DEVELOPING A CASE-BASED RETRIEVAL SYSTEM FOR SUPPORTING IT CUSTOMERS

Essam S. Hanandeh

Computer Information System, Zarqa University, Zarqa, Jordan

ABSTRACT:

The case-based reasoning (CBR) approach is a modern approach. It is adopted for designing knowledge-based expertise systems. The aforementioned approach depends much on the stored experiences. These experiences serve as cases that can be employed for solving new problems. That is done through retrieving similar cases from the system and utilizing their solutions. The latter approach aims at solving problems through reviewing, processing and applying their experiences. The present study aimed to shed a light on a CBR application. That is done to develop a new system for assisting Internet users and solving the problems faced by those users. In addition, the present study focuses on the cases retrieval stage. It aimed at designing and building an experienced inquiry system for solving any problem that internet users might face when using a case-based reasoning (CBR) dialogue. That shall enable internet users to solve the problems faced when using the Internet. The system that was developed by the researcher operates through displaying the similar cases through a dialogue. That is done through using a well-developed algorithm and reviewing the relevant previous studies. It was found that the success rate of the proposed system is high.

KEYWORDS:

IT support, similarity metric, case-based reasoning conversational case-based reasoning, IT problem management.

1. INTRODUCTION

When seeking to develop a CBR system, it is necessary to have the ability to learn from the previous experiences and cases. When employing CBR systems, one shall encounter problematic cases. If the solutions to such cases are examined and tested and a specific success rate is idenitfied for them, then those cases and their solution could be added to a cases base. Their solutions could be utilized for solving problem in the future. After adding the cases to a cases base, a CBR system can reason with having a variety of cases. In addition, a CBR system shall enable people to avoid committing the mistakes committed earlier. The latter system shall record failures (cases) and their reasons and solutions. Such information could be utilized for predicting any potential problem or failure in the future [3].

There are many methods used for describing the mechanism that a CBR system operates through. It should be noted that the CBR system consists from 3 stages. These stages are presented below [5]:

2. THE ADDITION OF NEW CASES TO THE MEMORY BASE

Cases are predictions or attributes used for representing the stored experiences. When saving the experiences into the memory, the CBR system can identify the case through searching for cases of similar indexed features. The indexed entries and features are usually constant [5].

3. THE IDENTIFICATION OF SIMILAR CASES

Throughout this process, the features used for identifying the new case must be specified and compared with similar indexed features of other cases.   After that, the CBR system shall retrieve the case that matches the new cases from the memory. The system depends on scale and symmetry in the retrieval of similar cases [9].

4. THE IMPLEMENTATION OF THE CASES THAT WERE MENTIONED

The system aims at measuring the extent of similarity between the new case and the cases saved in the cases base. That is done for selecting the best matching solutions and modifying them that can solve the new problem. The retrieval stage is the most important stage of the stages of a CBR system. To retrieve the matching case, the system must have a big cases base. Through the retrieval stage, the best matching case(s) shall be idenitfied and retrieved. Improving the retrieval process is a problem that many CBR-related studies dealt with [9]. Carrying out the retrieval stage requires using measurement techniques for measuring the extent of similarity between the current case and other similar cases. There are several measurement techniques used for measuring the extent of such similarity. Using effective measurement technique is importatn. That is because displaying inappropriate cases shall lead to producing erroneous results by the system [2].
Having a CBR system is important to solve the problems faced by anyone using internet. That is needed because most Jordanian Internet providers recruit few employees.

The present paper is organized into several sections. Section 2 presents a description for the related works that concern knowledge base and case-based reasoning. As for section 3, it presents a description for the study's methodology. Section 4 presents the study's objectives, whereas section 5 presents a discussion for the knowledge base system design. Section 6 presents a description for the conversational case-based reasoning system. Section 7 identifies the basic principles of the CBR process. Section 8 identifies the study's results and experiment. Section 9 display the study's recommendations and conclusion

 The present study aimed at proposing and designing an experienced inquiry system for solving the problems faced by internet users. This system is based on a database that includes cases and solutions saved in an earlier time. Through such a system, similar cases be retrieved and theirs solutions can be adjusted to fit the new problem. The researcher aimed to develop a system for achieving the following:

1. Solving problems efficiently and accurately with exerting the minimal amount of effort.
2. Identifying the most effective solutions for any problem faced in any cases.
3. Developing an interface for users that doesn't require having a prior knowledge nor prior experience.
4. Having a system that can be used at any time for solving problems and reducing the duration needed for searching

5. REVIEW OF LITERATURE

A researcher aimed at analysing and assessing the university students' performance. That is done through the application of different data mining classification techniques by using WEKA tool. The accuracy of the classifier algorithms depends on the nature and size of the data. Five classifiers were used; Bayesian Network, NaiveBayes, ID3, Neural Network and J48. Several

performance measures were employed for comparing the results of these classifiers. The results indicate that the Bayesian Network classifier shows the highest accuracy level [4].

A researcher [12] aimed to propose a hybrid CBR system. It was proposed through developing a reduction technique in feature selection and conducting cluster analysis in case organization. Through the study of [12], a minimal set of features is chosen from the problem domain. The redundant features are reduced through the neighbourhood rough set algorithm. After conducting the feature selection process, the growing hierarchical self-organizing map (GHSOM) is adopted as a cluster tool. This map organizes cases and divide the initial cases base into small subsets with a hierarchical structure. The system shall search in the case base for cases that have features that are similar to the new case. The effectiveness of the proposed approach was proved through conducting experiments on UCI datasets. It was also proved through conducting experiments on a practical case in electromotor product design. It was concluded that the developed techniques can enhance the performance level of the CBR system. [12].

A researcher aimed to propose an approach for measuring multiple intelligences (or MI) value as a co-determiner for predicting the academic achievement of students in a computer programming course. The predictive model's effectiveness is checked on three machine learning algorithms; a)- support vector machine, b)- artificial neural network and c)- the classic Naïve Bayes. After employing 3 validation schemes (i.e. 2, 5, and 10-folded cross-validations), it was concluded that the Mi-inclusive model can improve the accuracy of the predictions made about students' achievement. These prediction involve 3 levels; good, average, and poor levels. [7].

A researcher aimed at reviewing and assessing several data mining techniques that are used for predicting students' achievement level. He also aimed to shed a light on the way of utilizing prediction algorithm for identifying the most significant attributes of data. Through using educational data mining techniques, students' achievement and success can be improved. Such techniques shall be useful for students, educators and academic institutions. [1].

A researcher aimed to provide an approach for the retrieval of novel information. It was developed for conducting a personalized itinerary search. The proposed approach is derived from integrating 3 techniques; Choquet integral, Case Base Reasoning and ontology. The proposed approach is characterized with these features: (1) user-oriented ontology is employed as a knowledge source for extracting pertinent information about stakeholders' needs and preferences ; (2) semantic web rule language is employed for providing the system with enhanced semantic capabilities and supporting personalized case representation; (3) a CBR-personalized mechanism for retrieving data is developed for providing users with an optimum itinerary that can meet their preferences and personal needs . It was found that the developed system can enhance the accuracy of the case retrieval process. It was found that the developed system can reduce the time required for conducting the latter process. The developed system can fulfil the needs of stakeholders, facilitate the searching process and improve the accuracy of the case retrieval process [8].

Montani [17] proposes a CBR approach for supporting the management of exceptions in the business process execution. He developed a method for the semantic retrieval of data-centric workflows from a repository [16]. Conversational CBR has been applied in the tool CBR Flow [15] for guiding users in the process of adapting the workflow to change circumstances.

6. MATERIALS AND METHODS

It is necessary to have a system of high quality for receiving and handling thousands of calls. That is because solving problems via phone is costly and requires much time. Hence, there is a need to

develop an efficient system that is user-friendly and low-cost. The proposed system employs a CBR dialogue for solving problems faced by native speakers of Arabic when using internet. The latter approach also aims at:

1. Analysing and interpreting the text inserted by the user for expressing the problem he / she is facing when using the internet. That is done through describing the problem fully or partially. Then, the system shall separate the keywords from the text.

2. Retrieving the most similar cases that are saved in the system. Such cases may match the new problem fully or partially.
3. Making questions for providing a dialogue that shall enable the user to express and describe the concerned problem effectively to find the appropriate solutions.
4. Proving users with more opportunities for reviewing solutions related to the problem, and benefiting from the solutions saved in the system.

7. KNOWLEDGE BASE SYSTEM DESIGN

The researcher of the present study chosen a rule-based knowledge representation method for representing the relationship existing between rules and facts in the form of IF-THEN. Through the latter method, set of symbolic rules represent knowledge. In addition, inference in the system is carried out by a process of chaining through rules recursively, either by reverse or forward reasoning [13]. There are numerous programming tools and languages used for developing a knowledge-based system. Such languages and tools may include: LISP, Clip and Prolog [12]. Lisp and Prolog are considered excellent and prominent programing languages. LISP is a programming lannguages used for building a program that imitates humans' behavior. Prolog focuses on expert systems and knowledge bases. Lisp is an old high-level programming language. It is a general purpose language. It involves many sets of data types and a built-in data structures, such as: lists, maps and trees. On the other hand, Prolog is a declarative programming language that is specialized in inference-based reasoning. It employs a specialized logic for generating new knowledge from a set of rules called (knowledge base). The latter language is suitable for handling tasks related to the development of mathematical theorems, expert systems and knowledge.

8. CONVERSATIONAL CASE-BASED REASONING SYSTEM

It is a type of CBR interactive system through which users enter a partial description of the concerned problem. The latter description is entered in the form of a text. The CCBR-based system responds through presenting the relevant possible solutions of the stored problems in the cases base [9]. Through such a system, a decision tree is used for covering all the relevant cases to the query (figure 1).
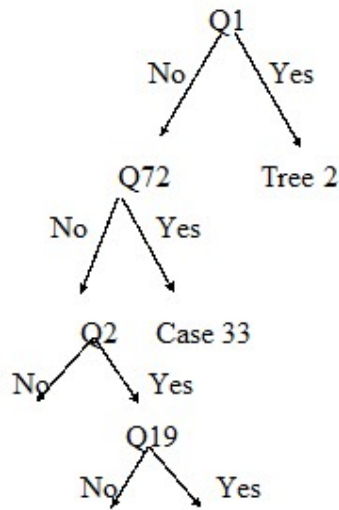
Figure. 1. The Decision Tree

CCBR systems gained success through enabling users to design and develop programs and applications. Such systems store knowledge as cases, rather than rules. Within most of the CBR systems, users are expected to enter the description of the entire problem at the outset. This requires having a prior detailed domain knowledge. In contrast, CCBR systems require inputting a brief free-text description of the faced problem only [5].

Then, the system shall conduct an interactive assessment for the problem. It should be done to construct a query. Through that conversation, the system shall rank and display the potential solutions and cases according to the extent of matching the case (and the displayed questions). Thus, the user needs to provide answers to the given questions only. In such a case, having a prior knowledge about the extent of relevancy isn't needed [6].

This method is complicated because it is a challenge to identify which questions and cases are relevant during the conversation. Displaying irrelevant choices for questions shall hinder the provision of effective diagnosis for the problem. Displaying irrelevant choices for cases shall also hinder the retrieval of an effective solution. The CCBR system must launch a conversation that is relevant. It must also provide relevant multiple answers according to provide a description for the faced problem. It is essential to have useful guidelines. Such guidelines require substantial expertise to be applied correctly [2].

The decision tree includes yes or no questions that are displayed for solving the concerned problem. There are 2 methods for carrying out the process of choosing the most relevant cases. The first methods refers to the use of keywords porvided in the text. The second method is the adoption of user's answers which shall leads to narrowing down the potential choices. Narrowing down the potential choices shall enable the user to find the sought solution quickly as it is seen in table 1.

Table 1. File of Questions & Cases

No. of questions	Yes Cases	No Cases
1	#23,34,77,.............	#11,#13,48,.......
88	#78,#13,......	#25,#55,#22,.......

Table (2) presents several records that are obtained from a case base.

Table 2. Records obtained from a case base

Case No.	Keyword 1	Keyword 2	Description	Solution
1	Modem	Speed	Change of Modem Speed to....	....
33	Network	Card	Network Cards not identified	....
71	Server	Error	Server Busy	....

9. THE BASIC PRINCIPLES OF CBR SYSTEMS

1. Summary: It refers to a description provided about the case. This description is brief
2. Attributes: They refer to a set of answers and questions which concern the case.
3. Solutions: They refer to a series of steps that are followed for solving a problem.

The system shall rank and index cases for facilitating the process of searching for similar cases and their solutions. Users of a CCBR system must only input a partial description for the concerned problem. In other words, the user doesn't need to input a full description for the concerned problem [11].

After that, the system shall clarify and illustrate the concerned problem through an interactive conversation. The dialogue shall end when the user selects a solution for his problem from the displayed solutions. It should be noted that users interact with the CCBR systems through a conversation. In such a conversation, a set of questions shall appear. The user must choose an answer for the displayed questions. Furthermore, the displayed cases shall be ordered in accordance with the extent of matching the features of the faced problem. Through displaying a set of potential cases, the user shall find a solution for the problem he is facing. Thus, the conversation shall be ended successfully [10].

The present study aimed at designging and implementing an experienced query system. That is done in the aim of helping internet users to solve the problems they face when using the internet. The researcher aimed to achieve that through providing the user with information about cases faced in an earlier time. Such information is stored in a system that simulates human behavior. This approach is considered effective and convenient. It doesn't require much effort nor knowledge engineering. The present study focuses on the use of multi-decisions trees attached to

specific cases in order to cover the threaded base cases. The use of several trees adds flexibility to the system. It also provides the user with many opportunities to find the sought solution

10. RESULTS AND DISCUSSION

A sample of ten problems were selected to test the effectiveness of the study's system. table 3 shows the results obtained from the implementation of the system to solve the sampled problems

Table 3. Results and by width ratio (1/1)

Watching	Success/failure	Steps	Number of cases presented	A Number of related cases
1	1	6	14	14
2	1	8	13	13
3	1	4	10	10
4	1	5	20	20
5	1	10	3	3
6	1	8	6	6
7	1	7	10	10
8	1	5	12	12
9	0	15	3	3
10	1	9	9	9

The success rate: 9/10 Steps rate: 7.7

Table 3 presents the success rate, and the numbers of steps, presented cased. It also presents the number of related cases with having a width ratio of (1/1)

Table 4. Results and by width ratio (1/2)

Watching	Success/failure	Steps	Number of cases presented	Number of related cases
1	1	6	7	14
2	1	5	7	13
3	1	8	5	10
4	1	6	10	20
5	1	7	2	3
6	1	6	3	6
7	1	7	5	10
8	1	7	6	12
9	0	12	2	3
10	1	7	5	9

The success rate: 9/10 Steps rate : 7.1

Table 4 presents the success rate, and the numbers of steps, presented cased. It also presents the number of related cases with having a width ratio of (1/2)

Table 5: Results and by width ratio (1/3)

Watching	Success/failure	Steps	Number of cases presented	Number of related cases
1	1	6	5	14
2	1	6	4	13
3	1	8	3	10
4	1	7	7	20
5	0	10	1	3
6	1	10	2	6
7	1	11	3	10
8	1	9	4	12
9	0	12	1	3
10	1	8	3	9

The success rate: 8/10 Steps rate : 8.7

Table 5 presents the success rate, and the numbers of steps, presented cased. It also presents the number of related cases with having a width ratio of (1/3).

11. CONCLUSION

It was concluded that the success rate is 80 %. The latter percentage represents the extent of the system's success in solving problems. It means that the system is able to solve 80 % of problems. These problems can be solved through displaying the cases that match the concerned case. Through using symmetric similarity metric and Spearman's Rank-Order Correlation, it was concluded that there's a positive strong correlation existing between the system's level of success and the number of relevant cases. However, it was concluded that there is a negative correlation between the number of steps and the rate of the display. It was also concluded that there is a negative correlation between the steps and the system's success. If the rate of display increases, the number of steps shall decrease and the success rate shall increase.

When the symmetric similarity metric is absent, there shall be a positive strong correlation between the system's success and the number of steps. In addition, there is a strong positive correlation between the system's success rate and the number of the relevant displayed cases. In simple words, when the number of displayed cases increase, the number of steps shall decrease and thus, the success rate shall increase. In short, the symmetric similarity metric contributes to raising the system's success rate. The absence of symmetric similarity metric shall affect the number of steps, available cases and efficiency of the system

ACKNOWLEDGMENT

REFERENCES

[1]   A. M. Shahiri and W .Husain, "A Review on Predicting Student's Performance Using Data Mining Techniques" Proceeding Computer Science, vol. 72, (2015), pp. 414-422.

[2] Mobyen A., Shahina B., Peter F., "Case Studies on the Clinical Applications using Case-Based Reasoning", Proceedings of the IEEE Federated Conference on Computer Science and Information Systems, pp. 3–10, 2012

[3] Marling C., Whitehouse P., "Case-based reasoning in the care of Alzheimer's disease patients", 4th International Conference on Case-Based Reasoning (ICCBR 01), pp. 702-715, 2001.

[4] Nichat, A.A. and D.A.B. Raut, Analysis of Student Performance Using Data Mining Technique. International Journal of Innovative Research in Computer and Communication Engineering, 2017. 2007(An ISO 3297): p. 5.

[5] Bichandaritz I., Montani S., "Advances in case-based reasoning in the health sciences", Artificial Intelligence in Medicine, vol. 51, no. 2, pp. 75-146, 2011.

[6] Corchado J., Bajo J., Abraham A., "Improving healthcare delivery in geriatric residences", IEEE Intelligent Systems, vol. 23, no. 2, pp. 19-25, 2008.

[7] Ninrutsirikun, U., Watanapa, B., Arpnikanondt, C. and Phothikit, N. (2016) Effect of the Multiple Intelligences in Multiclass Predictive Model of Computer Programming Course Achievement. 2016 IEEE Region 10 Conference.

[8] Bouhana, A., Zidi, A., Fekih, A. Chabchoub, H. and Abed, M. (2015). An ontology-based CBR approach for personalized itinerary search systems for sustainable urban freight transport. Expert Systems with Applications, 42, 3724-3741

[9] Begum S., Ahmed M., Funk P., Xiong N., Folke M., "Case-Based Reasoning Systems in the Health Sciences": A Survey of Recent Trends and Developments, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 41, no. 4, pp. 421 – 434, 2011.

[10] Montani S., Portinale L., Leonardi G., Bellazzi R., "Case-based retrieval to support the treatment of end stage renal failure patients", Artificial Intelligence in Medicine, vol. 37, no. 1, pp. 31-42, 2006.

[11] Jirapond T., Veera B., Puntip P., "A new similarity measure in formal concept analysis for case-based reasoning,‖ Journal of Expert Systems with Applications, vol. 39, no. 1, pp. 967–972, 2012.

[12] Zhu, G.N., Hu, J., Qi, J., Ma, J., Peng, Y.H.: An integrated feature selection and cluster analysis techniques for case-based reasoning. Eng. Appl. Artif. Intell. **39**, 14–22 (2015).

[13] Aronson J., and Turban E., (2007). Decision support system and intelligent systems, 7th Ed., New Delhi, Prentice-Hall of India Private Limited.

[14] Mohamed R., and Julie T., (2009). Knowledge Based Expert Systems in Bioinformatics, France. Retrieved from http://www.intechopen.com/source/pdfs/6704/InTechKnowledge_based_expert_systems_in_bioinformatics.pdf on September 28,2011.

[15] Weber, B., Wild, W., Breu, R.: Cbrflow: Enabling adaptive workflow management through conversational case-based reasoning. In: Funk, P., González Calero, P.A. (eds.) ECCBR 2004. LNCS (LNAI), vol. 3155, pp. 434–448. Springer, Heidelberg (2004).

[16] Gil, Y., Kim, J., Florenz, G., Ratnakar, V., Gonzalez-Calero, P.: Workflow Matching Using Semantic Metadata. In: Proc. of the 5th Int. Conf. on Knowledge Capture (K-CAP), pp. 121–128. ACM, New York (2009)

[17] Montani, S.: Prototype-based management of business process exception cases. Applied Intelligence, Springer Science + Business Media, LLC 2009 (2009) (Published online).