Research on machine learning Algorithm optimization based on 0-1 coding

Yuqi Qin*
School of Tianjin University of Technology, Tianjin, China
*Corresponding author e-mail: qinyuqi@tjut.edu.cn

Abstract. Machine learning algorithm is the core of artificial intelligence, is the fundamental way to make computer intelligent, its application in all fields of artificial intelligence. Aiming at the problems of the existing algorithms in the discrete manufacturing industry, this paper proposes a new 0-1 coding method to optimize the learning algorithm, and finally proposes a learning algorithm of "IG type learning only from the best".

Key words: Machine learning algorithm; 0-1 encoding; Genetic algorithm (ga).

1. Introduction
In recent years, with the transfer of manufacturing to Asia, China has gradually become a manufacturing power, but because of the technology is not absolutely leading and the imperfect management and other reasons, China has not become a manufacturing power. The most convenient way to improve the manufacturing industry is to improve the equipment used in production, formulate production plans scientifically, optimize production scheduling schemes and effectively use limited resources. So as to shorten the manufacturing cycle and reduce the cost of production.

Nowadays, order oriented, multiple varieties, small and medium batch, high quality, short delivery time and low cost have gradually become the obvious production characteristics of the manufacturing industry, which leads to the production mode of modern enterprises tend to be discrete production. Production control is also the core module of ERP system. Nowadays, most of the production control in China's manufacturing industry still depends on manual work and experience. Due to the increasing number of production orders, it is very complicated to carry out dynamic capability assessment after the constraints of related resources needed for production, such as manpower, equipment, raw materials, site, fixture and so on, so it is more difficult to carry out manual balanced allocation of production tasks. Therefore, the overall balance in manual management is only a rough estimate and does not actually play a balancing role.

Although there are many studies and reports on how enterprises allocate production tasks, human resource allocation and production plan execution [1-2], there are few studies and reports on rapid and balanced allocation of production tasks based on actual production capacity. This paper discusses the dynamic statistics of the total workload of all work centers, and then matches the capacity of the work center with the production task, generates a detailed production plan to adapt to the load after the balance, and selects the outsourcing processing mode for the production task exceeding the capacity load. This method will promote the rationalization and precision of manufacturing production plans, so as to
maximize the efficiency of production resources and increase the economic benefits of manufacturing. At the same time, it can also complete the production task on time without delay.

2. Discrete manufacturing production capacity equilibrium model

2.1. Analysis of production plan distribution mode

Production planning and scheduling are the key to production control. Generally, the production management department of medium-sized manufacturing enterprises often receives tens of thousands of production orders, which are transferred to outsourcing production or self-made production only by experience, as shown in Figure 1.

![Figure 1. The current mode of production organization in manufacture enterprise.](image)

2.2. Improvement of production plan distribution mode

In this paper, optimization mode [3] is adopted, that is, production tasks are directly delivered to the workshop (work center) after being selected by the system. Meanwhile, the system takes the delivery order information as the information source and balances the capacity of the workshop through the preparation of the workshop plan. As shown in Figure 2, the most critical link is to make judgment after scientific calculation, and then distribute part of the orders to outsourcing for processing. The mode adopts two-step optimization method: first, after the production department receives the order, it is divided into three types: parts that must be outsourced, parts that must be self-made and parts that can be outsourced. Then, the algorithm is used to match and select the parts that must be made by themselves and the parts that can be made by themselves to the capacity of each work center, and the orders that basically meet the capacity are left to be completed by themselves. And the orders that exceed capacity and the parts that must be outsourced are outsourced for processing, and the workpiece that reaches each workshop or work center is optimized for production scheduling.

The action of "software system optimization scheduling" can be performed by the production management department or the manufacturing executive department, but the scheduling must be optimized according to the resources such as the real-time changing capacity.
2.3. Balanced assignment model of production tasks

Due to the complexity of the production control process, various factors related to the production environment often change, and various uncertain events often occur [4]. Not sure which makes the production plan, lead to the workshop in the production scheduling is more complex, in this paper, based on this background is presented based on the matching mode of planning, capacity or orders issued to workshop production center, specific row by each workshop production scheduling algorithm is optimized according to the actual production site again, this method is called after picking directly to work center model of production planning.

For the convenience of illustration, some assumptions are used to set the capacity of the current sequence of the workshop as \( N = [A_1, A_2, \ldots, A_m] \), and \( M \) is the number of work centers or equipment; There are \( n \) kinds of orders to be processed with the same delivery time. The process of order \( J \) needs to be processed on machine \( I \). For example, the processing machine matrix \( S \) of the order is described as workpiece, behavior machine/sequence, and there are \( C_J \) of workpiece \( J \). The processing time required by the process of workpiece \( J \) on machine \( I \) is described as the processing time matrix \( T \), detailed as follows.

\[
S = \begin{bmatrix}
M_{11} & M_{12} & \cdots & M_{1n} \\
M_{21} & M_{22} & \cdots & M_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
M_{m1} & M_{m2} & \cdots & M_{mn}
\end{bmatrix}
\]

\[
T = \begin{bmatrix}
t_{11} & t_{12} & \cdots & t_{1n} \\
t_{21} & t_{22} & \cdots & t_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
t_{m1} & t_{m2} & \cdots & t_{mn}
\end{bmatrix}
\]

**Figure 2.** The improved mode of production organization in manufacture enterprise

**Figure 3.** Machining time matrix
In order to make the selected homemade order process as much as possible in line with their own capacity, the strategy is proposed: set the array of homemade order as \([a_1, a_2... a_n]\), where \(a_j (j = 1,2...N)\) is 0 or 1, 1 means that the order is self-made, and 0 means outsourcing processing. Establish the model as

\[
\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}a_{ij}M_{ij}t_{ij} - A_i \quad (1)
\]

The core of model (1) is to make the processing time of the selected order process as close as possible to the self-made process. The constraint of this model meets the time of delivery and it is suitable for order selection and master plan making in a certain time window. Next, a learning algorithm is proposed to optimize model (1).

3. Design of learning algorithm to optimize capacity balance

3.1. Overview of basic reinforcement learning algorithms

Reinforcement learning [5-7] is an effective machine learning method, among which q-learning algorithm is one of the most famous reinforcement learning methods. It was proposed by Watkins in 1989. It is a reinforcement learning method similar to dynamic programming algorithm and learns optimal action strategies through interaction with the environment [8]. The goal of Q-learning algorithm is to learn how to choose the best action of each step according to external evaluation signals in a dynamic environment, which is a dynamic decision-making learning process in essence. In the Q-learning algorithm, each Q(s, a) corresponds to a corresponding Q value. During the learning process, actions are selected according to the Q value. The Q value is defined as the expected return if the current relevant action is performed and a certain strategy is followed. The optimal Q value can be expressed as \(Q^*\), which is defined as the sum of the returns to be obtained by executing relevant actions and following the optimal strategy, which is defined as

\[
Q^*(s, a) = \gamma \sum S' (s, a, s') \max Q^*(s', a') + r(s, a) ,
\]

Where, S stands for state set; T(s, a, s') represents the probability of converting from an action A to state S 'under state S; R (s, a) represents the reward for performing action A in state S.

There are many research achievements on q-learning algorithm [9]. Literature [10] proposed a new multi-agent learning algorithm, which learns the behavioral strategies of other agents through the statistics of joint actions, and ensures the selection of joint optimal actions by using the full probability distribution of agent strategy vector [11]. Learning algorithms are widely used in optimization [12], robotics, intelligent control [13], intelligent transportation and multi-intelligence systems.

3.2. New learning algorithm principle description

This paper proposes a new learning algorithm for discrete manufacturing and applies it to production task assignment. The learning algorithm also borrows some ideas from genetic algorithm [14]. That is, the system is initialized to a set of random solutions, and the optimal value is searched by iteration. In learning algorithm, the potential solutions, called "students", the students of our students in the solution space to follow the optimal search, it is important to note in the process of learning can't abandon to outstanding students to learn that an important way, this kind of outstanding students of knowledge (history) optimal solution information come down, naturally formed the tracks of the optimal solution pattern extraction, You get the optimal pattern. The optimal pattern is the extremely useful knowledge information required by iterative search. Using the global optimal pattern can provide effective guidance for further solving the student collective space, thus speeding up the convergence rate of the algorithm.

3.3. "IG type learning from excellence" learning algorithm recursive equation

Suppose m students form a group in a d-dimensional target search space, and the i-th student is expressed as a D-dimensional vector \(\mathbf{X}_i = (x_{i1}, x_{i2},... , x_{iD})\), \(i = 1,2... M\). That is, the knowledge of the ith student in the d-dimensional search space is \(\mathbf{X}_i\). \(\mathbf{i}_\text{max} = (i_{11}, i_{22},... , i_{1D})\) is the optimal knowledge searched by the ith student so far; \(\mathbf{G}_\text{best} = (g_{11}, g_{22},... , g_{D})\) is the best knowledge searched by the whole student group.
so far. If "students" only learn from individual historical best and global optimal Gbest, it is referred to as "learning algorithm only from excellent ~ Ibest and Gbest". LAOE ~ IG for short, its recursive equation is as follows

\[ X^i_i(k) = X^i_i(k) \oplus I^i_{best}(k), \quad (3) \]
\[ X^G_i(k) = X^G_i(k) \oplus G^i_{best}(k), \quad (4) \]
\[ X_i(k+1) = X^i_i(k) \cap X^G_i(k) \quad (5) \]

In equations (3), (4) and (5), \( k \) is learning (iterative) algebra; \( X^i_i(k) \) is a student collective after \( X^i_i(k) \) learns from individual excellent group \( I^i_{best}(k) \) and selects the best. \( X^G_i(k) \) is a student collective after student collective \( X_i(k) \) learns from global excellent student \( G_{best}(k) \) and selects excellent students. The algorithm uses symbol coding or natural number coding, so students' knowledge of \( X^i_i(k) \) (\( X_{i1}(k), X_{i2}(k)..., X_{iD}(k) \)) is a sign or a natural number, something like that; \( \cap \) is the student group selection symbol, that is, to choose the best as the knowledge of the next generation of the student group.

3.4. Optimize the coding of capacity balancing learning algorithm

LAOE ~ IG algorithm optimizes different problems, and its encoding, decoding and learning methods are different. The algorithm uses machine code or process code to solve production scheduling problems. In this paper, LAOE ~ IG algorithm is applied to the problem of production task selection and assignment, so a new coding method, namely 0-1 coding, is selected.

**0-1 code:** Take \( N \) production orders to be processed as an example. First, \( N \) production orders were divided into three types: \( N_1 \) which must be made by oneself, \( N_2 \) which can be made by oneself but can be outsourced, and \( N_3 \) which must be outsourced. Randomly generate the following sequence

\[ \text{sequence} = [1, 0, 0, 1, 0], \quad (6) \]

In the sequence of (6), the position where numbers 0 and 1 appear represents the serial number of the production order, and 0 and 1 indicate that the production order adopts outsourcing or self-made processing method respectively. This paper proposes a learning method of LAOE ~ IG optimization model for students of new coding (11), which is described in detail as follows.

3.5. Optimize the learning method of production equilibrium learning algorithm

In order to optimize the LAOE-IG model (2), the following mutual learning methods are proposed. The following uses 12 production orders as an example to describe how to learn from each other.

First and second stage learning (ML1, ML2) : two learning points smaller than the production order number \( N \) are randomly selected, and the workpiece in the learning point of the upperclassman (1) is copied to the immature student, and then the order outside the learning point of the upperclassman (2) is copied to form the new student with the immature student. The learning method is shown in Figure 3. The learning segment is an order drawn with lines, and the mutual learning method of multiple segments is similar. "In Figure 3," Senior, "Immature" and "New" students refer to Senior, Immature and New students respectively. LP stands for learning point.
3.6. Steps of the "Learn from best only IG" learning algorithm

Step 1: Initialize the design. Let the learning algebra $k = 0$, the student group size is $P_{size}$, and the learning termination algebra is $Maxgen$. An initial student collective $X_i(1)$ with a size of $P_{size}$ is generated, and the adaptive values of each student in the student collective are calculated. Then the optimal student $G_{best}(1)$ of the first generation is selected from the student collective. At the same time, the individual historical optimal student collective $I_{best}(1)$ of the initial student collective is randomly generated.

Step 2: Use equation (1) to obtain $2 \times P_{size}$ students, calculate their adaptive values, and then select the $P_{size}$ students with smaller adaptive values to form $X^I(k)$. Then according to equation (2), $2 \times P_{size}$ students are obtained, and their adaptive values are also calculated. The $P_{size}$ students with smaller adaptive values are selected to form $X^G(k)$. Then, the students in $X^I(k)$ and $X^G(k)$ are compared in pairs by equation (3), and the students with the smaller adaptive value form the next generation of students collective $X_i(k + 1)$. It is judged that if the adaptive value of each student in the new student collective is lower than the individual historical minimum adaptive value, the new optimal student is used to replace the individual $P_{best}$ with the optimal historical adaptive value; similarly, if the adaptive value of the student in the new student collective is lower than the global optimal value, the current optimal student is used to replace the global optimal student $G_{best}$. If $k = Maxgen$, go to Step 3; otherwise, set $k = k + 1$, go to Step 2.

Step 3: Output $G_{best}$

4. Conclusion

In recent years, the study of production scheduling is more, to solve production scheduling algorithm is more, but for production scheduling is a np-hard combinatorial optimization problem, optimization of slow convergence speed, more time consuming and storage resources, is not a kind of can adapt to and solve the problem of all kinds of scheduling algorithms, the production plan and task equilibrium and less reported.

At present, many advanced manufacturing modes are based on production scheduling, and job scheduling is the core and important part of manufacturing system. Research on scheduling methods and optimization technologies has become the basis and key to the practice of advanced manufacturing technology [15]. In these production management systems with high degree of automation, the scheduling problem to make the production process run reasonably and efficiently has become very complicated, and it is necessary to establish an effective computer scheduling control strategy, so the research on the production balance problem of job scheduling has important practical significance.

This paper discusses the problem of production task allocation based on production capacity load, how to maximize production capacity, and puts forward "IG type learning from the best" learning
algorithm, using a new 0-1 code, successfully applied to solve the problem of balanced production task allocation.

References
[1] Sohal A. A longitudinal study of planning and implementation of advanced manufacturing technologies[J]. International Journal of Computer Integrated Manufacturing, 2000( 10) : 281 – 295.
[2] Anil M, Arunkumar P. Advanced technologies and humans in manufacturing workplaces: an interdependent relationship[J]. International Journal of Industrial Ergonomics, 2004, 33( 4) : 295 – 313.
[3] Zhu V, LIAN Z G, GUAN Z L, et al. Research on Workshop Scheduling Problem and Method of Equipment Manufacturing Enterprises [J]. East China economic management, 2008,22 (7) : 133-137.
[4] Lian Zhigang. Research and Application of Planning and Scheduling Technology and Manufacturing Execution System in Equipment Manufacturing Industry [D]. Wuhan: Post-doctoral Report of Huazhong University of Science and Technology, 2009.
[5] Tang Wenbin, Zhu Miaoliang. Multi-agent Systems for Reinforcement Learning [J]. Computer science, 2003,30 (4) : 16-17.
[6] Mitchell T M. Machine Learning [M]. Beijing: China Machine Press, 2003.
[7] Zhang Rubo. Research on the Improvement of Reinforcement Learning Speed [J]. Computer Engineering and Applications, 2001(22) : 38-40.
[8] Watkins C J. Learning from delayed rewards [D]. London: PhD Thesis of the King’s College of University of Cambridge, 1989.
[9] Ye Fei, LUO Jingqing, Yu Zhifu. An Improved SVM Learning Algorithm for Parallel Processing [J]. Microelectronics and computer, 2009,26 (2) : 40-44.
[10] Song Qing-kun, Hu Ziyiing. Q-learning Algorithm Based on Empirical Knowledge [J]. Automation technology and application, 2006,25 (1) : 10-12.
[11] Guo Rui, WU Min, PENG Jun, et al. A novel multi-agent learning algorithm [J]. Acta automatica sinica, 2007,33 (4) : 367-372.
[12] Zhao X, Yao Y, Yan L P. Learning algorithm for multimodal optimization [J]. Computers and Mathematics with Applications, 2009, 57: 2016 – 2021.
[13] dos Leandro S C, Rafael B G. Chaotic synchronization using PID control combined with population based incremental learning algorithm [J]. Expert Systems with Applications, 2010, 37: 5347 – 5352.
[14] Yin Aiying, Wu Yunbing, Yang Xiaohua. Hybrid sampling algorithm for manufacturing unbalanced data [J]. Computer engineering and design,2018,39(04):1053-1058.
[15] Poornima B, Rajesh G, Lance S. Accuracy of reinforcement learning algorithms for predicting aircraft taxi – out times: a case study of Tampa Bay departures [J]. Transportation Research Part C, 2010, 18: 950 – 962.