Tabular Reinforcement Learning for Revenue Management Problems

INFORMS Annual Meeting 2024

Adam Page

Supervisors: James Grant, Chris Kirkbride, Anna-Lena Sachs

Industrial Advisors: Jeremy Bradley, Alexander Green, Jordan Palmer

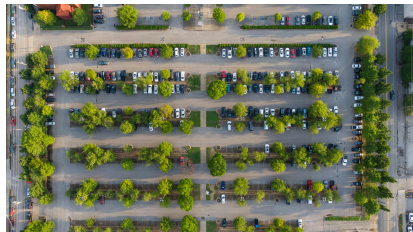October 20, 2024

STOR-i | Lancaster University   DATASPARQ

Problem

Airport Car Parks

A pre-bookable car park:

- Limited spaces available
- Set one price per day
- Each customer has:
    - A willingness to pay
    - A lead time
    - Length of stay
    - A buying preference

Revenue Management (RM)

Find an optimal pricing policy of a perishable good. Classically, solved via dynamic programming (DP) methods.

Set Up

Markov Decision Process

- **States:** Car parking spaces
- **Actions:** Prices we can offer
- **Transition Probabilities:** Number of spaces sold
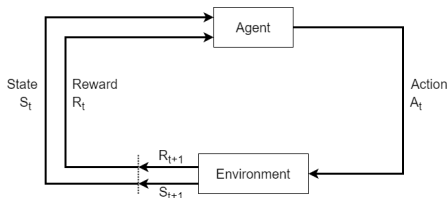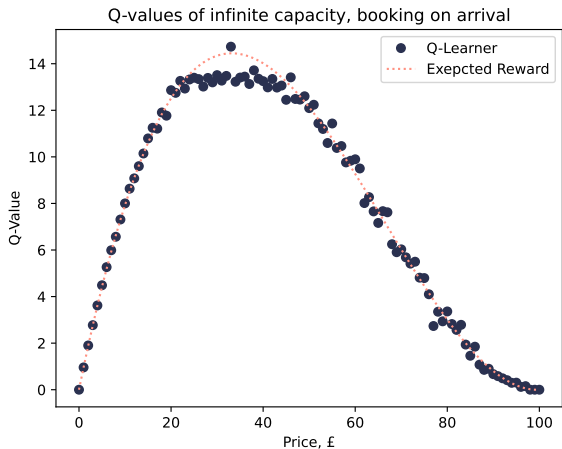- **Rewards:** Revenue per customer

Reinforcement Learning (RL)

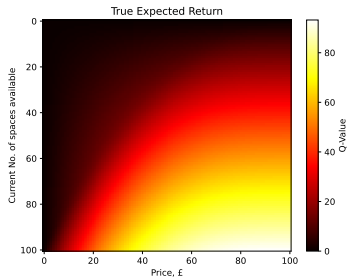Figure: Agent-Environment interactions (Sutton & Barto, 2018)

Can a Q-learner retrieve the expected reward function?

Q-values of infinite capacity, booking on arrival
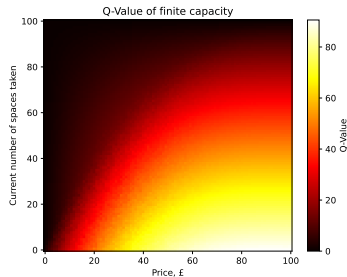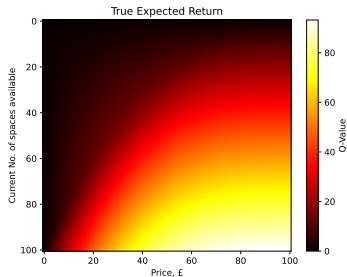
Car park now has a maximum capacity, so we look at return rather than reward.

Introducing States

Car park now has a maximum capacity, so we look at return rather than reward.

Thank You For Listening.
Any Questions?

Also happy to speak about:

- ▶ Interpretability/explainability of Q-Learning
- ▶ Intuitive 'model-based' RL methods