Engineering Applications of Artificial Intelligence 45 (2015) 46–56

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Integrating granular computing and bioinformatics technology for typical process routes elicitation: A process knowledge acquisition approach

Danchen Zhou*, Xuan Dai

Institute of Machinery Manufacturing Technology, China Academy of Engineering Physics, Mianyang 621900, China

ARTICLE INFO

Article history:
Received 24 February 2015
Received in revised form
6 April 2015
Accepted 16 June 2015
Available online 3 July 2015

Keywords:
Typical process route
Computer-aided process planning
Data mining
Granular computing
Sequence alignment
Neighborhood-based granulation

ABSTRACT

Computer-Aided Process Planning (CAPP) plays a significant role in modern manufacturing system, and knowledge-based CAPP system is one of the predominant trends of its development. How to discover and acquire valuable process knowledge from the existing process data by applying data mining methodology is always the key technology and bottleneck issue for improving the application level of knowledge-based CAPP systems. As an important knowledge of process flow, typical process routes have a vital influence on the accuracy and efficiency of part family oriented process planning. In this paper, a novel approach for elicitation of typical process routes through the combination of granular computing theory and bioinformatics technology is put forward. The sequence alignment technology in bioinformatics is used to establish the best alignment between two process routes, based on which their distance is exactly calculated. According to the distances between process routes to be analysed, the neighborhood-based granulation method in granular computing theory is applied to construct a series of process information granular layers with different granularity so as to acquire typical process routes from process information granules contained in an optimal granular layer. Two application examples not only adequately validate the applicability and effectiveness of the proposed approach, but also fully demonstrate its advantages in the quality and efficiency of typical process routes elicitation.

1. Introduction

As the bond and bridge between Computer-Aided Design (CAD) and Computer-Aided Manufacturing (CAM), Computer-Aided Process Planning (CAPP) is a kind of methodology that, through making full use of information technologies, can assist process technicians to plan the rational and feasible manufacturing procedures of a component from a blank to a finished product so as to transform the designing data of a product into the corresponding manufacturing data (Yusof and Latif, 2014; Xu et al., 2011). CAPP can be considered as a human-centered system composed of a person and a computer, in which both the geometry information (e.g., shapes and sizes) and the production information (e.g., material and batch) of a machined part are fed into a computer at first; afterwards, the matching and optimization between information of product design and information of all possible machining ability provided by the manufacturing environment is carried out by taking full advantage of the capability possessed by a computer such as numerical calculation, logic reasoning, and judgment; finally, a succession of feature machining methods, process route, machine tools, cutting tools, cutting parameters, and fixtures needed for the transition of a component from the original state to the last specified state are outputted by the computer automatically or by the human–computer interaction.

Process planning is an extremely complex thinking activity with high comprehensiveness and creativeness, and its efficiency and quality depends on the experiences and knowledge owned by the process technicians (Chen et al., 2011; Denkena et al., 2007; Helgoson and Kalhori, 2010; Grabowik et al., 2012). Consequently, the whole procedure of process planning cannot do without effective support of process knowledge. Knowledge-based CAPP system is exactly a kind of CAPP system by which the organic combination between computer intelligence and human decision-making ability via an appropriate human–computer interaction mode in the course of process planning, and more important, the reliable utilization and accumulation of process knowledge, can be realized on the basis of the acquisition, expression, arrangement, transmission, and sharing of process knowledge (Vosniakos and Giannakakis, 2013; Zheng et al., 2008; Tor et al., 2005; Shi et al., 2002; Law et al., 2001). In a word, its goal is to transfer the most

* Corresponding author. Tel.: +86 8162485612.
E-mail address: zdc69@163.com (D. Zhou).

http://dx.doi.org/10.1016/j.engappai.2015.06.014
0952-1976/

appropriate process knowledge to the right people at the right time so as to help them make rapid and accurate decisions on process planning, and to achieve the moderate automation and intellectualization of process planning.

How to obtain useful process knowledge promptly and accurately is a key problem for building a knowledge-based CAPP system. The existing process knowledge originates principally from books, manuals, related data, process documents, and experts. However, since these knowledge is deeply embedded in the minds of individual experts, hidden in the books and documents, but also implied in the specific manufacturing environments of the enterprises, it is quite difficult to find a suitable knowledge extraction method. In summary, because of the extensiveness, implicitness, diversity, and dynamicness of process knowledge, their automatic acquisition and accumulation has become one of the bottleneck problems in the development of knowledge-based CAPP system.

As a brand-new research direction in artificial intelligence field since 1990s, Data Mining (DM), also known as Knowledge Discovery in Database (KDD), has been widely used in manufacturing industry, such as engineering design, manufacturing systems, decision support systems, shop floor control and scheduling, fault detection and quality improvement, maintenance, and customer relationship management (Reyes et al., 2015; Wang, 2007; Harding and Shahbaz, 2006; Köksal et al., 2011; Ismail et al., 2009; Shahzad and Mebarki, 2012). Moreover, with the rapid development and great improvement of DM technology in mass data processing, mining efficiency, knowledge processing capability, and robustness of algorithms, the application of DM technology to the acquisition of process knowledge becomes possible. For example, Park (2003) suggested a knowledge capturing methodology for knowledge-based CAPP system, by which four knowledge elements for process planning including facts, constraints, way of thinking, and rules were derived. Denkena et al. (2014) presented an approach to the collection and systematic evaluation of manufacturing data by using DM technology for knowledge-based process planning related to the concept of gentelligent production. Liu et al. (2010) expounded the key technologies of process knowledge discovery in the process database, including process data treatment technology, process data mining technology, and process pattern evaluation technology. Jia et al. (2008) proposed the concept of Process Planning Knowledge Discovery (PPKD) for product process planning database in CAPP system. The technological architecture of PPKD was built based on object-oriented model driven technology, and its script was designed. Zhu (2008) put forward a method that integrated fuzzy clustering and variable precision rough set for acquisition of the unknown process planning rules in CAPP systems.

Usually, process knowledge can be divided into three types based on their forms, i.e. knowledge of process flow, knowledge of resource, and knowledge of calculation (Chen et al., 2011). As an important kind of knowledge of process flow, the typical process routes are generally applicable process sequences that have been summarized by a manufacturing enterprise through the long-term practice, and can be more frequently used for either broad classifications (e.g., shaft, box, gear, and sleeve) or narrow classifications (e.g., the shaft can be subdivided into plain shaft, stepped shaft, crankshaft, camshaft, etc.) of part family. Based on a typical process route of a part family, only minor modifications are needed for process routes planning of the new parts that belong to the corresponding part family so that the efficiency and accuracy of CAPP systems can be greatly improved. Hence, research on application of DM methodology in the intelligent elicitation of typical process routes has an important significance for the technological development of knowledge-based CAPP system.

The combination between the similarity calculation of process routes based on operation encoding and the clustering of the similar process routes previously is the predominant method for the acquisition of typical process routes. Liu et al. (2007) employed the Euclidean distance to calculate the similarity between two process routes, and applied the agglomerative hierarchical clustering method to form the clusters of similar process routes. Zhang et al. (2013) proposed a multi-level comprehensive method for the similarity calculation between two process routes, and used the Particle Swarm Optimization (PSO) to realize the intelligent clustering of analogous process routes. The aforementioned two studies provide the better solutions to the similarity analysis between process routes as well as to the automatic acquisition of typical process routes. Nevertheless, the available methods still suffer from two disadvantages:

(1) The operation encoding method is hard to comprehend because the straightforward operation descriptions in a process route are transformed into the abstract numbers.
(2) The number of clusters is determined by the experiences rather than the objective standards, which has certainly a negative impact on the quality of the extracted typical process routes.

Aiming at these two shortcomings, a novel approach for acquiring typical process routes is proposed in this paper. In the approach, the pairwise alignment method in bioinformatics is firstly used to compare two process routes, and in accordance with the best alignment, the distance (similarity) between two process routes is more precisely calculated; furthermore, the neighborhood-based granulation method in Granular Computing (GrC) theory is applied to build a series of process information granular layer, and by employing the multiple alignment method, the more high-quality and more diverse typical process routes can be extracted from the process information granules in an optimal granular layer determined based on an explicit and objective criterion.

2. Preliminaries

2.1. GrC and neighborhood-based granulation

GrC, in essence, is an emerging paradigm of information processing that brings together the existing formalisms of set theory (interval analysis), fuzzy sets, rough sets, shadowed sets, and quotient space theory under the same roof by clearly visualizing some essential commonalities and synergies such as information granules, information granulation, information granular layers, and information granularity (Bargiela and Pedrycz, 2003; Pedrycz, 2013).

Information granules, a fundamental concept in GrC, are collections of entities that are arranged together due to their similarity, functional adjacency, indistinguishability, coherency or alike. Information granules are examples of abstractions. As such they naturally give rise to a hierarchical structure: depending on the complexity of a problem or particular needs to be addressed, the same problem or system can be perceived and analysed at different levels of detail, which becomes an essential aspect facilitating a way of hierarchical processing of information with different levels of hierarchy (that is information granular layers) indexed by the size (that is information granularity) of information granules.

Information granulation, a key methodological and algorithmic issue in GrC, refers to the construction procedure of information granules with different granularity. It happens through a procedure of data organization and data comprehension that includes the selection of the formal framework of information granulation