Expert Systems with Applications 39 (2012) 8736–8743

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Reinforcement learning approach to goal-regulation in a self-evolutionary manufacturing system

Moonsoo Shin [a], Kwangyeol Ryu [b], Mooyoung Jung [c,*]

[a] Department of Industrial and Management Engineering, Hanbat National University, San 16-1, Duckmyoung-dong, Yuseong-gu, Daejeon 305-719, Republic of Korea
[b] Department of Industrial Engineering, Pusan National University, San 30, Jangjeon-dong, Geumjeong-gu, Busan 690-735, Republic of Korea
[c] School of Technology Management, Ulsan National Institute of Science and Technology (UNIST), Banyeon-ri 100, Ulsan 689-798, Republic of Korea

ARTICLE INFO

Keywords:
Self-evolutionary manufacturing system
Fractal organization
Goal-regulation
Reinforcement learning
Agent
Production planning

ABSTRACT

Up-to-date market dynamics has been forcing manufacturing systems to adapt quickly and continuously to the ever-changing environment. Self-evolution of manufacturing systems means a continuous process of adapting to the environment on the basis of autonomous goal-formation and goal-oriented dynamic organization. This paper proposes a goal-regulation mechanism that applies a reinforcement learning approach, which is a principal working mechanism for autonomous goal-formation. Individual goals are regulated by a neural network-based fuzzy inference system, namely, a goal-regulation network (GRN) updated by a reinforcement signal from another neural network called goal-evaluation network (GEN). The GEN approximates the compatibility of goals with current environmental situation. In this paper, a production planning problem is also examined by a simulation study in order to validate the proposed goal regulation mechanism.

1. Introduction

Up-to-date market dynamics has been strongly demanding flexibility and responsiveness from the manufacturing systems. Actually, the manufacturing enterprises strive to facilitate continuous and quick adaptation to the constantly varying customer requirements and the competitive environment. The conventional manufacturing systems, however, are not suited for the ever-changing environment because of their rigid organizational structure and preoccupied static goals (Frayret, D'Amours, & Montreuil, 2004; Heragu, Graves, Kim, & Onge, 2002; Renna & Ambrico, 2011). In order to ensure their competitive power, the manufacturing systems should have an advanced capability to dynamically organize their production resources and autonomously formulate their goals. Furthermore, a manufacturing system is required to have a capability of facilitating self-evolution according to its environmental situation (Shin, Mun, & Jung, 2009a).

In keeping with this line of thought, various multi-agent systems (MAS) have been proposed in the literature, including MetaMorph (Maturana, Shen, & Norrie, 1999), MetaMorph II (Shen, Maturana, & Norrie, 2000), PROSA (Van Brussel, Wyns, Valckenaers, Bongaerts, & Peeters, 1998), ADACOR (Leitão & Restivo, 2006), and FrMS (Ryu, Yücesan, & Jung, 2006), and r-FrMS (Shin, Mun, Lee, & Jung, 2009b). The agent-based manufacturing systems fit naturally into a decentralized control structure, whereby they have a flexible and reconfigurable organizational structure (Weiss, 1999). Based on such a principle, all these systems continuously adapt their organizational structure to the environment by means of self-organizing mechanisms toward achieving a goal. That is, goal-orientation is a main organizing rule.

In the area of MAS, researchers have been interested in reinforcement learning approaches to the problem of how an agent learns to select proper actions for achieving its goals through interacting with its environment (Wang & Usher, 2007). There have been several examples dealing with dynamic order acceptance (Arredondo & Martinez, 2010), production control (Csáji, Monostori, & Kádár, 2006), production scheduling (Wang & Usher, 2004, 2007; Zhang, Zheng, & Weng, 2007), and agent architecture (Tan, Ong, & Tapanuj, 2011). All these examples have shown successful approaches to goal-orientation, assuming well-defined goals. Every production resource represented as a resource agent, however, is required to regulate its own goal, not only to adapt to the changing environment but also to conform to its cooperators and competitors. An action oriented toward a fixed goal inappropriate to changed environmental situation results in an adverse effect on overall performance. Despite the various successful approaches to goal-orientation, the regulation mechanism of a predefined goal has not been fully explored.

In a manufacturing system, an aiming level with respect to various criteria (e.g. returned profit, utilization of resources, and processing lead time) can be considered as a goal. For example, if the

* Corresponding author. Tel.: +82 52 217 1011; fax: +82 52 217 1058.
E-mail address: myjung@unist.ac.kr (M. Jung).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2012.01.207

aiming level of returned profit is too high, the resource agent tries to undertake high profit tasks only whereas it ignores relatively low profit tasks. Thus, the agent tends to miss the opportunity to undertake many tasks, since the high profit tasks might be assigned to other agents which have much competitiveness (e.g. low cost, short processing time, and high quality). Consequently, not only utilization rate but also returned profit becomes low. In this case, the agent should lower the aiming level of returned profit. In other words, the resource agent should regulate its own goal so as to conform to its environmental status (involving competitive or collaborative features of the entire environment and the interrelated features with competitors).

This paper proposes an autonomous goal-regulation mechanism that adopts a reinforcement learning approach, aiming at implementation of self-evolutionary manufacturing system. The proposed regulation mechanism facilitates adaptation of a predefined goal to the changing environment. Individual goals are dynamically changed by a neural network-based fuzzy inference system, and the neural network is updated by a reinforcement signal from the environment. The remainder of this paper is organized as follows. Section 2 presents previous researches on self-evolutionary manufacturing systems, and discusses a reinforcement learning approach based on actor-critic learning. Section 3 is devoted to details of the proposed regulation mechanism. In Section 4, a case study on production planning is presented. Finally, the conclusions are drawn in Section 5.

2. Related works

2.1. Self-evolutionary manufacturing system

Shin et al. (2009a) have proposed a self-evolution framework for manufacturing systems. Here, the self-evolution means a continuous process of adaptation to the environment, by which not only the entire system but also its production resources autonomously regulate their own goal, and also the organizational structure of the production resources is dynamically changed. The self-evolution framework consists of three components (see Fig. 1): (1) fractal organization-based control architecture, (2) goal-oriented dynamic organizing mechanism, and (3) autonomous goal-formation mechanism. The term 'fractal' denotes a self-similar shape recursively constructed, implying '*a similar pattern inside of another similar pattern*' (Mandelbrot, 1982). Fractal organization is a fractal-like structured association of distributed entities in which a self-similar organizing pattern is recursively deployed into the whole system (Shin et al., 2009a).

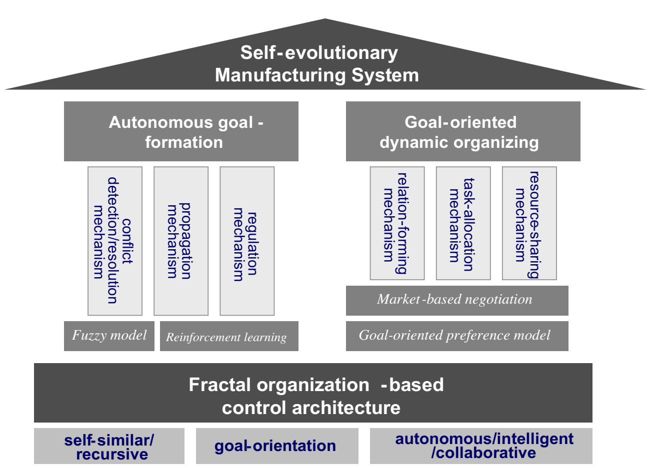

Fig. 1. Embodiment framework of a self-evolutionary manufacturing system. Shin et al. (2009a).

Shin et al. (2009b) have also proposed a relation-driven fractal organization for distributed manufacturing systems, namely, r-FrMS, along with its organizing mechanism. In the r-FrMS, production resources are represented as individual agents, namely, autonomous and intelligent resource units (AIR-units). Every AIR-unit is a goal-oriented decision-maker, and it collaborates and competes with other AIR-units in order to establish its own goal. Especially, the r-FrMS is organized through a market-based negotiation between the AIR-units.

Ryu and Jung (2004) have investigated the goal-orientation features of a fractal organization and proposed a goal-generation mechanism that operates in a goal-formation framework. The goal-generation mechanism mainly propagates a predefined goal into several sub-goals, focusing on the interrelation between a decision entity and its subordinate decision entities. Furthermore, Shin, Cha, Ryu, and Jung (2006) have presented a conflict detection and resolution mechanism, which is one of principal working mechanisms for autonomous goal-formation. In order to clarify the vague mutual interactions among goals, the conflict detection and resolution mechanism adopts an indirect evaluation scheme: (1) transformation of individual goals into sets of tasks, and then (2) a task-based simulation.

However, despite the visionary prospect of these research contributions, there are some limitations. The goal-generation mechanism proposed by Ryu and Jung (2004) is limited in its consideration for dynamic changes of individual goals. Moreover, the conflict detection and resolution mechanism proposed by Shin et al. (2006) has no consideration for conflict-free situations. Thus, a goal-regulation mechanism is required to adapt individual goals to the changing environment and also to continuously improve the goals and the resultant performance. While the goal-regulation mechanism is another principal working mechanism for autonomous goal-formation, it is still an open issue. In this paper, a reinforcement learning approach to autonomous goal regulation is proposed, based on the self-evolution framework addressed by Shin et al. (2009a) and its organizational model investigated by Shin et al. (2009b).

2.2. Actor-critic learning

The adaptive heuristic critic (AHC; Barto, Sutton, & Anderson, 1983; Sutton & Barto, 1998) is the best known actor-critic learning method, which was introduced to solve a pole-balancing control problem and implemented in such parameterized functional forms as neural network-based function approximators. The AHC model basically consists of two neural networks: the adaptive critic element (ACE) and the associative search element (ASE). The ACE plays the role of a critic, and it approximates an evaluation function in an adaptive way from the primary reinforcements given directly by the environment through rewards and punishments. The evaluation function represents an internal reinforcement, mapping states to expected values. The internal reinforcement is more informative than the primary reinforcement. On the other hand, the ASE plays the role of an actor, implementing and adjusting decision policies, and mapping states to actions (Mizutani, 1997). The ASE learns to select actions that lead to better critic values by using the internal reinforcement signal.

Jouffe (1998) has introduced the fuzzy actor-critic learning (FACL) which is dedicated to tune online the conclusion part of fuzzy inference systems (FIS). The FACL is an extension of the AHC, in which fuzzy logic is used as a function approximator. According to the Jouffe (1998), FACL allows the state representation to be more generalized than the original AHC. In other words, FACL is allowed to deal with discrete as well as continuous actions, whereas AHC is limited to discrete input space.