Market Making via Reinforcement Learning

Thomas Spooner
Department of Computer Science
University of Liverpool
t.spooner@liverpool.ac.uk

John Fearnley
Department of Computer Science
University of Liverpool
fearnley@liverpool.ac.uk

Rahul Savani
Department of Computer Science
University of Liverpool
rahul.savani@liverpool.ac.uk

Andreas Koukorinis
Stratagem Technologies Ltd,
and University College London
andreas@stratagem.co

RL for Market Making

Markov Decision Process (MDP):

State space:

- Agent state: e.g., remaining position
- Environment state: market signals - e.g., spread, LOB liquidity, price offsets, volatility

Action space:

- Limit order book levels at which market maker may choose to post liquidity

Transition probabilities from one state to another under certain action:

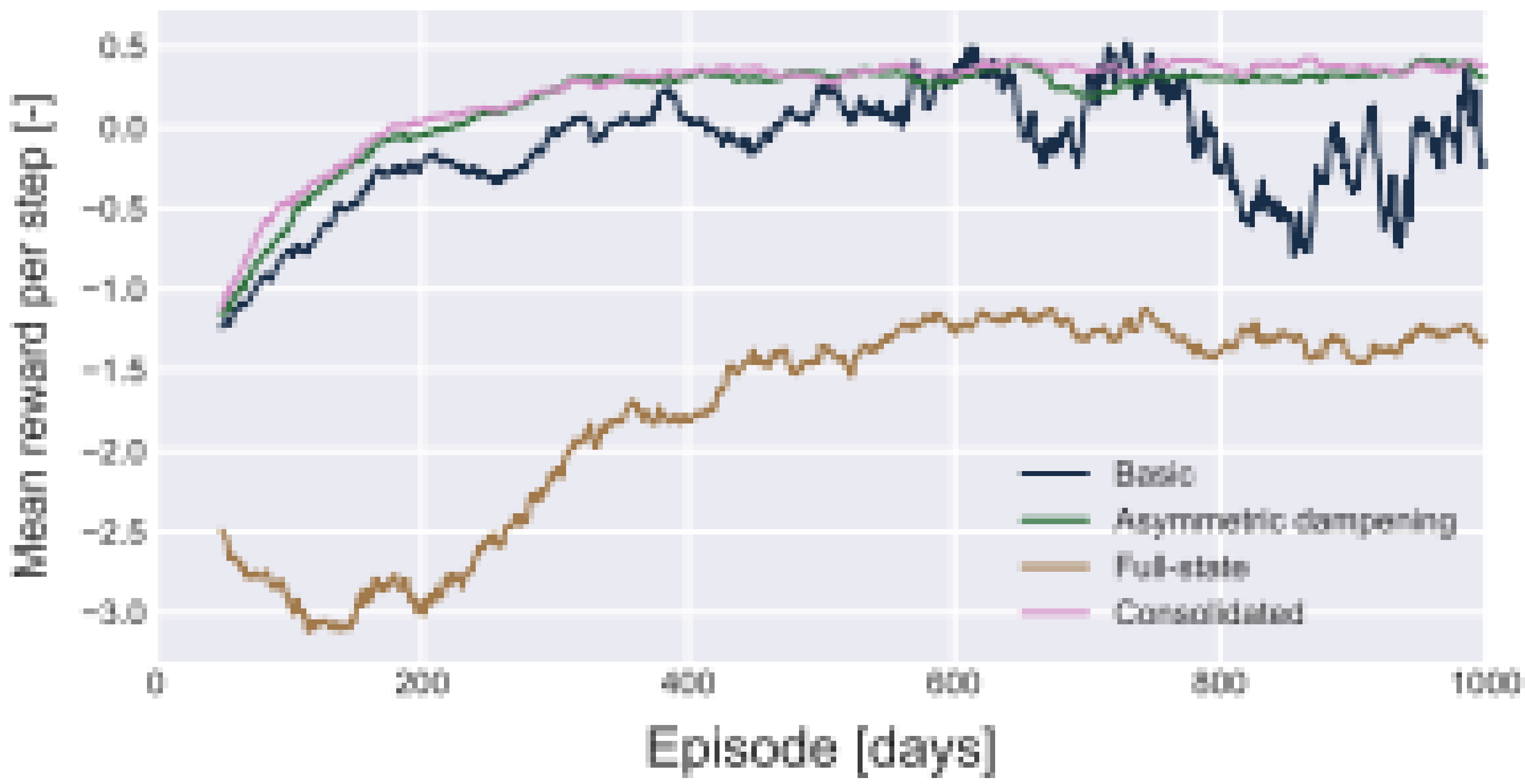
- Transition probabilities are implied by the **simulator**

Reward for taking action a at state s

- *Matched Order PnL + Inventory PnL*
- Another version: *Matched Order PnL + Inventory PnL – max(0, α*Inventory PnL)* – encourage more spread capture and introduces risk-averse behavior

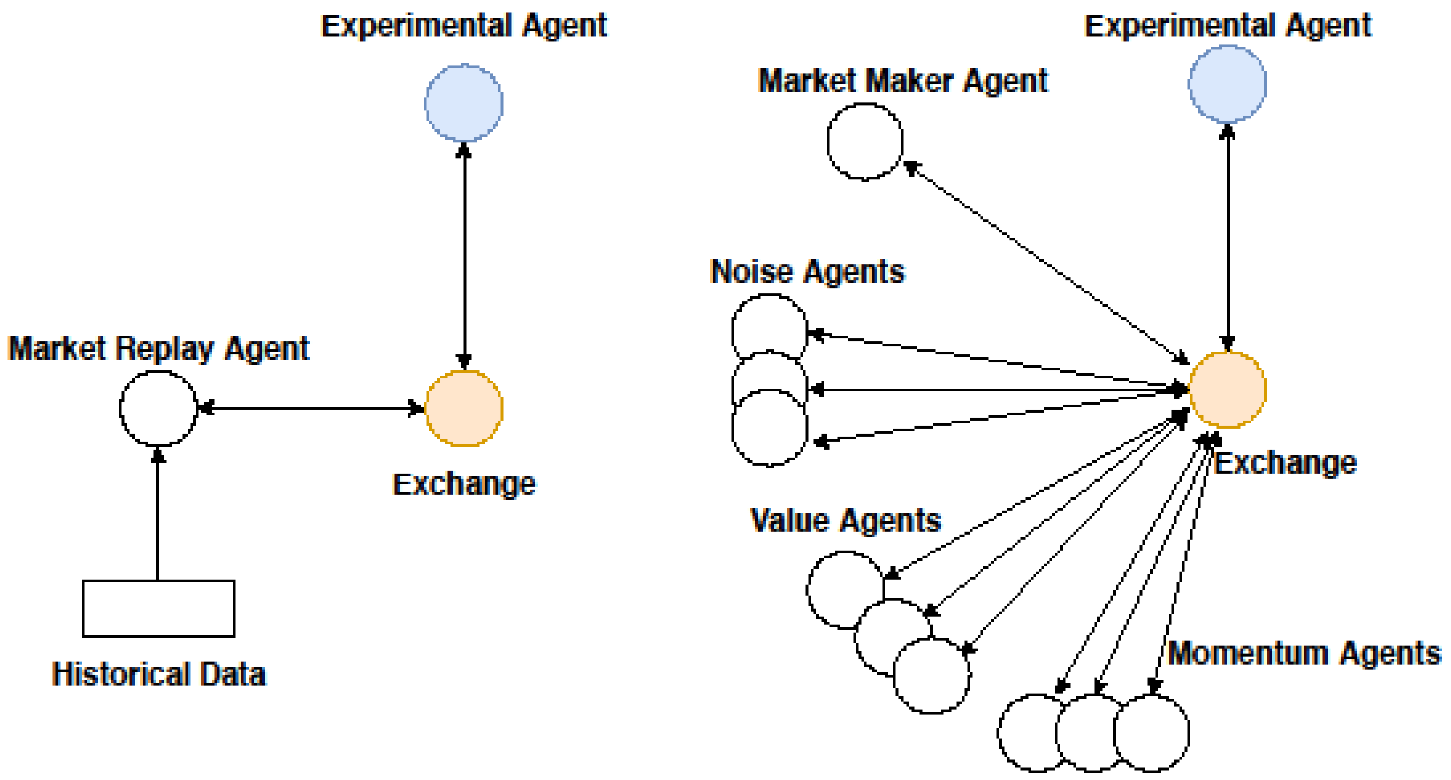
Objective:

Maximize expected cumulative rewards



0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
Mean reward per step [-]
0
200
400
600
800
1000
Episode [days]
Basic
Asymmetric dampening
Full-state
Consolidated

Multi-Agent Simulation



Why multi-agent simulation?

- Development and test of new trading strategies (especially reinforcement learning strategies) with the goal of simulating market response to agent’s actions
- Controlled trading strategy experiments
- Better understanding of rare events (e.g. flash crashes, market shocks) - regulatory applications

Simulation Realism

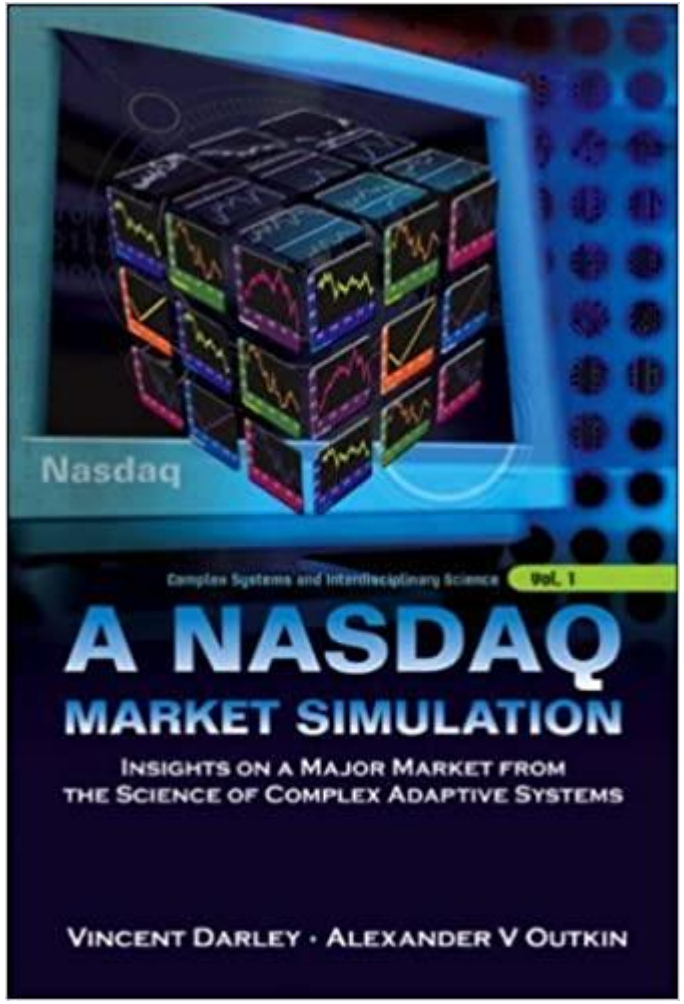
- Simulation is intended to model counterfactual scenarios
- Needs to be responsive the experimental agent's actions
- Multi-agent simulation useful as it's natural bottom-up approach

- What does it mean to "calibrate" a multi-agent simulator?
  - "Get real paper": simulation with respect to stylized facts
- How to validate/apply simulation results in practice?
- What is multi-agent simulation realism?

NASDAQ Market Simulation, 2007

- Until 1997, tick size $1/8
- In 1997, changed to $1/16
- In 2001, changed further changed to $.01

- Goal of research: understand how tick size reduction impacted the market



Market Model

- Exchange
- Agents:
  - **Investors** – observe the noisy version fundamental price and the market and decide whether to buy/sell (informed) or trade randomly
  - **Dealers (market maker)** – must post liquidity on both sides of the market
    - Variety of strategies
    - Of particular interest are parasitic dealers where liquidity is reposted inside the spread
    - Dealers can learn to make their strategies more profitable

Calibration

- **Statistical calibration**:
  - Confirm that stylized facts hold
  - Make sure volume distribution in a particular time period historical
- **Behavioral calibration**:

Calibrate dealer strategies to historical data (so that the composition is representative of the market over historical period)

Stock	Date	Time-stamp	MM	Bid	Bid vol.	Ask	Ask vol.
XYZ	970501	10:02:16	MMM1	14.625	10	15.125	10
...	...	...	...	...	...	...	...
XYZ	970501	10:04:58	MMM1	14.25	10	15.125	10
XYZ	970501	10:05:00	MMM1	14.25	10	14.75	10

Results

- Price discovery is impeded for smaller tick sizes (i.e. mid price does not track the fundamental)
- Dealers learn parasitic strategies

- The above were confirmed in a variety of agent scenarios – both historical and uncalibrated (period of low and high volatility, etc), proportions of agent with different strategies

- Results confirmed by NASDAQ report

Suggestions by authors:

- Do NOT use multiagent simulation for price prediction

- Qualitative rather than quantitative use cases are more appropriate
- Multiagent simulation is most appropriate to study macroscopic structural effects