# Reinforcement Learning for Production Planning with Demand Sensitive to Delivery Lead Time

**Chi-Yang Tsai, Erickson Liang**
Department of Industrial Engineering and Management, Yuan Ze University
135 Yuan-Tung Road, Chung-Li, Taoyuan 32003, Taiwan
iecytsai@saturn.yzu.edu.tw; s1055440@mail.yzu.edu.tw

## Extended Abstract

Machine learning techniques are developing at an accelerated rate in recent years and are being applied to solve complex problems in various areas. Among them, reinforcement learning does not require a pre-determined environment model and learns from experience of interacting with the environment in the past. It has the capability to solve complex sequential decision-making problems. Companies are facing greater competition in today's fast changing markets. Delivery lead time is becoming an important factor in competition between companies for customers. If a company tends to take longer time to deliver its products, customers may grow tired of waiting and turn to other places to satisfy their needs. In addition, customers can share the experience with others more easily and more quickly through the internet and thus influence future demand of the company. It is vital for companies to cautiously manage delivery lead time in order to maintain expected service quality. Facing fast changing demand, well developed production strategies are needed, especially with limited production capacity.

This research considers a production system facing stochastic demand that is sensitive to delivery lead time, that is, if it takes longer time to satisfy demand, potential customers will be discouraged and future demand will drop. In the studied problem, the amount of supply that can be used to meet the demand in a time period includes the production quantity in the period and the on-hand inventory at the beginning of the period. The length of the delivery lead time of the period is determined by the percentage of demand that can be met by the supply of the period. The base lead time is achieved if all demand is met. If supply is less than demand, some customers are unable to receive what they order within expected time. To these customers, the lead times are longer than the base the lead time and so is the average delivery lead time of the period. A greater percentage of unmet demand results in a longer average delivery lead time and consequently lesser amount of future demand.

Decisions on the production quantity need to be made in each period, given the current inventory/backorder status and finite production capacity. The objective of this research is to develop effective production strategies which map the best actions to every system state by utilizing reinforcement learning's capability to deal with complex stochastic systems. Computer simulation experiment is conducted with random customer demand to examine how reinforcement learning can adapt to uncertain environments. A zero-inventory production strategy is used on the same problems to evaluate the performance of the reinforcement learning algorithm. Under this strategy, the exact quantity is produced to provide enough supply to match the demand of the same period, therefore leaves no inventory at the end of each period. The results of the conducted numerical experiments show that the production strategy obtained by using the reinforcement learning algorithm reduces the amount of backorders by 15.81% and average profit is increased by 3.67%.