POSTER ID

THE END OF AIRLINE REVENUE MANAGEMENT AS WE KNOW IT?

QUAN NGUYEN*, NICOLAS BONDOUX*, RODRIGO ACUNA-AGOST, THOMAS FIIG
Innovation and Research Division
Amadeus SAS

Revenue Management (RM) is a key element for airlines to maximize revenue. RM systems performance heavily relies on the quality of the forecast and modelling. To overcome this, we use Reinforcement Learning (RL) which doesn't need any forecast or modelling. It is theoretically proven that RL will converge to the optimal solution. However, in practice, the system may require a lot of data (e.g., thousands of years of historical bookings) to learn the optimal policies. To overcome these issues, we present a novel model that integrates domain knowledge powered by a deep neural network. The results show very encouraging results with different numerical /simulated scenarios. We believe this opens the door to a new generation of revenue management systems that could automatically learn by interacting with competitors and customers, so it can react much faster to market changes.

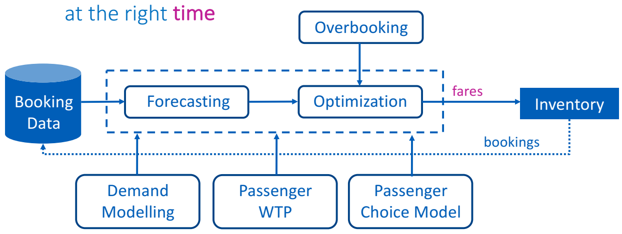
Research Questions

1. Could RL really learn RM's policy?
2. How much time/data does RL need to learn
3. Could RL do it better than RM?

Introduction

Revenue Management System

- Goal: maximizing **flight** revenue
- Action: set the right **fare** at the right **time**

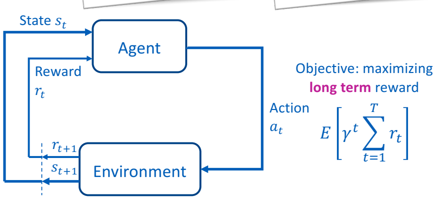


Reinforcement Learning

Learning from **interactions** to achieve a **goal**

Key elements:

- *Agent*: the learner and decision maker
- *Environment*: everything outside of the agent which has interactions with the agent
- *Rewards*: a signal whose values the agent tries to maximize over time.



RM and RL fundamentals

Optimality equation for the value function (Bellman 1952)

$$V(t,x) = Max_f\left[P(f)\big(f + V(t+1, x+1)\big) + (1 - P(f))[0 + V(t+1,x)]\right]$$

$$V(s) = Max_a \underbrace{\sum_{s'} P^a_{ss'}[R^a_{ss'} + V(s')]}$$

Optimality equation for the action-value function

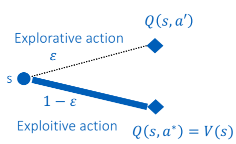
$$Q(s,a) = \sum_{s'} P^a_{ss'}[R^a_{ss'} + Max_{a'} Q(s',a')]$$

$$V(s) = Max_a\, Q(s,a)$$

Q-learning (Watkins 1989)

$$Q(s,a) \leftarrow \alpha[f + Max_{a'} Q(s',a')] + [1-\alpha]Q(s,a)$$

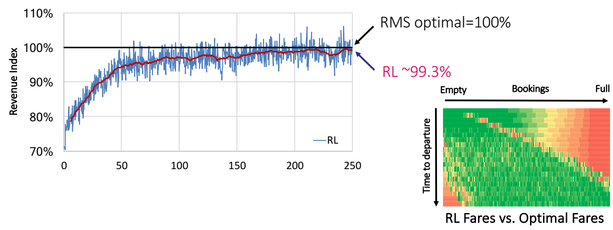
Learning rate



Experiments & Results

The results below were obtained using Amadeus Air Travel Market Simulator. The simulations were performed with a single-leg network.

Could RL learn RM's policy?



RL is able to learn RM's optimal policy in a monopoly scenario, so how about adding a competitor?

Did RL learn fast enough?

Yes, if we use deep neural network (Deep Q-Learning)!
But actually we don't care since we can pre-train the neural network with RM's policy using supervised learning.

It can copy RM's policy, but could it do better?

Conclusion

- Reinforcement Learning opens the door to a radical new approach
  - Model free – no forecasting and no optimization
  - Learns by direct price testing
- Why DQL vs. DQL underperforms?
  - The RM systems "agree" on the fundamentals, makes them able to extract more from the market
  - Prisoner's dilemma – incentive for first-mover

* Joint first authors