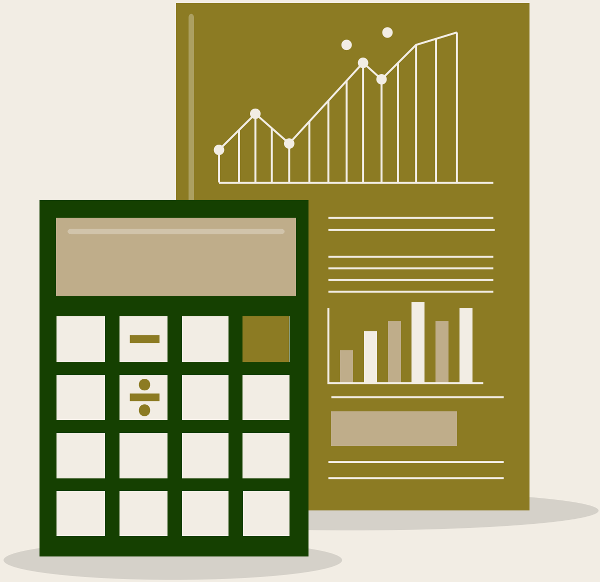
Application of Reinforcement Learning in Automated Market-Making

Made by Team 2



INTRODUCTION

The paper focuses on improving the reinforcement learning within Glosten-Milgrom model, especially by making a risk-sensitive RL.

Market Mechanisms in Financial Exchanges

ORDER-DRIVEN

Buyers and sellers submit their limit or market orders directly to the exchange.

Orders are processed by a price-time priority matching engine, operating in a double-sided auction format.

QUOTE-DRIVEN

Markets are typically characterized by market-makers who provide quotes for buying and selling securities.

Market-makers ensure liquidity and continuous trade flow by quoting prices at which they are willing to buy and sell.

MARKET-MAKING

Market-making involves simultaneously buying and selling stocks, aiming to profit from the bid-ask spread and providing liquidity to the market.

Asymmetry of the information

Somebody knows better the true value of an asset.

The risk of holding inventory

Market-makers aims to have zero inventory by the end of the time period

Market-makers aim:

- **To maximize profits while efficiently managing their inventory**
- **To focus on controlling the bid-ask spread to maintain market liquidity in a competitive trading environment**

THE AIM OF THE PAPER

Implementation and analysis of both risk-neutral and risk-sensitive versions of SARSA and Double-SARSA.

Introduction of a new framework for automated market making using risk-sensitive RL.

Modelling agent-environment interactions through a finite Markov Decision Process in discrete time, utilizing a linear reward function.

Simulation of the financial market based on the Glosten-Milgrom information model, extending the Chan-Shelton market-making model.

The novelty of the paper

"While all of these components have been studied before, our approach provides a novel combination, in particular the use of risk-sensitive learning as a way to develop a market-making strategy."

-Paper's Authors

SIMULATION OF FINANCIAL MARKET

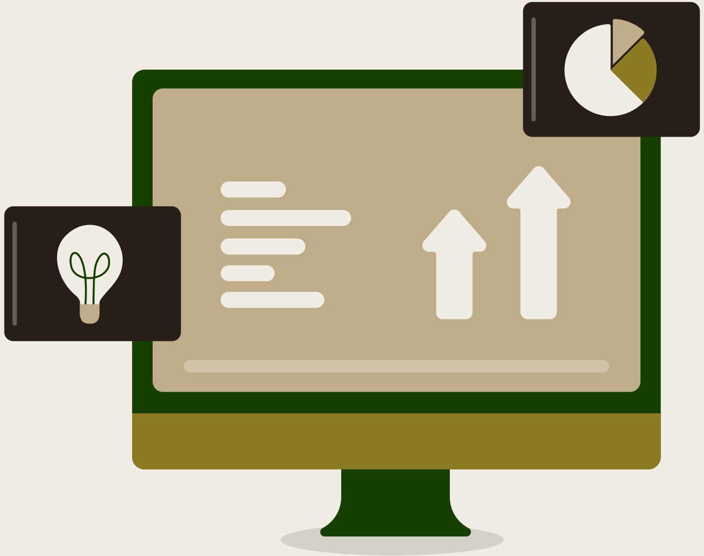
Glosten-Milgrom model with a market-making agent and groups of informed and uninformed traders

Uninformed traders randomly trade, while informed traders follow a buy-low sell-high strategy.

For full order execution:

- the market-making agent has passive limit orders
- other trading agents have aggressive market orders

Events, including trader arrivals and value shifts, occur at each time step with defined probabilities, maintaining a balance of $2\lambda_p + 2\lambda_u + \lambda_i = 1$.



02

ALGORITHMIC MARKET-MAKING

MARKOV DECISION PROCESS

Finite stationary MDP is represented as a tuple M = ⟨S, A, R, P⟩. It includes a set of states (S), actions (A), rewards (R), and transition probabilities (P).

$$R_t = w_{pro}\Delta PRO_t - w_{inv}|INV_t| - w_{spr}SP_t$$

$$\mathcal{P}(s', r \mid s, a) = \mathbb{P}(S_{t+1} = s', R_{t+1} = r \mid S_t = s, A_t = a)$$

w_pro, **w_inv**, **w_spr** – weight parameters for profit, inventory and market-maker's bid-ask spread.
IMB_t – the volume imbalance (number of **buy** orders **minus** the number of **sell** orders for a unit of share of the asset).

$$\Delta PRO_{t+1} = \begin{cases} ASK_t - P_t & \text{If a trader buys} \\ P_t - BID_t & \text{If a trader sells} \end{cases}$$

Reinforcement Learning

RL in Market-Making

RL is used to develop an optimal policy for dynamically updating bid and ask prices, aimed at maximizing long-term profit.

Stationary Policies

The study focuses on stationary, time-invariant policies, where actions are selected based on the current state in a Markovian fashion.

Policy Optimization

The goal is to find a policy that not only maximizes profits but also effectively manages inventory and the bid-ask spread.

Stochastic Policy Approach

Stochastic policies are employed to ensure sufficient exploration by the agent, balancing between learning from new experiences and utilizing known strategies.

RISK SENSITIVE RL

Risk-sensitive Approach

Risk-sensitive RL focuses on strategies that account for the uncertainty and variability in rewards, emphasizing the importance of risk assessment in decision-making processes.

Enhanced Decision-Making

By incorporating risk sensitivity, RL models are better equipped to make decisions that balance potential gains with associated risks, leading to more robust and adaptive strategies in dynamic environments.

Application in Market-Making

In the context of market-making, risk-sensitive RL aids in developing strategies that manage not only profit maximization but also effectively mitigate risks such as inventory holding and adverse selection.

METRICS

Average absolute price deviation is the difference between fundamental price and bid/ask values

$$\overline{\Delta P} = \sum_{t=1}^{T} | BID_t - P_t | + | ASK_t - P_t |$$

End of episode profit is the sum of all profits acquired at the end of each period

$$PRO_T = \sum_{t=1}^{T} \Delta PRO_t$$

End of episode inventory is the amount of asset market-maker has left, it should be close to zero for the efficient market-maker

Average episodic spread, SP, measures the average quoted spread for an episode

$$\overline{SP} = \frac{1}{T} \sum_{t}^{T} SP_t$$

Limitations of the Bayesian Learning

Although the Bayesian Learning is fully capable of tracking the fundamental value process over the long run, but it fails to immediately track this value following sudden and large jumps. This can potentially expose the market-maker to the risk of adverse selection, which our risk-sensitive RL framework handles better.

Softmax function

Risk-sensitive reinforcement learning algorithm driven by softmax action selection not only results in reduced variance for the three metrics, but also yields a higher average profit. Theoretically, convergence of softmax is not guaranteed.

Main limitations

Main limitations of the model is the simplicity of the Glosten-Milgrom model, which may perform worse in a more competitive environment with multiple market-makers and more realistically behaving trading agents.

Future research

Implementing Risk-Sensitive RL Functional Approximation Techniques

Developing a More Realistic Limit Order Book

Expanding State and Action Space

Integrating Advanced Information-Based Models

EXPERIMENTS

A1

Risk-neutral epsilon-greedy SARSA

A2

Risk-sensitive epsilon-greedy SARSA

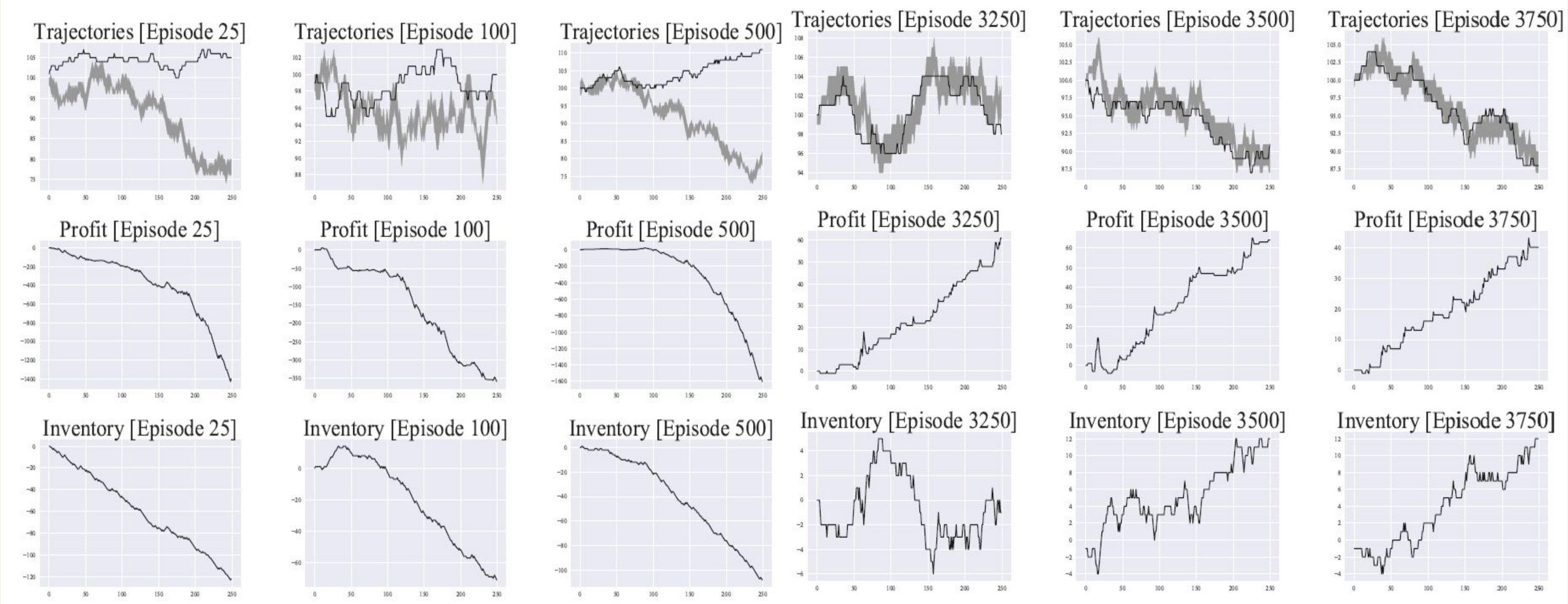
A3

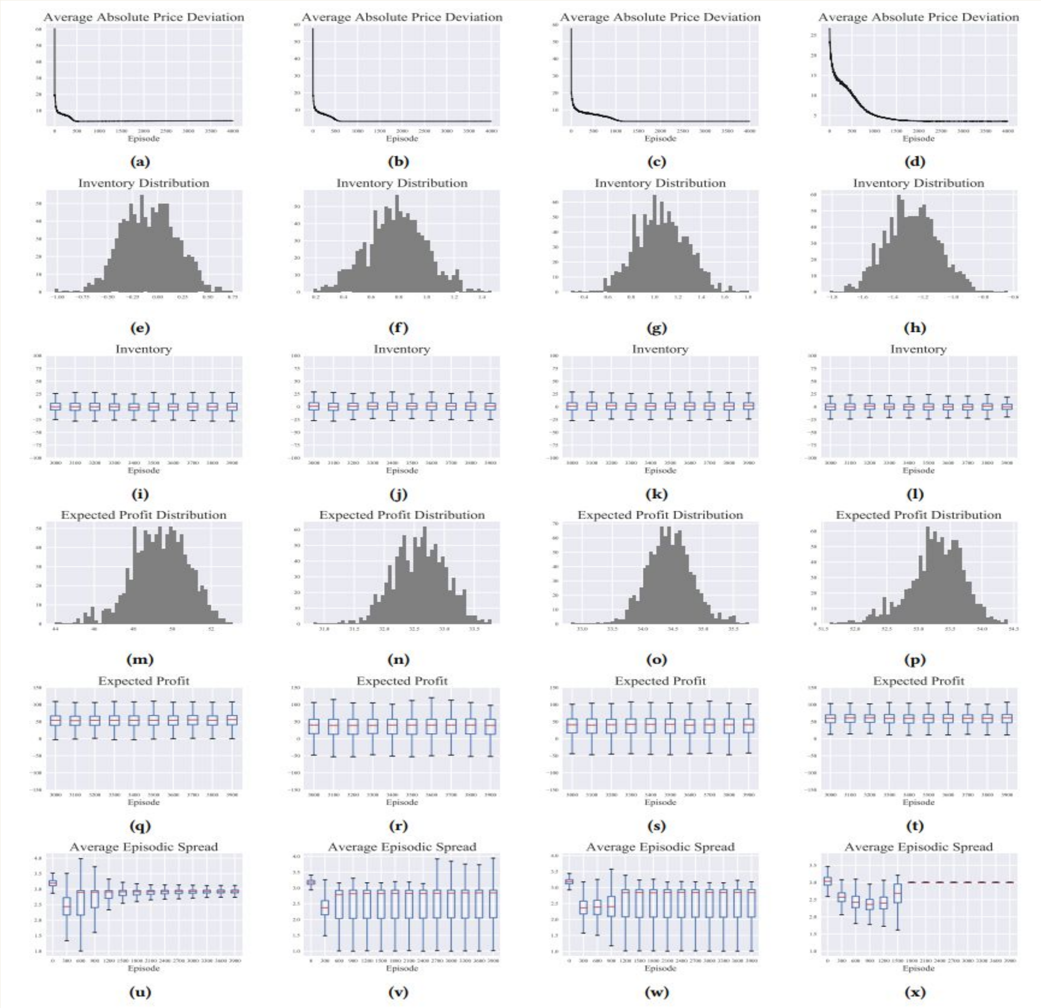
Risk-sensitive epsilon-greedy Double-SARSA

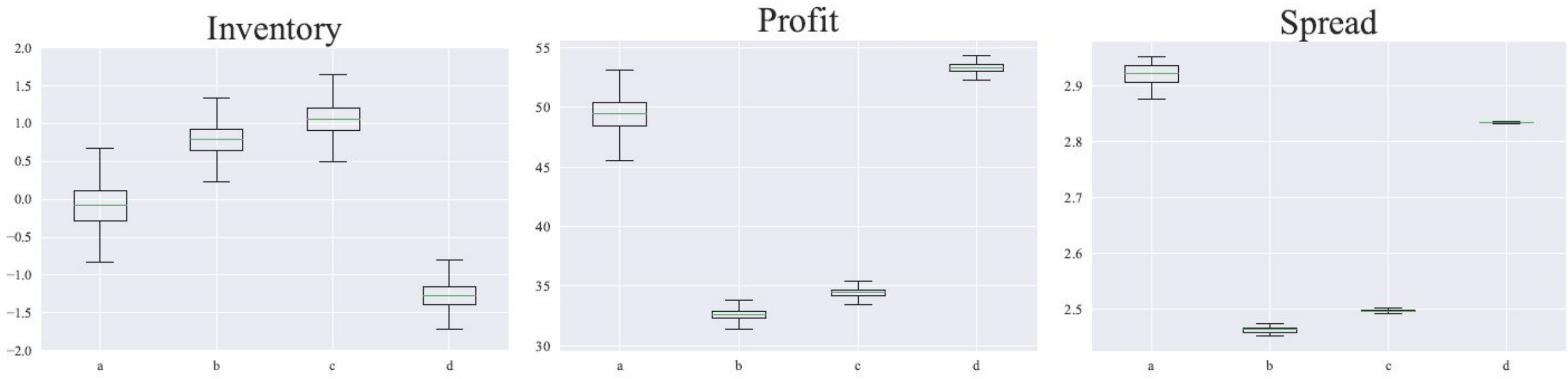
A4

Risk-sensitive softmax Double-SARSA

EXPERIMENT RESULTS




Average Absolute Price Deviation
Episode
(a)
Average Absolute Price Deviation
Episode
(b)
Average Absolute Price Deviation
Episode
(c)
Average Absolute Price Deviation
Episode
(d)
Inventory Distribution
(e)
Inventory Distribution
(f)
Inventory Distribution
(g)
Inventory Distribution
(h)
Inventory
Episode
(i)
Inventory
Episode
(j)
Inventory
Episode
(k)
Inventory
Episode
(l)
Expected Profit Distribution
(m)
Expected Profit Distribution
(n)
Expected Profit Distribution
(o)
Expected Profit Distribution
(p)
Expected Profit
Episode
(q)
Expected Profit
Episode
(r)
Expected Profit
Episode
(s)
Expected Profit
Episode
(t)
Average Episodic Spread
Episode
(u)
Average Episodic Spread
Episode
(v)
Average Episodic Spread
Episode
(w)
Average Episodic Spread
Episode
(x)



(a) (b) (c)

Figure 4: Comparison of (a) inventory, (b) profit and (c) spread at the end of episodes in the optimal region across the four different RL algorithms. Within each plot, the algorithms are, left to right, Risk-neutral ϵ-greedy SARSA (A1); Risk-sensitive ϵ-greedy SARSA (A2); Risk-sensitive ϵ-greedy Double SARSA (A3); and Risk-sensitive Double SARSA Softmax (A4).

CONCLUSION

Comparing different risk-sensitive and risk-neutral algorithms, we observed that risk-sensitive models have less average profit and lower spread, however they also have lower variance and lower average inventory.