Reinforcement Learning for Electronic Market-Making

Christian R. Shelton

Artificial Intelligence Laboratory and
The Center for Biological and Computational Learning
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: Many economic markets, including major stock exchanges, employ market-makers to aid in the transactions and provide a better quality market. Market-makers supply an advantage to the market. By consolidating the trading in a few agents, the market becomes more efficient. Traders wishing to buy and sell do not need to find each other or wait for each other's arrival. Additionally, by quoting a single price which is guaranteed for all traders, market-makers remove the price fluctuations that occur in markets where buyers and sellers must come to their own agreement for a price individually for each transaction. Markets with market-makers have greater volumes and better price stability.

Many major markets are now electronic. The NASDAQ is a distributed trading system completely run through networked computers. It uses competing market-makers (usually one per major trading company) to maintain a high quality market. However, the demands on human market-makers are high. A typical market-maker will be responsible for 10 to 20 securities. At any given moment, it is only feasible for the market-maker to be actively attentive to 2 to 3 of them. The market-maker is generally losing potential profit or volume on the other securities.

The last few years have also seen the growth of on-line trading systems. These systems are also entirely electronic and usually employ no market making. Orders are crossed against the other orders that happen to be present at that time of the trade and otherwise are dropped.

The goal of this research is to demonstrate the ability of reinforcement learning to fill the need for automated market-making. For the case of the NASDAQ, a learning system could fill the role of an "autopilot" by taking care of stocks in a more intelligent manner while being supervised by a human market-maker. This would allow a human market-maker to more successfully manage a large set of securities. In the case of small on-line trading systems, the system could replace the existing naive order crossing mechanism to provide a better market to its traders.

Previous Work: Theoretical market-making models have been developed in the context of stochastic dynamic programming. Bid and ask prices are dynamically determined to maximize some long term objectives such as expected profits or expected utility of profits. Models in this category include those of [3], [4] and [2]. The main limitation of these models is that specific properties of the underlying processes (price process and order arrival process) have to be assumed in order to obtain a closed-form characterization of strategies.

In our previous work [1], we used a temporal difference algorithm from the reinforcement learning literature to build an adaptive electronic market-making. However, temporal difference methods implicitly assume the environment is fully-observable. This places certain restrictions on the market-maker. By adopting different reinforcement learning algorithms, we have been able to overcome these constraints.

Approach: We use a novel reinforcement learning technique to construct a system that is dynamically responsive to changing market environments. In particular we build upon a new importance sampling estimator we developed to explicitly deal with stochastic and complex environments such as stock markets [5]. The estimator naturally allows the adaptive control policy to use memory and function approximation to aid in solving the complex task.

Additionally, the estimator provides an estimate over the entire policy space simultaneously. This allows our algorithm to provide for natural controls to trade-off between competing goals. For instance, the market-maker may want to reduce the spread of its quote (thereby providing a better quality market) while at the same time maximizing its profit. Figure 1 demonstrates this type of control. On a simulated market, we ran the algorithm on a series of run. On each run, the algorithm was given a different minimum allowable profit and was asked to minimize the spread subject to this constraint without any knowledge about the environment, using only experience

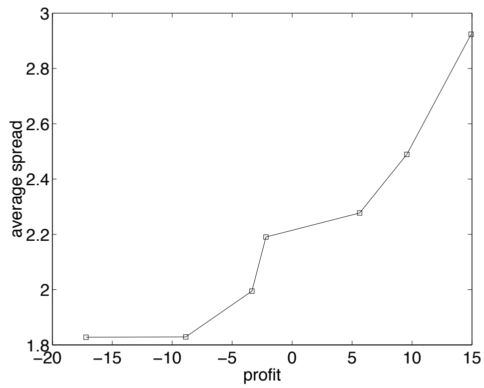

Figure 1: Spread verses profit / day found automatically by reinforcement learning. The lower-right corner is the most desired point: maximal profit with minimal spread. One comes at the expense of the other. See [6] for full details on the market simulation.

gained by interacting with the market. Figure 1 shows the average spread and profit for each of these policy (as calculated after learning by separate computation). This demonstrates our ability to provide the user of the system with natural parameters (minimum desired profit in this case) by which to control the behavior of the system and balance the multiple goals. The full details of this system are available in [6].

Impact: By providing automated market-making software, this research will allow for more efficient economic trading systems. Marking-making firms can use it to aid human market-makers and small exchanges can use it to provide better quality markets without the need to employ human market-makers.

Future Work: Our technique is quite general. It assumes nothing about the market environment and instead develops its strategy based on experience. We can therefore apply this method without change to other market situations. We are currently developing more realistic market simulations based on NYSE and NASDAQ trading data. Additionally, we are incorporating more information sources and trading actions to allow the system more complete control over its trading policy.

Research Support: This research was sponsored by grants from Office of Naval Research (DARPA) under contract No. N00014-00-1-0907, National Science Foundation (ITR) under contract No. IIS-0085836, National Science Foundation (KDI) under contract No. DMS-9872936, and National Science Foundation under contract No. IIS-9800032.

Additional support was provided by: Central Research Institute of Electric Power Industry, Center for e-Business (MIT), Eastman Kodak Company, DaimlerChrysler AG, Compaq, Honda R&D Co. Ltd., Komatsu Ltd., Merrill-Lynch, NEC Fund, Nippon Telegraph & Telephone, and Siemens Corporate Research, Inc., Toyota Motor Corporation, and The Whitaker Foundation.

References:

[1] Nicholas Tung Chan and Christian Shelton. An electronic market-maker. Technical Report AI-MEMO 2001-005, MIT, AI Lab, April 2001.

[2] Lawrence R. Glosten and Paul R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14:71–100, 1985.

[3] Thomas Ho and Hans R. Stoll. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics*, (9):37–73, 1981.

[4] M. O'Hara and G. Oldfield. The microeconomics of market making. *Journal of Financial and Quantitative Analysis*, 21:361–376, December 1986.

[5] Christian R. Shelton. Policy improvement for POMDPs using normalized importance sampling. In *Proceedings of the Seventeenth International Conference on Uncertainty in Artificial Intelligence*, pages 496–503, 2001.

[6] Christian Robert Shelton. *Importance Sampling for Reinforcement Learning with Multiple Objectives*. PhD thesis, Massachusetts Institute of Technology, August 2001. also available as AI Lab tech report 2001-003.