Decision Making in Agent-Based Manufacturing With A Reinforcement Learning Approach

SARA KHODAYARI, Graduate Student of Artificial Intelligence
CARO LUCAS, Professor,
Control and Intelligent Processing Center of Excellence (CIPCE)
Department of Electrical and Computer Engineering
University of Tehran,
North Kargar Ave., P.O. Box: 14395/515
IRAN

Abstract: - Development of production systems, presents new ways of manufacturing. Agent-based systems are of emerging interest in the specification and implementation of manufacturing systems. In this paper, we have simulated a manufacturing system to introduce the role of intelligent decision making in manufacturing. The main idea is to develop independent agents which would create a formation for producing a desired product .Agents do not have enough information to know the highest quality that a combination of manufacturers can achieve in production process. By applying reinforcement learning, the agents "learn" which combination of manufacturers is best for a specific task. The simulation is a basic one with focus on agent's decision making; therefore different manufacturing phenomena such as scheduling are not its concern.

Key-Words: - Agent-Based Manufacturing, Agent Decision Making, Reinforcement Learning, Manufacturing Simulation

1 Introduction

Manufacturing is taking its steps towards a modern decentralized approach instead of an integrated company which has the control over all the production and distribution aspects. Manufacturing industries are facing serious challenges in maintaining their existing markets and improving their capabilities in order to respond to rapidly changing customers' needs. Therefore, the change from conventional managements to a new one is somewhat inevitable [1, 3, 6]. There are many factors that will affect how successful a company will achieve such as the quality of product design, cost of product development, time to market, project management, marketing strategy, customers' preference and competitions etc. It is important to recognize that all these activities involve decision making and all other factors are beyond the company's control except for the decisions it makes [3]. Novel solutions have been applied to face such requirements.

Manufacturing can be modeled as a multi-agent system which is being studied under the area of Agent-Based Manufacturing. The nature of manufacturing and its wide ranges tends to bring new ideas in decentralizing and distributing the components [2]. One successful approach is to apply intelligence role in the domain. The core element of intelligence is agent. Designing the essentials of manufacturing in an agent oriented style, passes the intelligent features of decision making to manufacturing field.

There are certain types of decision makings concerned in manufacturing issues. One which is our main focus and is quite of importance is the combination of different parts of a product which is produced by varieties of manufacturers. A fact in agent based manufacturing is that a product is made of a complex combination of other manufacturer's products. Some products from a manufacturer go better with the product of another manufacturer. Finding the best combination of manufacturers for a specific product requires a lot of time and experience. Beside that, changes in production quality of producers are inevitable. Therefore, if there is any deficiency in the quality of a known manufacturer, it takes some time to realize the change and switch to another manufacturer. This is in conflict with the more increasing demands of the market and if such matter rises, it can bring about irresistible and intolerable consequences. Thus, a reliable method of discovering the best combination of manufacturers and realizing the changes on time is desired.

We have applied reinforcement learning techniques to allow each manufacturer learn the quality produced by different combinations of others and be apt to select that combination for the forthcoming requests of the same type. A simple simulation has been done to evaluate the impact of learning application on making the right decision.

In the next section, we discuss the necessity of agent based manufacturing in general. In section 3, general concepts of our simulation is described and afterwards in section 4 the issue of decision making for selecting the right combination is discussed. Then simulation results are presented in section 5. In the last section we come to a conclusion and future works are propounded.

2 Agent Based Manufacturing

Manufacturing is a vast domain which requires so much organizing and management. Today manufacturing is changing its conventional way to a modern one which the customer plays a vital role in the final product. Traditional manufacturing relies on schedules as means of forecasting what needs to be produced [1]. However in present marketplace a schedule is impractical. The customer defines his demand and the producer should adapt to those demands in order to compete with other manufacturers or else the business is lost. These elements, remove the mass production lines from manufacturing and presents the customer based fabrication to the industry. Such a competitive market would find fast and efficient decision making crucial. Therefore it seems somehow inevitable to enforce intelligence into this area.

On the other hand, agents are the main elements of intelligent decision making. Decision making in manufacturing have different aspects. The main decision we are focusing on, is the formation of agents for producing a customer's need, its efficiency and reliability. Since manufacturing is very complex and the range of scenarios and the possible combinations of parameters are infinite, it's impossible to anticipate and plan for every possible combination of products which the customer may demand. Even if it was possible to consider all the possible combinations, the cost would have been prohibitive [1]. A system which utilizes the history of the previous production and demands and has a judgment of what quality and time would be required for the present need, is more likely to make rational and reliable decision compared with a static system which decision making plays no role.

In general, multi-agent systems may be regarded as a group of intelligent entities called agents, interacting with one another to collectively achieve their goals. By drawing on other agents' knowledge and capabilities, agents can overcome their inherent bounds of intelligence. Independent of the domain problem, agents operate to solve problems by cooperating with a group of agents. Each agent must 1) operate within an organization of agents whether that organization is specified at design time or during run-time, 2) generate plans under that organization structure, 3) allocate tasks to proper agents, 4) integrated agent individual plans and schedules, and 5) execute the plans to solve the problem [5]. As for the industry, the organization is the group of manufacturing and their goal is to manufacture the most proficient product in the market.

In an agent based manufacturing system, agents are the elements to decide what should be produced; with which customer and what type of product can the organization profit more; which formation between the producing agents can perform more efficiently and matching to the customer's needs and many other parameters. All these subtle decisions require processing the knowledge gathered about history of all sections' performance. The goal or intention of agents in manufacturing is to make these decisions as efficiently as possible. An agent-based manufacturing organization is expected to perform with much better improvement and consistency in industrial applications.

3 General Concepts of Agent Based Manufacturing Simulation

In an agent based manufacturing environment, the production process is broken into its associated subtasks (operations) and they are given to different manufacturers. (In this context, "manufacturer" is our main decision- making element and we use "agent" and "manufacturer" interchangeably). Various combinations of manufacturers might be formed for performing each task (performing a task in this context is the same as constructing a product) and different combinations result in different qualities. This formation is desired to be known. When a customer enters an agent based manufacturing system and gives a request to a specific manufacturer, the manufacturer must choose the best combination for its production process. The history of the previous productions and results is a good basis to make decision. The most effortless way of selecting a combination is to choose randomly and allow the capabilities of each agent to be the only decision making criteria. Apparently, this would not come to the optimal solution with the highest possible quality. On the other hand, if each manufacturer would take other parameters into account and besides considering the capabilities of each manufacturer, rely on their previous history of work and how well the task has been accomplished, a better outcome is promising. This can be attained by applying learning techniques to use the hidden knowledge in the past history of other agents.

We have simulated a simple environment of manufacturing which a constant number of agents exist and each agent is capable of performing specific operations. Every once in a while a customer offers a task to an agent. The customer affirms the minimum quality requirement of the final product. The required capabilities of agents and tasks are set in a way that no agent can perform any task all by itself; that is, each agent needs others to accomplish a task (apparently this is a necessary condition for an agent based manufacturing system to force agents cooperate). The agent which has received the offer is to evaluate the required capabilities of the first subtask and the available capabilities in agents and decide to whom it should be given to. The task then is assigned to that agent. After finishing the subtask, the present agent should make a similar decision for the next subtask and this chain goes on until the task is completed.

Various agents perform the subtasks with different qualities. This affects the final product's quality. According to the difference of the declared quality by the customer and the quality obtained at the end, the selected combination is evaluated; if the quality is below the declared quality, the agents involved in the production process are punished.

A static lookup table exists in the system to allow the environment evaluate the final product's quality. This lookup table is considered to be the expert's knowledge of different groupings of manufacturers. The lookup table includes all the knowledge of expert about all possible combinations of agents for producing different products according to their capabilities. Every cell of the lookup table indicates the quality that a combination achieves by performing a specific task, with what probability that quality is obtained and in case the given quality is not satisfied, what the expected quality is in the other case. For example a task which requires the capabilities of #1, #4, #5 and #7 can be done by combination of Agent1, Agent3, Agent6 and Agent2 by probability of 0.7 with the quality of %70 and by probability of 0.3 with the quality of %55. We assume that the qualities obtained by each combination is semi-static; that is the lookup table does not change, therefore the qualities produced is limited to the values of the lookup table.

4 Agent Decision Making for Combination Selection by Reinforcement Learning

The basis of our considered manufacturing system is on the quality of the final product and whether it meets the requirements the customer has stated or not. As described above, this quality parameter is totally dependant on the combination of manufacturers who completed the task. Therefore, the agents try their best to find the most suitable combination which presents the highest quality. The agents apply learning techniques to use the obtained knowledge of the history of others to make this crucial decision as reliable as possible.

As stated before, when the task is offered to an agent, it breaks it into its subtasks and based on the capability required for each subtask and the capability the agents have, it should decide who would be most competent for performing the subtask. The agent who has received the offer begins with the first subtask and decides whom it should be assigned to- the decision making process would be discussed in detail. As the next agent is chosen and the subtask is performed, now the present agent should make analogous decision for the next subtask. Thus, the decision making is completely distributed among the agents and all the assigned agents are totally involved. For example the customer offers a task to Agent1 with subtasks #2, #3 and #5, and the required capability for subtask #2 is Capability2, for subtask #3 is Capability1 and for subtask #5 is Capability4. In a system with 3 agents, a capability matrix can be as follows:

$$Capability = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

The rows of the above matrix are agents and the columns are capabilities. Therefore, all the three agents can compete for Capability1, but only one agent can perform Capability3 and Capability4. In the given example, Agent1 might give subtask#2 to Agent2 or Agent3-based on the decision criteria. Imagine subtask#2 is given to Agent2. Now as the subtask is done, Agent2 should decide about subtask#3. It can be given to any of the agents. Note that these middle decisions, changes the formation entirely and the paths are distinctive. Varieties of formations can be taken into account depending on the order the subtasks are being accomplished.

As for the decision making process, we have applied reinforcement learning technique in general and Q-learning specifically, to allow the agent make the most suitable decision. The required Q-table cells are mapped to the agent who has performed a present operation of a given task which the quality is in a specific range and is to decide to whom the next operation should be given to. For example the Q-value of cell $Q(2,4,5,50,1)$ is the value of the table for Agent2 who has completed operation 4, the next operation is 5, the declared quality is in 50s and the next agent is to be Agent1. Apparently the decision parameter is the last one- the next agent to perform the next operation. The related to the agent who is incapable of performing the operation have been set to minus infinity. Therefore when selecting the maximum Q-value, it would not be chosen incorrectly.

The *present_operation* is the subtask the *present_agent* has in hands. The *next_operation* is the next subtask to be done. The *next_agent* is the agent to be selected for doing the next subtask. In order to make the problem easier and solve it with discrete Q-learning, we discretized the quality parameter which is continuous. The discrete intervals depend on the designer's ideal.

The action selection procedure is ε-greedy where ε is decaying in every iteration. Apparently, selecting the maximum Q-value is done over the next agent's parameter. So the next agent is selected as:

(1)

$$next\_agent = \max(Qtable(present\_agent, present\_operation, next\_operation, task\_quality, :))$$

which the last parameter is all the possible next agents.

This decision making continues and is passed from one agent to another until the task is fully completed. At this time, the task and the combination of agents which has performed it, is returned to the customer by the agent the first offer was given to. Here, the customer has access to the spoken lookup table and can evaluate the quality from the combination. According to the obtained quality, the agents involved in the task are to be rewarded or punished. In case the quality is higher than the customer's expectation, a positive reinforcement is applied and agents involved in the combination, are given an equal reward; if the quality is below the customer's expectation, a negative reinforcement is necessary and they would be punished equally. Based on the ascertained reinforcement the Q-table update is then applied to all the elements involved in the task. The Q update is:

(2)

for i = 1:Operation_no
Qtable(comb(i),i,i+1,task_quality,comb(i+1)) =
(1 - α *)* (Qtable(comb(i),i,i+1,task_quality,comb(i+1)))*
+ α ** (reward +*
γ **(max(Qtable(present_agent,present_op,*
next_op,task_quality,:)))
end

where:

comb=[agent1 for subtask 1,…, agent m for subtask n]

If the combination of agents be considered in vector *comb* and each column of vector indicates the subtask that the agent has performed, the Q update is applied to all the agents in combination for the subtask they had in hand and the decision they made for the next subtask.

It is to be noted that α parameter is decaying α and is saturated at the final iterations.

5 Result

Our main purpose of project definition and the simulation was to apply learning in agent based manufacturing for judging the possible formations and chains of production. In order to evaluate our work, we assume a combination to be the most proficient and desirable combination which results in the best possible quality for a specific product. During the learning process, we expect the agents to learn about each other's competencies and form the best possible formation. Since we consider a specific combination to be a desired one for our evaluation purposes, the ratio of selecting that combination to the total number of decisions made up until now is to grow in every iteration; that is, if we consider a "hit" parameter to be the number of selecting the right combination to the total number of iterations up until now, the "hits" should start to increase as learning is taking place.

(3)

$$hit = \frac{Correct\_Combination\_Selection}{Iteration\_Counter}$$

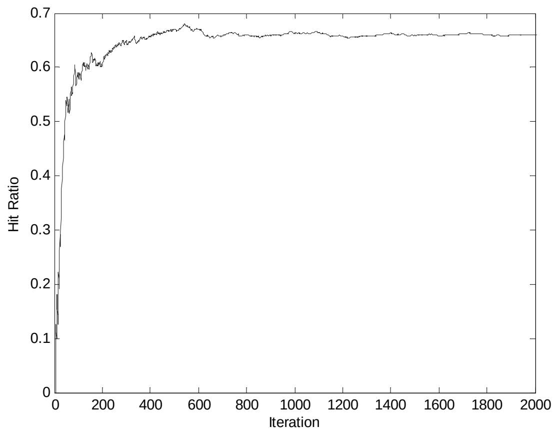

Fig. 1

As it can be seen in figure 1, there is an oscillation at the beginning iterations. But as the time goes by and more tasks are presented to the system, the agents learn more and select the right combination and hit ratio increases. When learning is achieved, the most likely selection is the combination with the highest quality as it was our desire.

We can define this parameter in reverse too with the meaning of decreasing error which is closer to the mind. The ratio of selecting the wrong combinations to the number of iterations up until now can be considered as "error" which is to be minimized.

(4)

$$error = \frac{Incorrect\_Combination\_Selection}{Iteration\_Counter}$$

This meaning is also presented in figure 2 and it has the opposite meaning as described in figure 1.

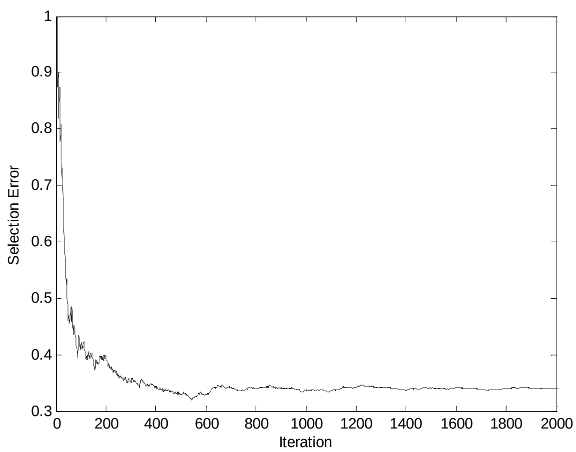

Fig. 2

6 Conclusion and Future Work

It has been claimed that learning methods such as reinforcement learning techniques, are practical tools for decision making in a multi agent application such as agent based manufacturing. It is admitted that simulations presented in this paper are far from being realistic; but still it shows the role of learning and its effect in making a vital decision such as selecting the right combination of manufacturers for a certain type of production. What has not been considered in this paper is the comparison of the results with other intelligent or classical methods. There might be works to be done for decreasing the time of convergence or diminishing the error rate. Other significant issues such as scheduling would realize the simulations much more and would make the results more reliable. On the other hand, there are other important decision makings in manufacturing which this method might be able to improve. Our main purpose was to propose a learning method for finding an optimal formation of manufacturers for producing a product and handling decision makings in manufacturing. Applying similar methods for improving other decision making areas in manufacturing is a work under research and development by authors and the results will hopefully be presented soon.

References:

[1] James Odell, *Agent-Based Manufacturing: A Case Study*, Journal of Object Technology, Vol. 1, no. 5, November-December 2002.

[2] Krzysztof Cetnarowicz, Jarosaw Kozlak, Institute of Computer Science AGH - University of Mining and Metallurgy, *Multi-Agent System for Flexible Manufacturing Systems Managemen*", Proceedings of the Second International Workshop of Central and Eastern Europe on Multi-Agent Systems. Eds. B. Keplicz, E. Nawarecki. Poland 2001. pp 51-60.

[3] Zuozhi Zhao, Arizona State University, *Design Decision Making in Gloable Manufacturing Enterprise in 2020*, http://www.public.asu.edu/~zzhao/Pdf/Zhao_Design_Essay02.pdf.

[4] William R. Lilegdon, *Manufacturing decision making with FACTOR*, Proceedings of the 26th conference on Winter simulation, 1994.

[5] K. S. Barber, T. H. Liu, D. C. Han, Agent-Oriented Design, Technical Report, Laboratory for Intelligent Processes and Systems, University of Texas at Austin, January, 1999.

[6] Leslie D. Interrante and Steven Y. Goldsmith, Sandia National Laboratories, *Emergent Agent-Based Scheduling of Manufacturing Systems*, Sandia National Laboratories Publications, 1998.

[7] Vladimir Marik, Pavel Vrba, Filip Macurek, *Agent-based Solutions for Manufacturing*, Rockwell Automation Research Center Prague, Czech Republic, September 25-26, 2003.

[8] Bing Qiao and Jianying Zhu, *Agent-Based Intelligent Manufacturing System for the 21st Century*, www.shaping-the future.de/pdf_www/152_paper.pdf.