1CM240

TU/e

AI in Logistics
Assignment 2 (50% of the course)

Preliminary information

- Hand in: Concise report with the main outcomes and findings, well documented code for the implementation of the RL algorithms and the simulator.
- The report should at least contain a formal mathematical MDP definition of the problem being studied, an explanation of the RL algorithms being applied, and some numerical analysis on the applied network.
- It is understandable that not all methods might be successful in the end. Assessment of this assignment is therefore not only based success of the implemented algorithms but also on the quality of the approach.
- This assignment considers an open challenge for which you need to implement a dynamic solution approach based on the principles of deep reinforcement learning. It requires at least the implementation of a deep reinforcement learning algorithm and a sufficient benchmark of that approach.
- Python code with a gym implementation is provided.
- An additional document with installation steps for the gym is provided.
- Note: The tutorial on March 10, 2023 will help you understanding the provided code.

The assignment

We consider a dynamic inventory routing problem in which stores need to be replenished from a single warehouse dynamically.

Upon replenishing the stores, a fixed order cost depending on the replenished stores is incurred. These fixed order costs represent the transportation cost for the respective vehicle route among the replenished stores. This cost consists of a fixed cost for using a truck, and a fixed cost for each distance unit traveled. For simplicity, you receive a simulator with a Gym, implemented in the popular OpenAI gym framework. Within the gym, a vehicle routing solver is present to determine the shortest vehicle route among a set of stores.

In addition to the transportation costs, you have lost sales and holding costs for sales not being fulfilled and inventory that carries over to the next period (in a similar fashion as Q5 of assignement 1). All discussed cost parameters are part of the gym, and are clearly indicated in the supplied python code.

You are self-responsible for the set-up of a suitable network of stores and demands. As an inspiration, the gym contains all locations of jumbo stores in the Eindhoven area with demand distributions inspired by the size of the stores. The current demand distribution are very simplistic. You need to improve upon this. Furthermore, once you have found a good approach to solve this problem for a (sensible) fixed set of parameters, you can experiment with your set-up and understand the consequences on the policies developed.

Figure 1: Supermarket Network

Assessment

- **Quality of the assumptions, set-up of the study, and motivation of choices: 15%**
  - The problem needs to be formalized. Proper notation should be introduced and an MDP needs to be defined.
  - If assumptions are made they should be realistic and not reduce the problem to much.
- **Implementation of DRL method and benchmarks. Quality, correctness and novelty. 40 %**
  - For the benchmark we expect at least an (s,S) policy to be implemented.
  - You get rewarded for a good benchmark - so do not try to reduce the quality of your benchmark to make your DRL look better.
  - In general, we prefer a good DRL solution over a good benchmark (in case you need to decide where to put your time)
  - For the DRL, we will in particular look at how you deal with the action space.
  - The routing algorithm we gave you is slow and is very likely too slow to train DRL. You may need an approximation for the routing cost.
- **Quality of the numerical study: 35%**
  - Good choice of hyperparameters.
  - Good comparison of benchmark and DRL.
  - Comparison of training and testing.
- **Quality of the report: 10%**
  - We expect proper formalization of the problem and use of consistent notation.
  - The report should be self-contained. It should not be necessary to go into the code to understand your implementation.
  - The report should be properly formatted (Latex is highly advised).