Available online at www.sciencedirect.com
ScienceDirect
Robotics and Computer-Integrated Manufacturing 25 (2009) 32–41
Robotics and Computer-Integrated Manufacturing
www.elsevier.com/locate/rcim

A fuzzy-knowledge resource-allocation model of the semiconductor final test industry

Kung-Jeng Wang[a,*], Y.-S. Lin[b], Chen-Fu Chien[b], J.C. Chen[c]

[a]*Department of Industrial Management, National Taiwan University of Science and Technology, ♯43, Sec.4, Keelung Road, Taipei 106, Taiwan, ROC*
[b]*Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 300, Taiwan, ROC*
[c]*Department of Industrial Engineering, Chung Yuan Christian University, Chung-Li 320, Taiwan, ROC*

Received 23 August 2006; received in revised form 25 February 2007; accepted 22 May 2007

Abstract

The operations of the semiconductor final test industry are complicated and characterized by multiple-resource constraints that require simultaneous considerations. One of the most challenging production-planning decisions in the industry concerns an efficient allocation of resources that results in high manufacturing performance. Firms in the industry are thus eager to discover resource-allocation knowledge from large manufacturing databases. This study develops a novel model via the extraction of fuzzy-business rules from databases for obtaining resource-allocation knowledge as well as allocating resources efficiently. The proposed model uses both a genetic algorithm to find the best priority sequence of customer orders for resource allocation and, in accordance with the priority sequence of orders, a fuzzy-inference model to allocate the resources and to determine the order-completion times. Experiments showed that the proposed model can significantly reduce task tardiness.

Keywords: Capacity management; Fuzzy inference; Genetic algorithm; Semiconductor final test; Resource-allocation decisions

1. Introduction

Resource allocation is crucial to high-tech manufacturing industries. This problem is sophisticated owing to task-resource relations and tight tardiness requirements. The revenue model of the semiconductor final test industry differs from a general manufacturing industry insofar as the former gains profits by selling the capacity of both testing machines and services to its customers. Within the industry's overall revenue-oriented process, the wafers (chips) from semiconductor manufacturing fabs are raw materials, most of which are urgent orders that customers make and compete with each others for limited resources. This scenario creates a complex resource-allocation problem.

In the semiconductor final test industry, a chip test requires both a functional test and a package test. Testers are the most important resource in performing chip-testing operations. Handlers, programs, loadboards, and toolings are auxiliary resources that facilitate testers' completion of a testing task. All the auxiliary resources are connected to testers so that they can conduct a final chip test. Handlers upload and download IC chips from testers and do so with an index device and at a predefined temperature. Loadboards feature interfaces and testing programs that facilitate the diagnosis of IC chips' required functions. Customers place orders for their product families that require specific quantities, tester types, and testing-temperature settings. These simultaneous resources (i.e., testers, handlers, and loadboards) conflict with the capacity planning and the allocation of the chip final test because products may create incompatible relations between testers and handlers.

Conventional resource-allocation model is not fully applicable to the resolution of such sophisticated capacity-allocation problems. Nevertheless, these problems continue to plague the semiconductor final test industry. Thus, one should take advantages of business databases,

*Corresponding author. Tel.: +886 2 2737 6769; fax: +886 2 2737 6344.
E-mail address: kjwang@mail.ntust.edu.tw (K.-J. Wang).

0736-5845/$ - see front matter
doi:10.1016/j.rcim.2007.05.004

which consist of huge potentially useful data and attributes implying certain business rules and know-how regarding resource allocation. Traditionally, people have used statistics techniques to carry out the classification of such information to induce useful knowledge [1]. However, some implicit interrelationships of the information are hard to discover owing to noises coupled with the information as well as the fuzzy property of some information.

In contrast, artificial intelligence (AI) and data mining techniques enable an enterprise to explore its huge business databases and find possible clues for improvement. Because profits gained from such an investment are promising, many modern manufacturing enterprises begin to deal with knowledge engineering to develop an intelligent system for strengthening their productivity and competitiveness. AI techniques including genetic algorithm (GA) and fuzzy inference are promising approaches of resource allocation by which enterprises might (1) efficiently form proper working sequences of orders (2) efficiently allocate resources to the orders to solve the addressed problem.

The purpose of this study is thus to extract, in the context of the semiconductor final test industry, valuable business rules by using GA and fuzzy-inference techniques that facilitates capacity-allocation decision making. The study focuses on the following issues.

1. How should an enterprise, given a set of manufacturing attributes, use fuzzy-membership functions to describe the attributes that possess uncertain properties?
2. How should an enterprise induce resource-allocation rules from a large production database?
3. How should an enterprise use those rules to create optimal resource allocation such that the tardiness of orders can be minimized?

This study developed a fuzzy-rule knowledge-based resource-allocation model that features a GA for optimization. We applied a fuzzy c-means (FCM) algorithm to the model in order to identify the production-system statuses by using fuzzy-members' parameters. We constructed a fuzzy-inference neural-network system (FINNS) to facilitate efficient learning and induction from a large number of production-history examples and to determine the timing of releasing orders. We compared the performances of FINNS with the performances of a decision-based inference system. Using FINNS, we then developed a GA to form a resource-allocation model that, denoted as GA-FINNS, serves the simultaneous capacity allocation of multiple resources. Here, we examine the performance of GA-FINNS in comparison with the performance of the decision tree-based resource-allocation model using GA, denoted as GA-DTR. The proposed GA has successfully selected the best sequence of orders in terms of minimal total tardiness, and the fuzzy-inference module also conveyed the resource allocation of each order.

The rest of this article is organized as follows. Section 2 is the literature review. In Section 3, we detail the building of the inference model for resource allocation. Using an FCM clustering method, we define the parameters of input and output fuzzy attributes, and we develop and evaluate FINNS. In Section 4, we detail the development of GA-FINNS and evaluate the performance of GA-FINNS. Section 5 is a summary of this study.

2. Literature review

Enterprises that try to make resource-allocation decisions sometimes encounter problematic manufacturing knowledge. In recent decades, AI techniques, such as GA and fuzzy-inference models, have been emerging as solutions to such problems.

FCM is a method of clustering that allows one datum to belong to two or more clusters. In this study, we use this method to cluster data. The details of the FCM algorithms are in Bezdek [2]. In addition, fuzzy inference is the process whereby one maps inputs to outputs by using fuzzy logic. In general, a fuzzy-logic inference system can successfully perform predictions under an uncertain environment. For those attributes that take the representational form of numerals but that actually have non-number meanings, a fuzzy system may capture the internal implications more precisely than a deterministic representation. Usually, enterprises use two types of fuzzy-inference models: the Mamdani type and the Takagi–Sugeno type. Mamdani-type inference [3] is, in general, not computationally efficient because it requires one to integrate across a continuously varying function in order to find the centroid. Sugeno and Takagi [4] suggest the use of a single spike—a singleton—as the membership function of the rule consequent. A fuzzy singleton is a fuzzy set with a membership function that is unity at a single point in the universe of discourse and that is zero everywhere else.

In the Mamdani-type inference system, in order to conjunct the rules, we apply the AND fuzzy-operation intersection: $\mu A(x) \cap \mu B(x) = \min[\mu A(x), \quad \mu B(x)]$. It needs to convert a fuzzy set (produced in the composition stage) into a crisp value. The only difference between the Takagi–Sugeno type and the Mamdani type is in the rule consequent. Instead of a fuzzy set, the Takagi–Sugeno type features a mathematical function for the input variable: IF x is A and y is B then z is $f(x,y)$, where x, y, and z are linguistic variables; A and B are fuzzy sets in the universe of discourses X and Y, respectively; and $f(x, y)$ is a mathematical function.

While the applications of conventional GA to individual capacity planning and to task allocation can be found in the literature of, for example, [5], a small number of researchers have reported on their successful use of fuzzy inference both in academics and industry. For example, Cordón, Herrera and Villar [6] proposed a new approach wherein one can use a GA to learn from a database and to build a knowledge base. Wang, Fang and Nuttle [7] developed a due-date bargainer, a useful tool that supports negotiation on due dates between a manufacturer and its