old case.

The following figures show the result when using Semantic case-Based Reasoning (SCBR) where the user insert the text (This an test of pregnant to and have asthma and stomach gastric and intestinal ulcer and kidney problem be University by sugar child and high pressure data will be adults only Gordon and maybe he is a children heart) then the open NLP extract the precautions from the input text (asthma, stomach, intestinal, peptic ulcer, kidney, blood pressure, heart) then use dbpedia-spotlight to understand and pick the important precaution and symptoms and represent them as a query finally apply four steps of CBR to determine the best three medicine as shown in figure 8.
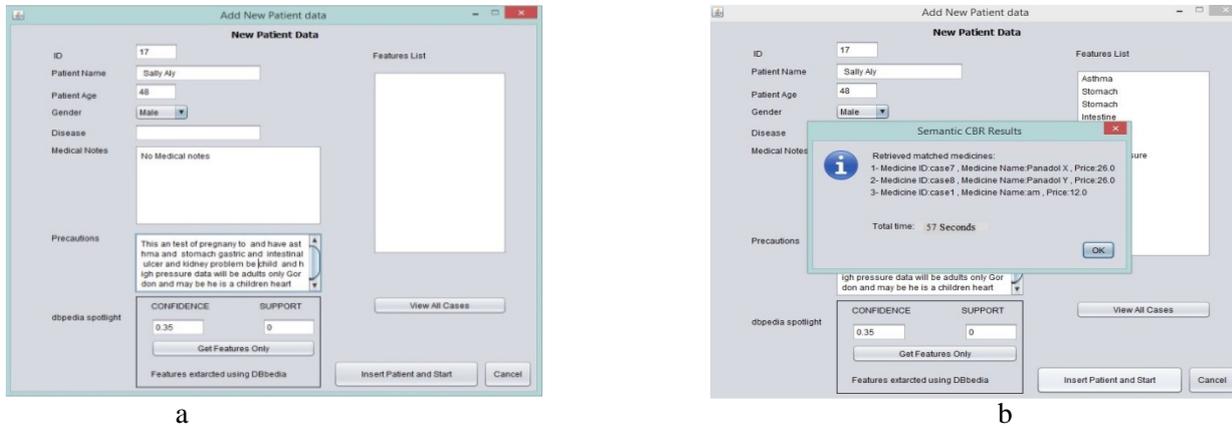

a   b

Fig. 8 a,b Show the result when using SCBR

VI. RESULTS

The case based used for testing the previously mentioned software obtained from the UC Irvine Machine Learning Repository, which contains details for 1000 cases for used patients (Diagnosis) [30].

This paper introduces a comparative study after testing and comparing the CBR applications mentioned previously in table 1 using the same case based

There are a number of major concerns when studying case-based reasoning approach. These major concerns are listed below:

- What is the structure of the cases?

- What is the numbers of correct match cases?

- What are (Recall, F-Measure and Accuracy)?

- What is the time taken to find target solution?

TABLE 1 CBR shell comparison

CBR Shell	Case Structure	Correct Match Cases	Precision	Recall	F-Measure	Accuracy	Total Time
CBR Shell	Textual	70	0.666	0.7	0.682	0.703	3 minutes
Free CBR	Textual	81	0.784	0.81	0.786	0.813	2.3 minutes
jCOLIBRI	Xml /text	93	0.902	0.93	0.961	0.934	1.58 minutes
myCBR	Object	90	0.881	0.90	0.906	0.891	2.1minute
eXiTCBR	Custom CSV	61	0.603	0.61	0.606	0.613	4 minutes
SCBR System	Textual	99	0.933	0.99	0.961	0.994	57 seconds

The results show that the highest accuracy reached and the number of cases retrieved and matched also take the least time of cases retrieved and matched through the SCBR followed by jCOLIBRI, myCBR, FreeCBR,CBR Shell and eXiTCBR respectively.

VII. COMPARATIVE STUDY BETWEEN TRADITIONAL CBR AND SEMANTIC CBR

This section introduces a comparative study after testing and comparing the CBR applications mentioned previously in table 1 using the same case based.

There are a number of major concerns when studying case-based reasoning approach. These major concerns are listed below:

- What are the selection strategies for finding similar cases?

- How is the case being retrieved?

- How is the selected case being revised?

- How is the suggested case being stored in case based?

- How is the suggested case being indexed for faster access?

- How to deal with noisy data or missing values?

According to the previous mentioned points, a comparative study between the traditional CBR software and case based reasoning using semantic (SCBR) mentioned previously in table 1. Next paragraphs describe the effect of each factor to each CBR software respectively.

No interfaces to external systems and DB are available in myCBR. It is valid regarding the interfaces to real-time or diagnostic systems. On Retain phase, myCBR allows saving the Query as a new case, also to use an old case as a basis for new Query. MyCBR is entirely based on GUI, providing a ready-windows templates and forms for defining classes, attributes, SFs, queries to the case-based DB, visualization of found results and more.

myCBR platform can be used for non-complex CBR applications development with partial CBR R4 cycle and with small number of cases in text file. For CBR application development, no time for programming is needed but it is needed only for case configuration. MyCBR is not suitable to be applied with large number of attributes with text solution, especially when they must be visually presented in one window. Table 1 summarizes the comparisons between the selected CBR software.

SCBR system has a very simple and powerful GUI; it represents cases semantically in a very simple way so the cases don't need to be created by hand. There are a number of case retrieval algorithms applicable in case based reasoning. These algorithms are based on the similarity metric that allows resemblance between cases stored in case based. The nearest neighbor retrieval algorithm & induction retrieval algorithms are two chief algorithms used in this process. Nearest-neighbor retrieval is a straightforward approach that computes the similarity between relevant cases found through indexing. According to using ontology in retrieval stage theses algorithm don't take large processing time to find similar cases in case-based and CBR systems generally give good or reasonable solutions and possibly case don't need adaptation algorithm and if it needs SCBR can make maintenance easy and justification through precedent.

TABLE 2 CBR shell comparison.

CBR Shell	Case Structure	Selection strategies	Case retrieval	Case revised	Case stora-ge	Case indexed	Graphical User Interface (GUI)	Dealing with uncertain data
CBR Shell	Text-ual	distance method	Two methods KNN Threshold	Manual	Text	No	Very simple GUI	Can't handle
FreeCBR	Text-ual	weighted Euclid distance	Simple matching	Manual	Text	No	Simple and easy but limited	Can't handle

jCOLIBRI	Xml /text	Similari-tyfunc-tions	method k-NN, Threshold, Ontology, Textual, OpenNLP and GATE Recommender	Automatic	CSV XML	Yes	Simple and power-ful Use wizard to simplify	Handle as null
myCBR	Object	similarity functions	Query model	Manual	CSV XML	No	user can customize the GUI and handle most of things	Handle as _unknown_ or _undefined
eXiTCBR	Custom CSV	Distance method or similarity measure	Simple Querying	Manual	Text	No	Very simple , no options	Can't handle

Table 3 can be summarize the difference between traditional case based reasoning and case based reasoning using semantic knowledge as proven in proposed SCBR system.

Table 3: Diff. between traditional CBR &SCBR

Traditional Case based Reasoning(CBR)	**Semantic Case based reasoning(SCBR)**
• Can take large storage space for all the cases • Can take large processing time to find similar cases in case-based • Cases may need to be created by hand • Adaptation may be difficult • Needs case-based, case selection algorithm, and possibly case-adaptation algorithm • if you require the best solution or the optimum solution CBR may not be for you • CBR systems generally give good or reasonable solutions this is because the retrieved case often requires adaptation	• CBR is intuitive - it's how we work • no knowledge elicitation to create rules or methods this makes development easier • systems learn by acquiring new cases through use • this makes maintenance easy • justification through precedent • Adaptation may be easy • SCBR system give the target solution this is because it depend on semantic • Complex case structures • Knowledge-based learning

VIII. CONCLUSION AND FUTURE WORK

Case-based reasoning systems (CBR) have some drawbacks such as: occupies a large storage space for all the cases, take large processing time to find similar cases in case-based and cases may need to be created by hand. This paper proposed a case based reasoning mechanism with semantic knowledge to handle these problems where the new system mainly depends on defining the cases semantically. New cases are semantically represented before being matched to the stored experiences.

This paper also introduces a comparison among most common used traditional CBR software and the proposed SCBR system. It also mentions the advantages and disadvantages of each software. Moreover, this paper applies the same case based to the six CBR software to compare and evaluate the results using the predetermined factors and calculating Precision, Recall, F-Measure and Accuracy for each one. As a conclusion CBR, Free CBR and eXit CBR are very simple software including simple GUI and only include the selection and retrieval of similar cases using traditional techniques. On the other hand, Proposed SCBR system, myCBR and jCOLIBRI are more complex and can be used for complex CBR.

The proposed system can prove that combining Ontology technology and CBR has a positive impact on search results and the more cases are stored to increase system performs. We recommend applying this approach to cases on Wikipedia in other fields. Also, for future work we will investigate the methodology for building ontology from unstructured data such web pages and documents. Moreover, more investigation can be done for reducing the case storage size and time.

References

1. Aamodt and E. Plaza, "Case-based reasoning: foundational issues, methodological variations, and system approach" AI Communications 7(1), 39–59, 1994.
2. Janet L. Kolodner, "An Introduction to Case-Based Reasoning" Artificial Intelligence Review 6, 3-34, 1992.
3. Abadi, D., A. Marcus, S. Madden, and K. Hollenbach (2009). SW-Store: a vertically partitioned DBMS for Semantic Web data management.
4. Intelligence and Applications (AIA 2009), IASTED, Innsbruck, Austria, Editor(s):V. Devedžic, February, 2009.
5. Bichindaritz I, Marling C. Case-based reasoning in the health sciences: What's next? In Artificial Intelligence in Medicine. 36(2), 2006.
6. Simon C.K. Shiu, "Case-Based Reasoning: Concepts, Features and Soft Computing" Applied Intelligence 21, 2004.
7. Zhi-We Ni, Shan-Lin "Integrated Case-based Reasoning" Proceedings of the Second International Conference on Machine Learning and Cybernetics, XI; 2-5 November 2003.
8. Leake, David, "CBR in Context: The Present and Future (http:/ / www. cs.indiana.edu/~leake/papers/p-96-01_dir.html/paper. html)", In Leake, D., editor, Case-Based Reasoning: Experiences, Lessons, and Future Directions. AAAI Press/MIT Press, 1-30, 1996.
9. R. Bergmann, K.-D. Altho®, S. Breen, M. GÄoker, M. Manago, R. TraphÄoner, and S. Wess. Developing industrial case-based reasoning applications: The INRECA methodology. LNAI 1612. Springer, 2nd Edition, 2003.
10. M. Lenz and K. Ashley, editors. Proceedings of the AAAI98 Workshop on Textual Case-Based Reasoning. AAAI Press, 1998.
11. D. Aha, L. A. Breslow, and H. Munoz-Avila. Conversational case-based reasoning. Applied Intelligence, 2001.
12. I. Watson. Applying case-based reasoning: techniques for enter-prise systems. Morgan Kaufmann Publishers Inc., 1998.
13. Antoniou, Grigoris, & Frank van Harmelen, " A Semantic Web Primer. Cambridge", MA: MIT Press.Bridgman, & Percy, W. (1922). Dimensional Analysis. New Haven, CT: Yale University Press,2004.
14. S. Bechhofer, I. Horrocks, C.A. Goble, and R.Stevens,"OilEd: a Reason-able Ontology Editor for the Semantic Web", In Proceedings of Description Logics, 2001.
15. D. Dou, D. McDermott, and P. Qi, "Ontology Translation on the Semantic Web", Presented at on Data Semantics Journal, 2005.
16. D. Fensel, "Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce", Springer, 2001.
17. D. Tidwell, "Web Services-The Web's Next Revolution", IBM Web Service Tutorial, 29 Nov. 2000, http://www- 106.ibm.com/developerworks/edu/ws-dwwsbasics-i.html.
18. H. Mihoubi, A. Simonet, and M. Simonet, "An Ontology Driven Approach to Ontology Translation", In Proceedings of DEXA, 2000.
19. B. Fuchs and A. Mille, "Une modélisation au niveau connaissance du raisonnement à partir de cas", in L'HARMATTAN, Ed., Ingénierie des connaissances, 2005.
20. M. D'aquin, J. Lieber and A. Napoli, "Artificial Intelligence: Methodology, Systems, and Applications", Lecture Notes in Computer Science, chapter Case-Based Reasoning Within Semantic Web Technologies. Springer Berlin / Heidelberg, vol. 4183, 2006.
21. I. Bichindaritz, Thesis: "Case Based Reasoning Meets the Semantic Web in Biology and Medecine", Advances in Case Based Reasoning, 7th European Conference, LNAI, Spain, 2004.
22. B. D-Agudo and P. G-Calero, "An architecture for knowledge intensive CBR systems", Advances in Case-Based Reasoning". (EWCBR'00). Springer-Verlag, Berlin Heidelberg New York, 2000.
23. B. D-Agudo and P. G-Calero, "CBROnto: a task/method ontology for CBR". In S. Haller and G. Simmons, editors, Procs. of the15th International FLAIRS'02 Conference. AAAI Press, 2002.
24. M. Brown, C. Förtsch, and D. Wissmann. Feature extraction - the bridge from case-based reasoning to information retrieval. In Proceedings of 6th German Workshop on Case-Based Reasoning, 1998.
25. S. Brüninghaus and K. D. Ashley. The role of information extraction for textual CBR. In Proceedings of the 4th International Conference on Case-Based Reasoning, Springer-Verlag, 2001.
26. Croft et al., Bruce Croft, Donald Metzler, and Trevor Strohman. Search Engines: Information Retrieval in Practice. Addison Wesley, 2009.
27. K. M. Gupta and D. W. Aha. Towards acquiring case indexing taxonomies from text. In Proceedings of the 17th Int. FLAIRS Conference, Miami Beach, FL, 2004.
28. E. Hatcher and O. Gospodnetic. Lucene in Action (In Action series). Manning Publications Co., Greenwich, CT, USA, 2004.
29. Bird, Steven, Ewan Klein, and Edward Loper. Natural Language Processing with Python, Sebastopol, CA: O'Reilly, 2010.
30. http://archive.ics.uci.edu/ml/

AUTHOR(S) PROFILE

Dr. Passent M ElKafrawy, is an Associate Professor since 2013; she got her PhD from the University of Connecticut in United states on 2006 in Computer Science and Engineering. In the field of computational geometry as a branch of Artificial Intelligence. Then she taught in Eastern State University of Connecticut for one year. In 2007 she worked as a Teacher in Faculty of Science, Menoufia University, Mathematics and computer science department since that time till now.

She has over 20 publications and member of ACEE, ECOLE and TIMA research organizations. One of the organizing members of the following conferences: SPIT and ESOLEC. Supervising over 10 research studies between PhD and MSc. Member of the faculty projects for education development as CIQAP, DSAP, and Question Bank.

Rania Ahmed, Assistant Lecture in Modern University for Technology and Information, was born in Giza, Egypt, in 1985.She received the Bachelor in Computer Science degree from the Helwan University in 2006 and the Master in Computer Science degree from the Helwan University in 2010.She is currently pursuing the PH.D degree with the Department of Computer Science.