Knowledge Reuse for Design of a Complex System

Wendi DING, 2022, Embedded Systems, ISAE
Juan José MONTERO JIMENEZ, Sebastien Schwartz, ISAE
Rob VINGERHOEDS, DISC, ISAE, Toulouse, France

1 INTRODUCTION

Case Based Reasoning (CBR) is a branch of symbolic artificial intelligence techniques that focuses on solving a new problem by reusing information and knowledge from a previous similar situation. This reuse of experience allows to reduce time and effort and to have a better confidence in the solution. CBR is intuitive for humans, given that human beings think and reason by using analogies and examples. It is a promising paradigm for developing support systems for complex system, and has been used in a wide variety of knowledge-management applications.

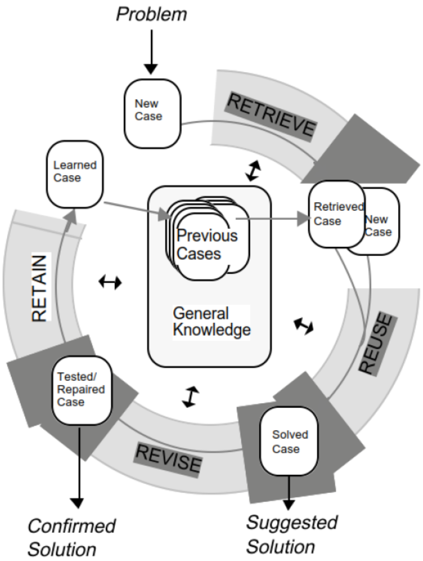

Figure 1: A general CBR cycle [1].

A general CBR cycle can be described by four processes, as depicted in Fig. 1. A new problem is solved by [1]:

- **Retrieving** previously experienced cases that address similar problems.
- **Reusing** the solutions developed from similar cases.
- **Revising** the returned solutions based on the reuse and the evaluation of obtained results
- **Retaining** the new knowledge by incorporating it into the existing case base.

This project is dedicated to a Predictive Maintenance CBR system built in ISAE-SUPAERO and is aimed at developing a decision support system based on CBR for complex system design. The goal is to improve diversity between retrievals and to optimize the case base performance through exploring Case Base Maintenance (CBM) methods. Existing approaches are studied and applied to the predictive maintenance case base, and a Modified Condensed Nearest Neighbor (CNN) method which integrates separation, generalization and reindexing of descriptions and solutions is proposed and proved to be effective in diversifying the retrievals while keeping the advantages of traditional competence-preserving CBM methods.

2 STATE OF ART

Problems such as the utility problem and lack of diversity between retrievals arise with the expansion of the case base, causing long computation time for the retrieval process and redundant retrieval results, which are undesirable for users. Approaches have been explored by researchers to figure out these problems. K. Bradley and B. Smyth [2] defined the diversity measure of a set of cases and proposed Bounded Greedy Selection (BGS) algorithm.

However, the diversification methods integrated into the retrieval process could make the retrieval process very heavy and time-consuming, while Case Base Maintenance (CBM) methods may have a perspective to improve the diversity of retrievals. M.K. Haouchine et al. [3] proposed a competence-preserving case deletion CBM strategy called CM-CNN to guide the construction of smaller case bases. Other attempts include the case addition RC-CNN algorithm proposed by B. Smyth et al., and the case deletion strategy Footprint Deletion (FD) proposed by B. Smyth and M. T. Keane.

3 DEVELOPMENT

3.1 Tools of implementation

The implementation of the CBR program is based on the open-source similarity-based retrieval tool myCBR with its Software Development Kit (SDK). The SDK is integrated in java and the output xml files of case base and retrieval information are read in python. The tool of vpython is employed to realize Case Base Visualization (CBV).

3.2 Investigation of CBM Methods

We first choose to apply some competence-preserving CBM methods. Based on the representativeness assumption [15], we randomly divide the 230 cases in the original case base into 184 cases and 46 queries. A testbench is built in java to collect statistics and experiments are carried out to test the diversification effects of competence-preserving CBM methods.

Then we propose a new CBM method called Modified CNN method which integrates CNN algorithm and hierarchical case base memory structure to further improve diversity between retrievals. We tested this new method in our testbench and collected results concerning case base coverage and competence, similarity and diversity between retrievals compared with existing competence-preserving CBM methods.

3.3 A Modified CNN method

To further improve diversity between retrievals, it is necessary to separate the description and solution spaces to reduce redundant solutions. The Modified CNN method takes the advantages of classical CNN method which aims at reducing the case base size while preserving its competence. The generalization process of descriptions and solutions consists of the following steps:

1) Generalization of Solutions: Generalize the solution with its nearest neighbour if $\mathrm{sim}(v_i, v_j) \geq \theta_{sol}$. Choose the solution with higher performance as parent and make the others as its children.
2) Reindexing of Descriptions and Solutions: The descriptions originally connected to children solutions are reconnected to the parent solution.
3) Generalization of Descriptions: Descriptions are generalized to form Generalized Cases (GCs) if $\mathrm{sim}(u_i, u_j) \geq \theta_{des}$. Original descriptions are stored as subcases of their GCs.
4) Reindexing of GCs and Solutions: All the solutions of subcases of a GC are reindexed to the GC (see Fig. 2(b))

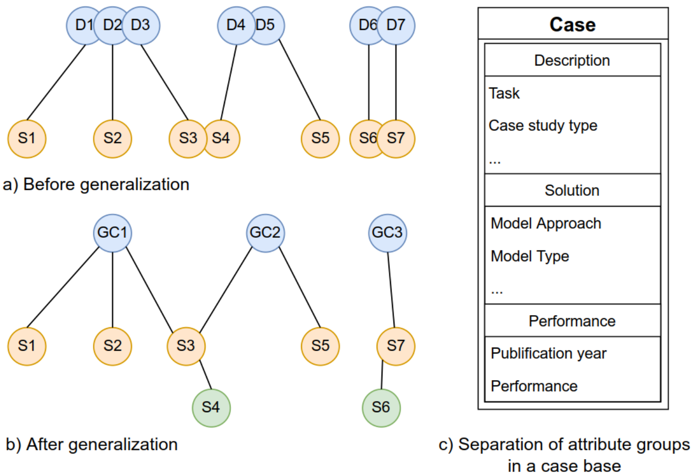

Figure 2: Separation and generalization of solution space and description space

3.4 Force-Directed Graph-Drawing CBV algorithm

We have applied a Force-Directed Graph-Drawing Algorithm to visualise case-bases. It is aimed to model the similarity relationships between cases as on-screen distances. The key issue is to map the n-dimensional cases onto a two-dimensional screen while preserving the similarity relationships between pairs of cases as on-screen distances. The cases are modeled as balls connected by springs whose length is determined by the similarity between cases. Through the iteration of motion functions on vpython canvas, we are able to model the distribution of descriptions and solutions in a three-dimensional space and to show the retrieval results on the canvas.

4 RESULTS

Statistical testing results show that it is possible to reduce the predictive maintenance case base size by up to 65.2% while maintaining the same level of competence by applying competence-preserving CBM methods. In the meantime, diversity between retrievals also rises from 0.5 to 0.7-0.8, which is a desired phenomenon.

Fig. 3 show that the average diversity between retrieved solutions is generally not influenced by θ_{des}, which verifies our assumption that the diversity of retrievals is determined by the distribution of solutions in the solution space, except for the situations of $\theta_{sol} \geq 0.95$, in which case the generalization of solutions occurs so rarely that the solution distribution is mostly indirectly determined by the position and generalization of their descriptions. Compared with competence-preserving CBM methods, the diversity between retrieved solutions is effectively increased with the decrease of the solution generalization threshold θ_{sol} applying our Modified CNN method.

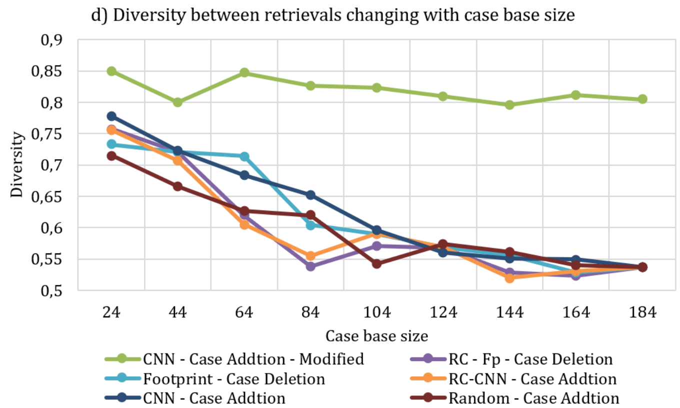

Figure 3: Average similarity and diversity using different methods changing with the number of retrievals.

5 CONCLUSION

In this project, we proposed a Modified CNN method to separate, generalize and reindex the descriptions and the solutions, which can be integrated to competence-preserving CBM methods to diversify CBR retrieval results while reducing case base size and preserving competence at the same time.

In conclusion, we have proved the Modified CNN method to be effective in diversifying the retrievals. The diversity of retrievals can be controlled by the solution generalization threshold θ_{sol}. It is also feasible to reduce the case base size and to preserve case base competence by adjusting the description generalization threshold θ_{des}.

REFERENCES

[1] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches", *AI Communications*, Amsterdam: IOS Press, vol. 7, no. 1, pp. 39–59, 1994.

[2] B. Keith and B. Smyth. "Improving Recommendation Diversity", *Proceedings of the 12th National Conference in Artificial Intelligence and Cognitive Science*, California: AAAI Press, pp. 75-84, 2001.

[3] M.K. Haouchine et al., "Case Base Maintenance Approach", *International Conference on Industrial Engineering and Systems Management*, 2007.