복잡한 분야의 한정된 데이터 상황에서의 사례기반 추론: 공정제어 분야의 적용

김 형 관(한국과학기술원 테크노경영대학원)

Case Based Reasoning in a Complex Domain with Limited Data: An Application to Process Control

Steven H Kim and Jonghyung Joo
Graduate School of Management
Korea Advanced Institute of Science and Technology
Seoul, Korea
Email: skim@msd.kaist.ac kr

Abstract

Perhaps one of the most versatile approaches to learning in practical domains lies in case based reasoning To date, however, most case based reasoning systems have tended to focus on relatively simple domains The current study involves the development of a decision support system for a complex production process with a limited database

This paper presents a set of critical issues underlying CBR, then explores their consequences for a complex domain. Finally, the performance of the system is examined for resolving various types of quality control problems

Key words: Case based reasoning, adaptation, process control, learning system

Acknowledgements. We are indebted to Mr Key-Mok Shon of the LG Production Research Center In addition, we would like to thank Mr Myung-Kuk Park and Mr. Inbom Lee.

INTRODUCTION

The increasing pace of business activity and the complexity of manufacturing processes highlights the need to deploy learning systems to support decision making Perhaps one of the most versatile approaches to learning in practical domains lies in case based reasoning

To date, however, most case based reasoning systems have tended to focus on relatively simple domains. The current study involves a project to develop a decision support system for a complex production process The case based reasoning module plays a central role in a multifaceted system which includes production rules and statistical techniques. The software tool is designed to reduce the learning curve required by human engineers to attain proficiency in supervising the production of a computer display device

BACKGROUND

Case-based reasoning (CBR) occurs in all areas of human behaviour: engineering, medicine, science law, and politics, to name a few. In any field, practitioners rely on precedents to guide their decisions, whether the problem at hand is mundane or one that calls for innovation (Kim and Novick, 1993) Case retrieval and adaptation are essential functions in case-based reasoning systems (Kolodner, 1985)

At first glance, a straightforward approach to the application domain would seem to lie in the use of neural nets. However, the large number of inputs render the application domain unsuitable for neural networks The profusion of degrees of freedom (i.e. weights) in a neural network would lead to overfitting in conjunction with the small case base of 99 records. For this reason a CBR methodology appeared to be most pertinent approach.

A learning system should make increasingly useful decisions as it accumulates experience This is the express goal of the work in case-based reasoning (CBR) As with any approach, CBR has its advantages and limitations

CASE STUDY

The case study involves the manufacture and assembly of components relating to the color display tube (CDT) of a computer monitor. One of the key components in a CDT is the deflection yoke (DY). The DY has horizontal and vertical coils which generate a magnetic field to deflect electron beams horizontally and vertically. The deflected electron beams are directed onto the CDT screen. The convergence pattern is key to the quality of a CDT. The coil in the DY requires precise positioning for the proper operation of the CDT

Due to variations in the composition of the copper coil as well as the manufacturing process, it is a challenging task to produce a high quality tube. The standard procedure for manufacturing a DY involves coil winding, assembly, and convergence check. A good convergence pattern is the criterion for evaluating the quality of a DY. The quality depends primarily on the coil winding process, a procedure which requires years of experience

The troubleshooters rely on a process of trial and error to expand their knowledge and perfect their skills. During the apprenticeship, novice operators have to spend a lot of time to observe experts and accumulate an appropriate knowledge base.

At present, plans to expand manufacturing capacity at the company are constrained by the availability of such human expertise Moreover, these experts will be retiring in the years ahead, leaving a shortage of skills due to the lack of young engineers interested in this field

For these reasons, the goal of the adaptive system is to support process control by retrieving appropriate exemplars from the case base The CBR software is part of a larger hybrid system which procures precedent cases including the heuristic strategies of experts as well as statistical techniques for process design. While many routine tasks are automated on the factory floor, a number of complex activities such as process supervision are effected manually. Due to heavy demands on their schedule, the human troubleshooters could spare enough time to manually construct only a small database of 99 records for developing the adaptive system

While many routine tasks are automated on the factory floor, a number of complex activities such as process supervision are effected manually Due to heavy demands on their schedule, the human troubleshooters could spare enough time to manually construct only a small database of 99 records for developing the adaptive system.

During the problem solving process, *recognition* stage involves the identification of exceptions from the raw data corresponding to each assembled unit. Any deviations beyond the design threshold is flagged by automated equipment

The *diagnosis* stage involves a mapping of the tolerance exceptions into an identifiable symptom. For instance, excess deviations at a particular set of points might be judged as an "AmPlified Horizontal" (APH) symptom

Unfortunately, multiple symptoms often occur simultaneously In that situation, the nonlinearities and interdependencies among variables make it difficult to identify the correct set of symptoms within the large space of combinations of basic symptoms. For this reason, a novice engineer requires about 4 months to become an expert in the diagnosis stage.

Finally, the *prescription* stage involves the specification of a solution based on the diagnosis For instance, the solution to an APH symptom might be to "Increase the depth of traversal on the coil plunger by 1 mm". In the case of multiple symptoms occurring concurrently in a single defective unit, a straightforward effort to remedy one problem may exacerbate the other problems. Due to the complexity of the problem domain, 3 years of experience are required to attain a moderate level of proficiency and 5 years to become an expert troubleshooter in the prescription stage

In this project, each case consisted of a problem pattern specified in a tabular format This *problem table* is generated during the convergence check process. Every case consists of the observations from over a dozen check points At each check point, the divergence of every pair of colors along the *x* or *y* axis is specified.

Another table of data relates to the *process parameters* which are adjusted by a process control expert. The adjustment of process parameters in response to a problem pattern leads to a *result pattern*

For each attribute of a particular case, the *divergence* is the deviation of that attribute from the nominal value If the divergence is small, no problem exists. On the other hand, a large value of divergence beyond a threshold indicates a problem

The severity of the problem is called the *criticality* If the divergence is less than the threshold for that attribute, the criticality is zero; otherwise the latter is positive An attribute with a positive value of criticality is indicative of a problem The problem is known as a *critical symptom*, or simply *symptom* when no confusion arises

The problem base contains the records of 99 problematic patterns along with their solutions. Another attribute of each problem pattern is its level of severity, which is also stored with each record (Park et al 1997, 1998) The result table includes the evaluation value for the pattern.

RESULTS

The CBR system identifies a set of symptoms based on the raw data, then suggests a remedy based on precedent cases In this way, a relative novice can easily infer an appropriate solution to each problem as it arises.

The potential benefits of the system are even more significant in light of the increasing mobility of employees. In this volatile environment, an engineer may well leave the company without ever attaining an expert level of competence. A pool of semi-qualified engineers is of limited help in running an existing factory or establishing a new one With the software at hand, however it becomes feasible to staff a new factory with troubleshooters having less experience

The performance of the CBR program was evaluated by using the first 70 records of the database as the training set and the remaining 29 records as the test set The results are presented in Table 1. For instance, the last column indicates that there were 9 records exhibiting multiple symptoms In testing the program, 1 case yielded the identical solution for both the target and precedent cases. In addition, 5 cases produced different but workable solutions which represent alternate ways to rectify the problems. Moreover, 3 failures resulted from the lack of a match (i.e. no precedent with the same symptoms) in the casebase.

The first column in Table 1 pertains to a modified program for resolving only the most critical problem. Since the standard system was designed to handle multiple symptoms, it was modified for the test against the most critical symptom. The CBR program produced a match and retrieved a solution in 75 % of the test cases which involved a single symptom. In addition, a slightly different but still acceptable solution was produced for one particular test case. The remaining 4 cases yielded improper matches. Although the system is still under development, its performance is adequate to be of help in guiding novice engineers on the factory floor.

CONCLUSION

The system addresses three benefits. First. the efficiency of the control process is improved by automatically retrieving relevant precedents. Second, the system identifies effective control strategies used by multiple process experts and thereby helps to enhance manufacturing performance. Third, the prescriptive procedure produced by the system is designed to reduce the learning curve for a novice engineer

The goal for the future is to produce a fully automated system for production supervision. In addition, approaches such as neural networks or multistrategy learning may be compared with the current CBR approach in terms of accuracy, efficiency, and robustness.

REFERENCES

Gennari, J. H., et al. "Models of incremental concept formation." *Artificial Intelligence*, v 40(1-3), 1989: 11-61

Ishikawa T. "Analogy by Abstraction: Case Retrieval and Adaptation for Inventive Design Expert System" *Expert Systems with Applications*, v. 10(3/4), 1996 351-6

Kim, S H. and M. B. Novick. "Using Clustering Techniques to Support Case Reasoning." *International J of computer Applications in Technology*, v. 6(2/3), 1993 57-73

Kolodner. J. L "A process model of case-based reasoning." *Proc of the IJCAI-85*, 1985 284-90

Kolodner, J. L. *Case-Based Reasoning* Morgan Kaufmann Publishers, Inc., 1993

Miyashita, K. "Case-based knowledge acquisition for schedule optimization' *Artificial Intelligence in Engineering*, v. 9, 1995: 277-87.

Park, M., K. Shon, S Yoo, and C. Park "Feedback and Diagnostic Aiding System for Deflection Yoke of CDT" *Proc of the 6th International Conference On Production Engineering. Design and Control*, 1997: 715-24

Park, M., I. Lee, S. H Kim, and J. Joo. "A Support System for Process Control: Enhancing Product Quality through Case-Based Reasoning" *Int Conf on Computational Intelligence and Multimedia Applications* '98, Churchill, Australia, Feb. 1998.

Table 1. Analysis of results for CBR system against the test set of 29 records An *identical* solution is one where the retrieved solution contained the same prescriptive procedure as employed in the test case. An *alternate* solution is one which is dissimilar to that in the test case but still rectifies the problem (note that a problem admits multiple types of solutions) For multiple symptoms, the 3 incorrect results were due to the paucity of the knowledge base (i.e. no record in the learning set had the same pattern of symptoms).

	1 symptom	≥ 2 Symptoms
Total # records	**20**	**9**
Acceptable results, of which :	**16**	**6**
* Match	15	1
* Alternate	1	5
Incorrect results	4	3
Hit rate	80 % = 16/20	67 % = 6/9