13th Global Congress on Manufacturing and Management, GCMM 2016

Research on matching method for case retrieval process in CBR based on FCM

Zhao Yamin[a]*, Zhang Mengmeng[a], Guo Xiaomin[a], Zhou Zhiwei[a], Zhang Jianhua[a]

[a]Management Engineering Department, ZhengzhouUniversity, Zhengzhou, China

Abstract

Era of knowledge economy, how to effectively mining, the use of knowledge is the enterprise growing concern. CBR system from the field of artificial intelligence is a self-learning system to manage tacit knowledge (case). Case retrieval link is the core link, the advantages and disadvantages of search methods directly affect the efficiency of case retrieval and case matching accuracy. Therefore, this paper proposes a new case matching process: when the size of the case database is small, it searches based on the case similarity algorithm; when the case database is large, it searches based on the FCM secondary retrieval model. And illustrates the fastness and efficiency of FCM in matching large-scale case database.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license
(http://creativecommons.org/licenses/by-nc-nd/4.0/).
Peer-review under responsibility of the organizing committee of the 13th Global Congress on Manufacturing and Management
Keywords: CBR; case matching; similarity algorithm; FCM

1. Introduction

The focus of knowledge management is on the management of tacit knowledge. Externality of tacit knowledge is the key to the success of the organization to create and utilize new knowledge effectively, and case is an important way for the explicitization of tacit knowledge. CBR (Case Based Reasoning) system is widely used in all kinds of knowledge-based systems because of its advantages such as easy explicit tacit knowledge explicit, high knowledge utilization rate and good self-learning ability. In general, the CBR process includes: case representation and organization, case retrieval, case adaptation and revision, case study and management. In the CBR process, case

---

* Corresponding author. Tel.: 15690875485.
  E-mail address: 1217200588@qq.com

retrieval is the most core activities, the results of the search results will directly affect the efficiency of the entire CBR system operation and accuracy. The quality of the retrieved cases depends on the accuracy of the case similarity calculation. A good retrieval algorithm can quickly achieve the retrieval, and the retrieved cases are as similar as possible and the number is as small as possible. Based on these, the focus of this study is how to effectively apply FCM to CBR, so as to solve the case matching problem in the case retrieval process.

2. Literature review

At present, scholars have made a lot of research on case retrieval in CBR, and have also obtained a lot of valuable theoretical and case retrieval methods. Including four aspects: First, the case for attribute reduction, thereby enhancing the speed of retrieval. Zhu Haodong, Zhong Yong (2010) [1] proposed to attribute reduction and class correlation algorithm combined, and then retrieve the case. The algorithm not only simplifies the retrieval process, but also improves the accuracy of the retrieval. However, it needs to discretize the continuous attributes, resulting in information distortion.Second,by optimizing the representation structure of the case, thereby reducing the difficulty of retrieval. Li Linlin, Sun Jiyin et al. (2007) [2] proposed a case retrieval algorithm based on decision tree knowledge representation: transforming expert experience into tree structure and then searching. This method is easy to retrieve, but when the case database changes need to re-create and storage, resulting in greater cost. (2012) [3] analyzed and designed the object-oriented case-based knowledge structure of the case knowledge, and on this basis, the design of KM case knowledge representation subsystem model, and explain the working mechanism of the model, Full support for case knowledge representation. Then,by improving the degree of recognition algorithm, to achieve the precise case retrieval. (2012) [4] proposed an improved Nearest Neighbor method to calculate the similarity of cases, considering the local similarity of attribute values and the local similarity of attribute weights, but the algorithm is not applicable to all attributes Value. Zhang Jianhua (2014) [5] on this basis, the attribute values were divided, from 7 different situations on the calculation of the degree of understanding were described, and verified by an example. This also helps to compute the degree of recognition in this paper.Finally,first classification algorithm with the case, thereby improving the speed of retrieval. Xiao Feng, Xin Daxin (2002) [6] proposed a neural network-based method for retrieving multimedia database content to adaptively classify cases and then identify and match them. This method can significantly improve the retrieval speed, but when the case attribute is more time-consuming retrieval is longer.

Although the existing literature on the common case matching algorithm to make improvements, but will increase as the case or case attributes, and increase the cost and time. At this time, it is necessary to classify the cases in the case base by using clustering algorithm before case matching. In the real world, people like to be divided into different categories of objects, such as the biological is divided into sectors, doors, classes, heads, families, genera, species. This division process is the process of clustering. Up to now, researchers have put forward a lot of clustering methods [7-11], the specific content as shown in Table 1.

Table 1. Cluster analysis of the advantages and disadvantages

Name	Advantages	Disadvantagrs
Division	Effective and simple	The number of clusters is difficult to determine,and it is difficult to obtain the global optimum.
Level	Effective simple and Easy to understand	Irreversible.
Density-based approach	The resulting clusters are of arbitrary shape.	The threshold value is not easy to determine.
Grid-based approach	High speed	The underlying granularity of knowledge is not easy to grasp.
Model-based approach	Simple, easy to operate	As the type of case attribute increases, the model needs to be continually transformed.