**Reinforcement Learning - Supply Chain Ordering Management: An application to the beer game**

**Introduction**

In order to illustrate the ideas of dynamic systems, Jay Forrester of MIT first created the iconic Beer Distribution Game in the late 1950s. In this instance, the dynamic system is a supply chain that transports beer from a brewery to the final customer. The supply chain structure and game rules are straightforward, but the ensuing behavior is highly complex, which is what makes the game so exciting.

A way to creating simulation models of real-world systems is agent-based modeling. Agents that are used in agent-based modeling capture entities and exhibit behavior in response to internal or external events. In the beer distribution game, the players will be represented by agents. Instead of teaching the agents a set of predefined behaviors, the agents will be trained to play the beer distribution game using machine learning techniques, particularly a technique known as Reinforcement Learning. Reinforcement learning emphasizes interaction-based learning. The agent must experiment with different actions to determine which ones have the best results rather than being instructed which to do. One approach to reinforcement learning is Q-learning.

In this paper, the goal is to implement a Q-learning algorithm to deal with inventory management challenges such as changing demands, varying lead times, and/ or disruptions.

The structure of the essay is as follows: The reinforcement learning problem is briefly described in section 2. The generated data to test and train the algorithm is shown in section 3. In section 4, the Q-learning algorithm will be implemented and adjusted to account for Deep Q-networks. In section 5, the findings on model performance after testing different parameters will be summarized and it will be discussed, on how well reinforcement learning suits the applied problem setting.