CASE BASED REASONING SYSTEM FOR COURSE ADVISOR IN HIGHER INSTITUTION

Okeoma Chinwendu A
okeoma23@gmail.com

Mgbeafulike Ike
ike.Mgbeafuike@gmail.com

ABSTRACT
Each university or college provides degrees as bachelors and masters degrees. Each degree consists of courses or subjects that are taken in specific period (semester). Academic advising is process of selecting the courses that academic student will register in each semester to fulfill the degree requirement. The academic advisor suggests to the student which course to register. One of the important processes in student educational life is the advising process. This paper proposes and implemented a case based reasoning (CBR) system that recommends to the student the most suitable major in his case, after comparing the historical cases by the student case. The system converts each departmental course into a group of concepts for each course. The system checks the similarity between the student taken course and the stored course in each department. CBR system had proven its effectiveness in the transfer cases between major. Based on the CBR system recommendation, the student can take a decision which major is the best based on the achieving level.

1.0 INTRODUCTION

Haring information is an essential step in the educational services. Knowledge sharing, knowledge management are considered emerging concepts that different researchers discuss. Academic students should register courses based on their profiles and each major they prefer. For this reason, academic advising makes the students achieve their educational goal. There a relationship between the student and the academic advisor in which the advisor plays the facilitator role. Manual academic advising has many drawbacks as labor intensive, time consumption, human advisors use their accrued knowledge and the large number of students compared to the number of

The aim of this research is to produce a Content Based Recommender (CBR) system that avoids the manual academic process and converts this process into an automated one. The implemented system aim to reduce the following manual advising process: limited number of advisors (labor intensive), complication of the advising double curricula (double major) process.

2.0 Literature Review

Case-based reasoning (CBR), broadly construed, is the process of solving new problems based on the solutions of similar past problems (Wang & Shao, 2004). It has been argued that case-based reasoning is not only a powerful method for computer reasoning, but also a pervasive behavior in everyday human problem solving; or, more radically, that all reasoning is based on past cases personally experienced. This view is related to prototype theory, which is most deeply explored in cognitive science. While much of the inspiration for the study of case-based reasoning (CBR) came from cognitive science research on human memory (Schank, 1982), the resulting methodology has been shown to be useful in a wide range of applications (Watson, 1997). Unlike most problem solving methodologies in artificial

intelligence (AI), CBR is memory based, thus reflecting human use of remembered problems and solutions as a starting point for new problem solving. An observation on which problem solving is based in CBR, namely that similar problems have similar solutions (Leake, 1999), has been shown to hold in expectation for simple scenarios (Faltings, 1997), and is empirically validated in many real-world domains. Solving a problem by CBR involves obtaining a problem description, measuring the similarity of the current problem to previous problems stored in a case base (or memory) with their known solutions, retrieving one or more similar cases, and attempting to reuse the solution of one of the retrieved cases, possibly after adapting it to account for differences in problem descriptions. The solution proposed by the system is then evaluated (e.g., by being applied to the initial problem or assessed by a domain expert). Following revision of the proposed solution if required in light of its evaluation, the problem description and its solution can then be retained as a new case, and the system has learned to solve a new problem.

In a nutshell, CBR is reasoning by remembering: previously solved problems (cases) are used to suggest solutions for novel but similar problems.

Kolodner(1996) lists four assumptions about the world around us that represent the basis of the CBR approach:

1. Regularity: the same actions executed under the same conditions will tend to have the same or similar outcomes.
2. Typicality: experiences tend to repeat themselves.
3. Consistency: small changes in the situation require merely small changes in the interpretation and in the solution.
4. Adaptability: when things repeat, the differences tend to be small, and the small differences are easy to compensate for.

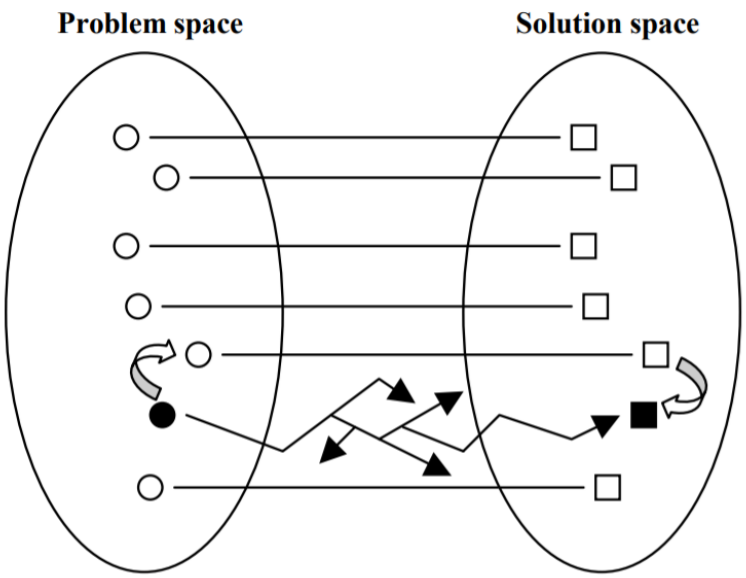

Figure 1 Relationship between problem and solution spaces in CBR. Adapted from (Leake, 1996).

- ● = description of new problem to solve
- ○ = description of solved problems
- □ = stored solutions
- ▪ = new solution created by adaptation

As will be clear from the literature review we present in this article, aspects of reuse and retention, and to a lesser extent revision, have also attracted significant research interest. Several books examine fundamental aspects of CBR and present case studies of research and applications (Riesbeck & Schank, 1989; Kolodner, 1993; Leake, 1996; Watson, 1997; 1998; Bergmann, 2002). However, our aim in this article is to present a concise summary of research that focuses on the problem-solving cycle in CBR, including some of the most recent advances.

Once the currently encountered problem is described in terms of previously solved problems, the most similar solved problem can be found. The solution to this problem might be directly applicable to the current problem but, usually, some adaptation is required. The adaptation will be based upon the differences between the current problem and the problem that served to retrieve the solution. Once the solution to the new problem has been verified as correct, a link between it and the description of the problem will be created and this additional problem solution pair (case) will be used to solve new problems in the future. Adding of new cases will improve results of a CBR system by filling the problem space more densely. (Smyth & Keane, 1998).

1. Process of Case-Based Reasoning

Case-based reasoning has been formalized for purposes of computer reasoning as a four-step process (Althoff, 1994):

1. **Retrieve:** Given a target problem, retrieve from memory cases relevant to solving it. A case consists of a problem, its solution, and, typically, annotations about how the solution was derived. (Kolodner, 1993). For example, suppose Fred wants to prepare blueberry pancakes. Being a novice cook, the most relevant experience he can recall is one in which he successfully made plain pancakes. The procedure he followed for making the plain pancakes, together with justifications for decisions made along the way, constitutes Fred's retrieved case. (Schank.1990; Leake, 1992). One approach to reducing retrieval time, as in the pioneering work of Stanfill & Waltz (1986), involves the use of massively parallel computers. While the requirement for expensive hardware is an obvious drawback, the approach still guarantees finding the maximally similar cases by performing an exhaustive memory search. Stanfill & Waltz describe the implementation of a memory-based reasoning algorithm on a fine-grained SIMD parallel machine. Their Connection Machine performs a highly parallel search for similar cases and was applied to the problem of pronouncing English words using a case memory containing thousands of examples of correctly pronounced words. Another approach to reducing retrieval time relies on the organization of cases in memory. For example, Wess (1994) propose an approach to retrieval in which organization of the case memory is based on similarities between cases. A binary tree called a k-d tree is used to split the case memory into groups of cases in such a way that each group contains cases that are similar to each other according to a given similarity measure. To ensure that the most similar cases are retrieved, the retrieval algorithm computes similarity bounds to determine which groups of cases should be considered first. Smyth & McKenna (1999; 2001) propose an alternative model of case retrieval that is informed by the availability of an explicit model of case-base competence (Smyth & McKenna, 1998; 2001). The so-called footprint-based retrieval algorithm is a two-stage retrieval approach that searches two distinct populations of

cases. First, it involves the search of a small subset of so-called footprint cases which have been identified as providing a covering set for the entire case base (i.e., can solve the same set of problems)

2. **Reuse:** Map the solution from the previous case to the target problem. This may involve adapting the solution as needed to fit the new situation. In the pancake example, Fred must adapt his retrieved solution to include the addition of blueberries. Aamodt & Plaza's (1994). Hammond (1990) describes the reuse of recipes in CHEF, a menu-planning system. Substitution adaptation is used to substitute ingredients in the retrieved recipe to match the menu requirements (e.g., when a recipe containing beef and broccoli is retrieved for a meal requiring chicken and snow peas, the meat component is replaced by chicken and the vegetable component is substituted by snow peas). Transformation adaptation may also be needed to amend the proposed recipe further by adding or removing steps in the recipe that result from any ingredient substitutions (e.g., for chicken, rather than beef, a new skinning step should be added). Further transformations may occur at the revise stage where critics analyze the failure of a recipe and repair strategies are applied to the proposed recipe to add or remove steps in the failed recipe. CHEF's learning of critics introduced the topic of case-based planning and many of its themes (e.g., indexing, use of cases in memory, failure-driven learning). SWALE (Schank & Leake, 1989) is a case-based explanation system for story understanding that reuses old explanations by applying substitution adaptation to amend the actor, their role or the action in the retrieved explanation. Transformation adaptation may again be needed to add or remove components in the current explanation resulting from these substitutions. Déjà Vu (Smyth 1995) is a CBR system for the automated design of plant control software. It builds on some of the ideas proposed by utilizing transformation adaptation knowledge in the form of general adaptation strategies and more specialized adaptation specialists.

3. **Revise:** Having mapped the previous solution to the target situation, test the new solution in the real world (or a simulation) and, if necessary, revise. Suppose Fred adapted his pancake solution by adding blueberries to the batter. After mixing, he discovers that the batter has turned blue an undesired effect. This suggests the following revision: delay the addition of blueberries until after the batter has been ladled into the pan.

4. **Retain:** After the solution has been successfully adapted to the target problem, store the resulting experience as a new case in memory. Fred, accordingly, records his newfound procedure for making blueberry pancakes, thereby enriching his set of stored experiences, and better preparing him for future pancake-making demands. In the classic review paper by Aamodt & Plaza (1994), retention is presented as the final step in the CBR cycle, in which the product of the most recent problem-solving episode is incorporated into the system's knowledge. To a great extent this has traditionally translated into a variety of approaches for recording the product of problem solving as a new case that can be added to the case base. Of course, there are various issues concerning how best to learn a new case and different systems record different types of information in their cases. Most, for example, simply record the target problem specification and the final solution, with the implicit assumption that

the outcome was successful. For example, when CBR is integrated with a generative problem solving system for speed-up learning, the success of the system's solutions may be guaranteed When outcomes are less reliable or when the criteria for success are more complex, case representations must include additional information on the outcome of the solution, which may also include fine-grained information on how well the solution addressed many dimensions of the system's goals Aleven, (2003)).

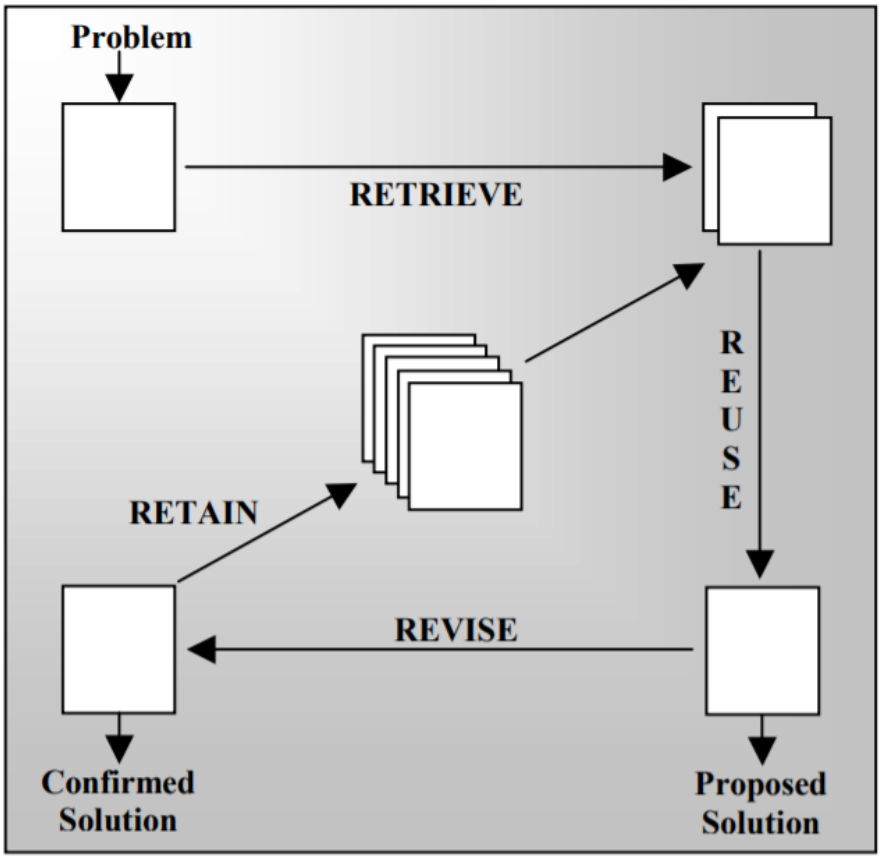

Fig. 2: The CBR cycle. Adapted from (Aamodt & Plaza, 1994).

Figure 2. shows Aamodt & Plaza's (1994) classic model of the problem solving cycle in CBR. The individual tasks in the CBR cycle (i.e., retrieve, reuse, revise, and retain) have come to be known as the "4 REs". Because of the pivotal role of retrieval in the CBR cycle, a considerable amount of research has focused on retrieval and similarity assessment. As illustrated in Figure 2, Leake (1996b) expresses the role of similarity through the concepts of retrieval and adaptation distances. Also captured in Leake's diagram is the relationship between problem and solution spaces in CBR.

2. The utility problem in CBR

In the past the prevailing view of case learning in CBR was based on the assumption that learning would occur as a by-product of every problem solving episode. However, as CBR systems were developed and deployed for real-world application scenarios, the potential pitfalls of long-term case learning became apparent, especially in relation to the impact of case-base growth on retrieval costs. This is an example of the utility problem identified in

explanation based learning research (Arcos, (1997)This problem refers to the performance degradation experienced by speed-up learners as a result of learning control knowledge. In brief, Minton demonstrated how rules learned for reducing problem solving time, by directing the search more carefully, might ultimately degrade overall system performance as the time spent considering the application of a speed-up rule eventually overtakes the time needed for first principles problem solving. For example, overly specific rules that are seldom applicable, or rules with a high match cost, or rules that offer limited speed-up were all found to contribute to a decline in problem solving efficiency. At the heart of the utility problem is a natural trade-off between the benefits of speed-up knowledge and the cost of its application. A similar trade-off also exists in CBR systems (Smyth 1995). Cases correspond to a form of speed-up knowledge in the sense that retrieval and reuse of similar cases are expected to provide more efficient problem solving than first-principles methods, with additional cases increasing the range of problems that can be solved rapidly. However, this rather naive view of case knowledge fails to consider retrieval costs. In CBR systems the utility problem is caused by the conflict between the average savings in adaptation effort due to the availability of a particular case, which tends to increase efficiency as the case base grows, and the average retrieval time associated with a given case-base size, which tends to decrease efficiency. Smyth (1996) demonstrate the inevitability of the utility problem in CBR under reasonable general assumptions about the retrieval and reuse characteristics of a CBR system. They show that as a result of case learning, retrieval efficiency (mean retrieval time) tends to degrade while adaptation efficiency (mean adaptation time) is seen to improve, but at an ever decreasing rate. Initially, as a case base grows each newly learned case can have a significant impact on adaptation as it is more likely to improve overall case-base coverage. However, as the case base grows new cases are more likely to overlap with existing cases and so offer little in the way of new coverage and minimal adaptation savings. As new cases are added retrieval costs become progressively greater but adaptation savings progressively less. Eventually the increase in retrieval time as a result of a new case addition is greater than the adaptation savings offered.

Advantages and Disadvantages of Case-Based Reasoning

Advantage:

a) Ability to encode historical knowledge directly
b) Achieving speedup in reasoning using shortcuts
c) Avoiding past errors and exploiting past successes
d) No strong requirements for an extensive analysis of domain knowledge.
e) Added problems solving power via appropriate indexing strategies
f) Highlights important features
g) Domains do not need to be completely understood
h) Solutions are quickly proposed.

Disadvantage:

a) Old cases may be poor
b) Library may be biased
c) Most appropriate cases may not be retrieved
d) Retrieval/Adaptions Knowledge still needed.

RELATED WORK

The following subsections will define the following main ideas: researches in (educational services, academic advising and different techniques for automated advising).

A. Educational Services E-advising has many benefits as stated in E-advising enhances online student retention; handle the communication between the student and the advisor in an easy way with different tools: text/chat mode, an audio mode. E-advising provides timely, quality advising services and it provides an innovative response to questions using personalizing communication in anytime, anyplace since the information will be available on the web. In e-learning systems, student can search for courses using the recommender of e-learning system. The recommendation system collects information from the log of student history and course ontology that shows detailed information of each course and its pre-requisites courses. Learning recommendation and course modification can assist students in their learning performance, also it can handle the students' learning behavior, evaluating the in e research a web education tool was created for covering the course materials and getting feedback from teachers and students. Data mining definition is the extraction of hidden information from large databases Data mining is used in academic advising based on the following researches. Another field that can use data mining can is the major selection Student syllabi is the document that include all the courses taken by the student, and their grades, also it includes the rest of the courses that should be taken to graduate from the specified major. E-advising process is the process of advising automation, different users of the e-advising process should evaluate the quality of e-advising.

B. Academic Advising

Academic advising "step is defined as the process of supporting, motivating student's thought-out university's study plan and along the achievement of their educational goals" Major selection is a very important step in the academic student life. In the researchers created a decision support system. The mechanism of the system works as follows: the system calculates the supporting degree (passing mark) for each course; second the system counts it to reach the whole major supporting degree. At the end, the system suggests the major of highest supporting degree to the student. The expertise of the academic advisor is so important. My Majors is a web application that can be used by the students or the academic advisors. My Majors proposed benefits are: reduces the time consumed in advising process, it improves the enrollment of 24 hours a day, every day. In medical Iranian school, research in observes the academic advising process. Most of medical students do not know the main tasks of the faulty advisors. Educational workshops were established for the sake of students and to understand brief information about roles.

C. Case based Reasoning in Academic Advising

Previous researches identified the most of academic advising problem and obstacles. The techniques include: Decision Support System, Decision Tree, Data Mining techniques, Decision Matrix, and Rule-based Reasoning. Knowledge management definition depends on the nature and needs of the business or organization; however knowledge is the management of the cognitive production factors accommodated in a business or government organization. Knowledge provides number of the benefits in education: student's information will be increased through capturing, storing and sharing of

knowledge. The proper and efficient knowledge sharing cannot be done using the human interaction only; however a systematic approach should be used. Different automated techniques such as ontology (known tools are used for the enhancement of the educational service and the academic advising. CBR used for many purposes as robotics, electronics, mechanics and real estate, consumer services, urban planning, medicine, tourism, software development-computer engineering, environmental planning, civil engineering. Textual Case based reasoning (TCBR) is defined as "a research area that deals with solving new problems by reusing previous similar experiences documented as text". Choosing the right data mining technique in real application is a very hard decision. There are two parameters to choose the data mining technique the first is goal of the problem to be solved and the second is the structure of the available data. Classified data mining techniques based on a conceptual map into 4 categories: descriptive model, associative model, discriminator model and predictive model. CBR is classified in the discriminate model based on CBR is a predictive model based on historical data. RubricAce is software used for testing. RubricAce TCBR is developed in for recommending textual feedback for students' assessment. The assumption of this system is "students with similar grades should be given similar feedback".

CONCLUSION

The paper provides a CBR system for academic advising in university system. The CBR system provides solution for transfer cases between university majors. The comparison process indicates the similarly of course contents and such course similarities was used in major matching and calculating the achievements level. A survey was spread and filled in by academic advisors to evaluate its results and compares it with the traditional manual system. The CBR system could be extended to include different majors of different colleges and differentiated course syllabus. Transfer cases between different colleges of different majors could be handled using this paper contribution. The CBR system could include student preference as well prior to recommendation phase in order to filter the results and propose the most convenient major to his/her preferences.

REFERENCES

Aghajari S, "Comparison of Knowledge Management Technologies in Academic environment", In Proceedings of International Conference on Education and Management Technology (ICEMT), Cairo, Egypt, 2010.

Biletskiy Y, Brown A, & Ranganathan G, "Information extraction from syllabi for academic e-Advising",Expert Systems with Applications vol. 36, pp. 4508–4516, Elsevier, 2009.

Binh N, Duong H, Hieu T, Nhuan N, "An integrated approach for an academic advising system in adaptive credit-based learning environment", VNU Journal of Science, Natural Sciences and Technology, vol. 24 ,pp. 110-121, 2008.

Borges A, M. Corniel, Giln R, Ramos L, & Contreras L, "Ontological Model as Support Decisions Making in Study Opportunities: Towards a Recommendation System", In Proceedings of the 4th WSEAS/IASME International Conference on Education Technologies (EDUTE08),2008 .

Deorah S, Sridharan S, Goel S, "SAES- Expert System for Advising Academic Major", IEEE, 2010.

Grupe F, "An Internet-based expert system for selecting an academic major", Internet and Higher Education,pergamon, 2002.

Werghi N, & Kamoun F, "A decision-tree-based system for student academic advising and planning in information systems programmes", Int. J. Business Information Systems, Vol. 5, No. 1, 2010. http://dx.doi.org/10.1504/IJBIS.2010.029477

Zhou Q, Yu F, "Knowledge-Based Major Choosing Decision making for remote students", IEEE International Conference on Computer Science and Software Engineering, 2008.

Henning M, "Students' Motivation to Learn, Academic Achievement, and Academic Advising", PhD dissertation, AUT University, New Zealand, 2007.

Liu C, Chiang S, Chou C, & Chen S, "Knowledge exploration with concept association techniques", Emerald, vol. 34 No. 5, pp. 786-805, 2010.