Three Automated Stock-Trading Agents: A Comparative Study

Alexander A. Sherstov and **Peter Stone**

Department of Computer Sciences
The University of Texas at Austin

July 19, 2004

Overview

- 3 autonomous stock-trading agents
  - **RL** = reinforcement learning
  - **TF** = trend following
  - **MM** = market making

Alexander Sherstov
Peter Stone

Overview

- 3 autonomous stock-trading agents
  - **RL** = reinforcement learning
  - **TF** = trend following
  - **MM** = market making

- Fully implemented and tested
  - **Individual** simulations
  - **Joint** simulation
  - **PLAT Stock-Trading Competition**

Alexander Sherstov
Peter Stone

Overview

- 3 autonomous stock-trading agents
  - **RL** = reinforcement learning
  - **TF** = trend following
  - **MM** = market making

- Fully implemented and tested
  - **Individual** simulations
  - **Joint** simulation
  - **PLAT Stock-Trading Competition**

- MM: winner in fall 2003, runner-up in spring 2004

Motivation and Background

- Motivation for autonomous stock trading
  - **On-line** bid submission
  - **Real-time order-book** info
- Penn-Lehman Automated Trading (PLAT) simulator
  - Merges **virtual** (agent-generated) and **real** orders
  - **Price dynamics** affected by virtual orders

The PLAT Simulator

BUY ORDERS

SHARES	PRICE
500	**24.062**
6,000	24.061
5,000	24.055
...	...
3,000	**24.040**

SELL ORDERS

SHARES	PRICE
500	**24.069**
500	24.069
200	24.070
...	...
2,800	**24.100**

- Activity occurs in cycles
  - **withdraw** placed orders
  - **place** new buy/sell orders

Alexander Sherstov
Peter Stone

The PLAT Simulator

BUY ORDERS

SHARES	PRICE
500	**24.062**
6,000	24.061
5,000	24.055
...	...
3,000	**24.040**

SELL ORDERS

SHARES	PRICE
500	**24.069**
500	24.069
200	24.070
...	...
2,800	**24.100**

- Activity occurs in cycles
  - **withdraw** placed orders
  - **place** new buy/sell orders
- Order matching most to least competitive

Alexander Sherstov
Peter Stone

The PLAT Simulator

BUY ORDERS

SHARES	PRICE
500	**24.062**
6,000	24.061
5,000	24.055
...	...
3,000	**24.040**

SELL ORDERS

SHARES	PRICE
500	**24.069**
500	24.069
200	24.070
...	...
2,800	**24.100**

- Activity occurs in cycles
  - **withdraw** placed orders
  - **place** new buy/sell orders
- Order matching most to least competitive
- Benchmark strategy: SOBI [Kearns '03]
  - Provided as **example**
  - Computes volume-weighted $\overline{s}$, $\overline{b}$
  - Sells when $|p - \overline{b}| > |p - \overline{s}|$, buys otherwise

Competition Details

- No position limits; **can sell short**
- Large **penalty for leftover shares** at market close
  - Position restrictions limit **agents' impact on real economy**
- **daily-score** $=$ **profit** $+ \$0.002 \cdot$ **rbt-shares** $- \$0.003 \cdot$ **fee-shares**
  (no commission; fees exactly as on Island ECN)

Alexander Sherstov
Peter Stone

Performance Criterion

- Obvious metric **inadequate: aggregate profit**
  - **disregards stat. significance**
  - **ignores risk** (variance in earnings)

Performance Criterion

- Obvious metric **inadequate: aggregate profit**
  - **disregards stat. significance**
  - **ignores risk** (variance in earnings)

- PLAT metric:

$$\text{Sharpe ratio} \stackrel{\text{def}}{=} \frac{\text{ave. score}}{\text{std. deviation}}$$

  - measures **statistical significance** of earnings
  - "**most widely used** measure of risk-adjusted return"
    $\Rightarrow$ well-suited for **day trading**

Related Work

- Prior research in automated stock trading
  - automated market making [Chan '01, Das '03, Feng '04]
  - use of RL for on-line parameter adjustment [Chan '01]
  - "reverse" strategy [Yu '03]
  - VWAP trading

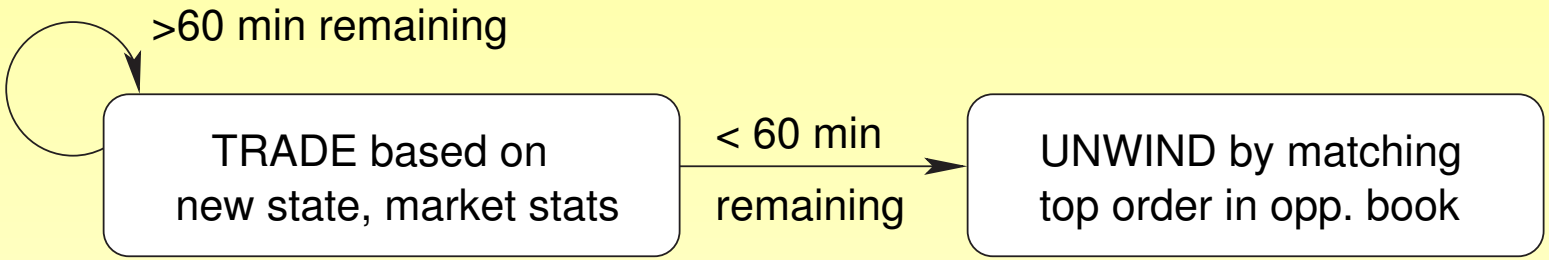
Related Work

- Prior research in automated stock trading
  - automated market making [Chan '01, Das '03, Feng '04]
  - use of RL for on-line parameter adjustment [Chan '01]
  - "reverse" strategy [Yu '03]
  - VWAP trading

- Contributions
  - study of **heterogeneous** strategies in a **joint** economy
  - use of a highly **realistic stock simulator**
  - use of **Sharpe ratio**
  - **NOT** aiming for a deployable strategy

Approach and Assumptions

- Generic architecture:



- Profit maximization, unwinding **assumed independent**

- **Trading strategy** abstracted in TRADE module:
  RL, TF, MM

Reinforcement Learning Agent: Intro

- Motivation
  - **on-line adjustment** to diverse economy
  - **minimal expertise** coded in

    $\Rightarrow$ RL makes decisions every tick

Reinforcement Learning Agent: Intro

- Motivation
  - **on-line adjustment** to diverse economy
  - **minimal expertise** coded in
    $\Rightarrow$ RL makes decisions every tick

- Problem specification: quadruple $\{\mathcal{S}, \mathcal{A}, T, R\}$
  - $\mathcal{S}$, **environment's states**; $\mathcal{A}$, **agent's actions**
  - $T : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$, **transition function**
  - $R : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$, **reward function**
  - T, R unknown to agent
  - Goal: **policy** $\pi : \mathcal{S} \rightarrow \mathcal{A}$ that maximizes return, $\sum_{t=0}^{\infty} \gamma^t r_t$

Reinforcement Learning Agent: Design

- **Model**
  - **Challenge:** include **relevant vars**, keep task **manageable**
  - $\mathcal{S}$: $\Delta p_t = p_t - \overline{p_t}$, where $\overline{p}_t = \beta \overline{p}_{t-1} + (1-\beta) p_t$
  - $\mathcal{A}$: share volume $\in [-900, 900]$
  - R : diff. in present value (cash + shares @ last price)
  - **tile coding** function approximation

Reinforcement Learning Agent: Design

- **Model**
  - **Challenge:** include **relevant vars**, keep task **manageable**
  - $\mathcal{S}$: $\Delta p_t = p_t - \overline{p_t}$, where $\overline{p}_t = \beta \overline{p}_{t-1} + (1-\beta) p_t$
  - $\mathcal{A}$: share volume $\in [-900, 900]$
  - R : diff. in present value (cash + shares @ last price)
  - **tile coding** function approximation

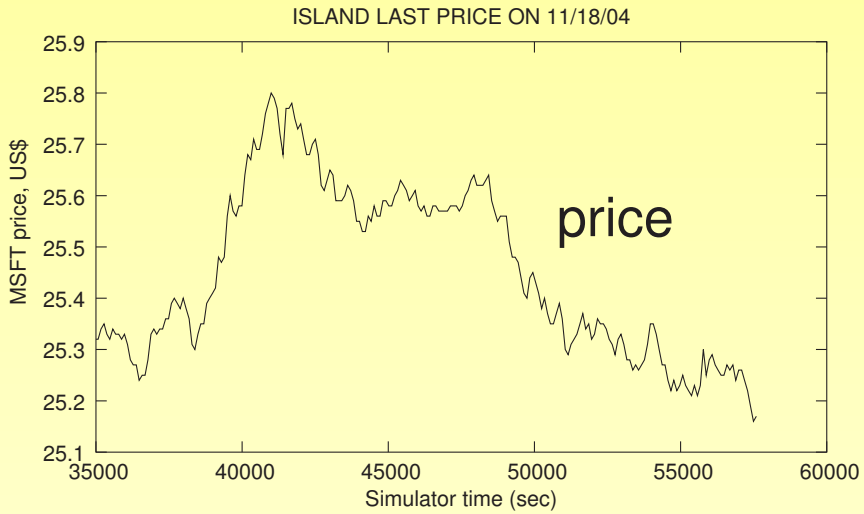
- **Parameter choices**
  - Sarsa
  - $\alpha = 0.04$, $\gamma = 0.8$, $\epsilon = 0.1$, $\lambda = 0.7$
  - $\beta = 0.999$

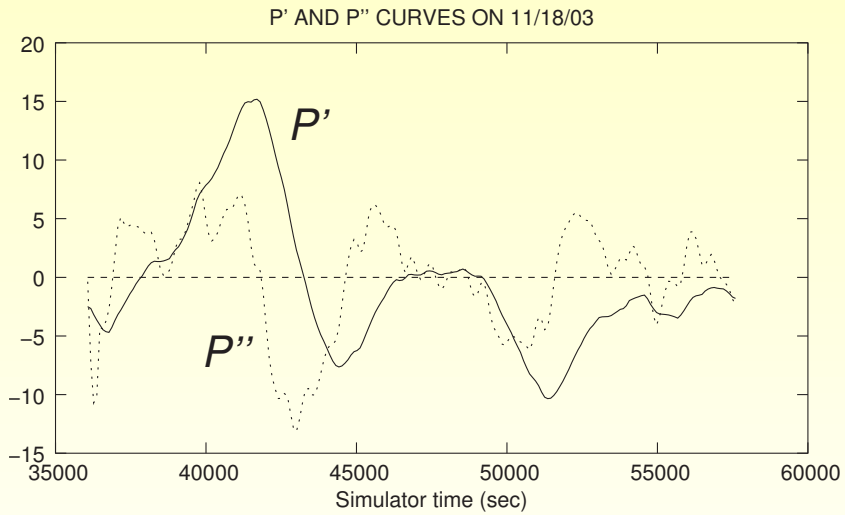
Trend Following Agent: Intro

- **Heuristic** strategy, as opposed to RL
- Uses linear regression to **identify price trends**
- Works by approximating P' and P''

Trend Following Agent: Intro

- **Heuristic** strategy, as opposed to RL
- Uses linear regression to **identify price trends**
- Works by approximating P' and P''





Trend Following Agent: Design

- Use of P' and P'':
  - P' indicates **current trend**
  - P'' signals **trend reversal**

Alexander Sherstov
Peter Stone

Trend Following Agent: Design

- Use of P' and P'':
  - P' indicates **current trend**
  - P'' signals **trend reversal**

- Strategy
  - If $P' > 0$ and $P'' > 0 \longrightarrow$ **buy stock**
    (price increasing at an increasing rate),
  - If $P' < 0$ and $P'' < 0 \longrightarrow$ **sell stock**
    (price falling at an increasing rate),
  - Otherwise $\longrightarrow$ **unwind**
    (trend reversal underway)

Market Making Agent

- Capitalizes on **small fluctuations**, not longterm trends
- Keeps **near-zero share position** instead of unwinding during reversal

Market Making Agent

- Capitalizes on **small fluctuations**, not longterm trends
- Keeps **near-zero share position** instead of unwinding during reversal
- Implementation
  - Uses **same prediction model** (P', P'') as TF
  - Buys and sells as TF, but...
  - ...places **orders in pairs**, adding a **small** profit margin to take advantage of current trend

Alexander Sherstov
Peter Stone

Market Making Agent

- Capitalizes on **small fluctuations**, not longterm trends
- Keeps **near-zero share position** instead of unwinding during reversal
- Implementation
  - Uses **same prediction model** (P', P'') as TF
  - Buys and sells as TF, but...
  - ...places **orders in pairs**, adding a **small** profit margin to take advantage of current trend
  - Example: when $P' > 0$ and $P'' > 0$, place (BUY,p) and (SELL,p + PROFIT-MARGIN)

Evaluation Considerations

- **Any strategy** will do well on some days, poorly on others.
- Ex post **optimality not attainable** w/o knowledge of price behavior
- Reasonable approach: evaluation on a set of **representative market dynamics**

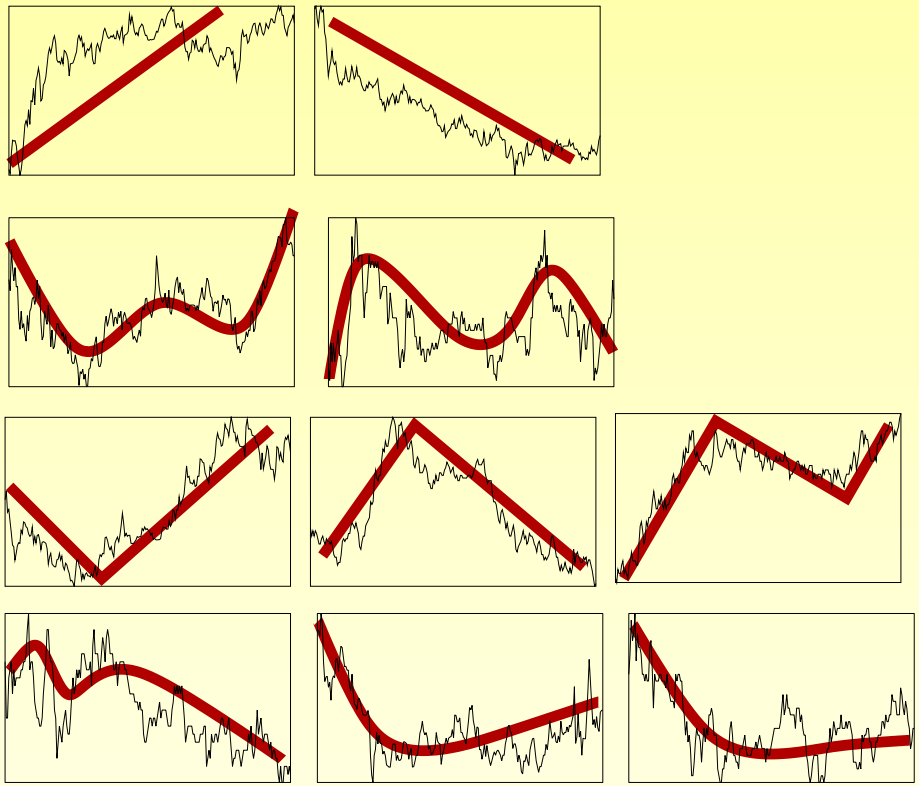
Evaluation: Market Conditions

Monotonic (M)

Substantial fluctuation (F)

Zigzag behavior (Z)

Mixed/other (O)



Alexander Sherstov
Peter Stone

Individual Evaluation: RL

		RL	SOBI
*	M	$11,134	$21,935
*	M	$45,680	$56,308
	F	$5,142	$55,710
	F	$50,529	$17,464
	Z	$69,683	$230,715
*	Z	$358,774	$96,387
	Z	$284,563	$11,059
*	O	$49,621	$13,805
	O	$3,407	$25,026
	O	$2,302	$29,015

M monotonic **F** fluct.
Z zigzag **O** other
***** agent wins

- Effect of price diff parameter:
  - successful under M
  - fails under F (freq. reversals)
  - satisfactory under Z, O (trends longer)
- Major stumbling block: exogenous transition model
- More "focused" RL [Tesauro '02]:
  - construct market model
  - use DP to compute order

Alexander Sherstov
Peter Stone

Individual Evaluation: TF

		TF	SOBI
*	M	$4,015	$29,686
*	M	$3,591	$44,216
	F	$4,292	$108,476
	F	$1,533	$19,958
	Z	$4,390	$155,539
	Z	$3,163	$32,383
*	Z	$479	$1,964
*	O	$5,494	$12,063
	O	$4,139	$118,016
	O	$4,692	$23,098

M monotonic
Z zigzag
***** agent wins

F fluct.
O other

- A single profitable day!
- Steady value loss typical
- Analysis
  - beats SOBI under M
  - fails under too-short/too-long trends (M,F)
  - strongest under medium-duration trends (Z)
  - Problem: premature unwinding

Alexander Sherstov
Peter Stone

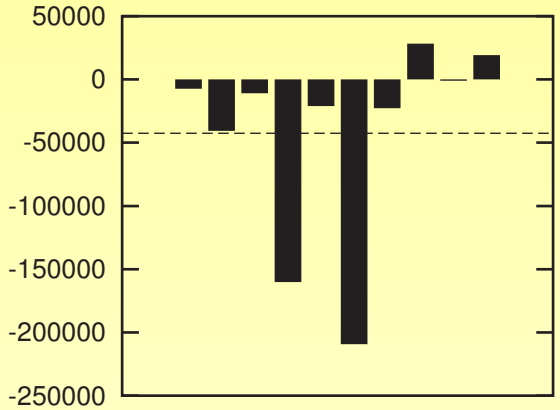
Individual Evaluation: MM

		MM	SOBI
*	M	$529	$30286
*	M	$972	$52255
	F	$471	$117192
	F	$1131	$24908
	Z	$518	$154082
	Z	$3370	$15605
*	Z	$744	$2417
*	O	$654	$22632
	O	$638	$85099
	O	$1224	$27467

M monotonic
Z zigzag
***** agent wins
F fluct.
O other

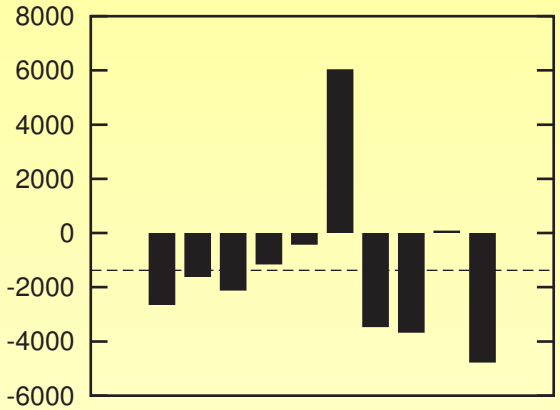
- Comparative evaluation
  - profitable 70% of the time
  - beats SOBI on 4 days
  - small but consistent profits
- Notes
  - fails under Z: share imbalance due to trend reversal never eliminated

Joint Simulation Results



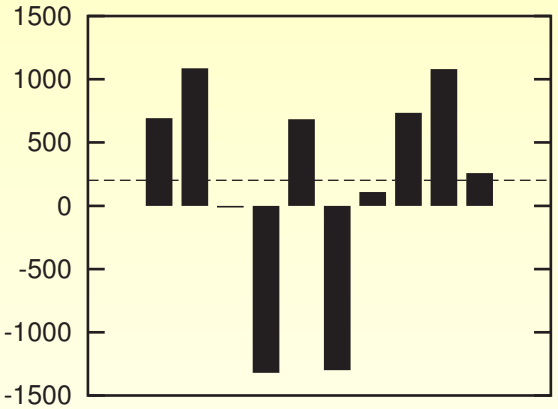
RL

ave=$42,553 **Sharpe**=−0.54



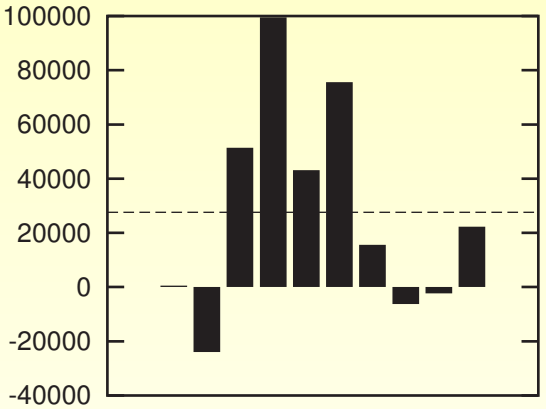
TF

ave=$1,377 **Sharpe**=−0.46



MM

ave=$201 **Sharpe**=0.23



SOBI

ave=$27,546 **Sharpe**=0.70

Live Competition Results: Top 3

December 2003

Date	MM	#2	#3
12/9	135	7447	4106

Alexander Sherstov
Peter Stone

Live Competition Results: Top 3

December 2003

Date	MM	#2	#3
12/9	135	7447	4106
12/10	381	3006	3254
12/11	436	1365	5971
12/12	140	848	322
12/13	62	2536	1334
12/16	439	3716	3940
12/17	359	3501	7924
12/18	411	1037	2163
12/19	430	4617	119
12/20	679	1692	645

Live Competition Results: Top 3

December 2003

Date	MM	#2	#3
12/9	135	7447	4106
12/10	381	3006	3254
12/11	436	1365	5971
12/12	140	848	322
12/13	62	2536	1334
12/16	439	3716	3940
12/17	359	3501	7924
12/18	411	1037	2163
12/19	430	4617	119
12/20	679	1692	645
Ave	347	1487	1411
St. dev.	185	3378	3772
Sharpe	1.88	0.44	0.37

Live Competition Results: Top 3

December 2003

Date	MM	#2	#3
12/9	135	7447	4106
12/10	381	3006	3254
12/11	436	1365	5971
12/12	140	848	322
12/13	62	2536	1334
12/16	439	3716	3940
12/17	359	3501	7924
12/18	411	1037	2163
12/19	430	4617	119
12/20	679	1692	645
Ave	347	1487	1411
St. dev.	185	3378	3772
Sharpe	1.88	0.44	0.37

April 2004

Date	#1	MM	#3
04/26	3433	271	1045
04/27	1374	538	4729
04/28	2508	242	243
04/29	2928	248	6694
04/30	3717	13	12508
05/03	3444	636	11065
05/04	1322	386	2377
05/05	3300	452	5708
05/06	2199	461	9271
05/07	966	121	11755
Ave	2519	239	4725
St. dev.	1009	316	6551
Sharpe	2.50	0.76	0.72

Alexander Sherstov
Peter Stone

Live Competitions: Analysis

- Winner in December 2003, runner-up in April 2004

Live Competitions: Analysis

- Winner in December 2003, runner-up in April 2004
- Earnings much smaller than competitors' but more consistent

Live Competitions: Analysis

- Winner in December 2003, runner-up in April 2004
- Earnings much smaller than competitors' but more consistent
- Solid profitability record
  - Only agent to make money on all 10 days in Dec. 2004
  - 18/20 overall profitability record, with minor losses on two other days

Conclusions

- studied **heterogeneous** strategies in a **joint** economy
  - RL, TF, MM
- used a highly **realistic stock simulator**
- used **Sharpe ratio**

Conclusions

- studied **heterogeneous** strategies in a **joint** economy
  - RL, TF, MM
- used a highly **realistic stock simulator**
- used **Sharpe ratio**
- **NOT** aimed for a deployable strategy, but...

Conclusions

- studied **heterogeneous** strategies in a **joint** economy
  - RL, TF, MM
- used a highly **realistic stock simulator**
- used **Sharpe ratio**
- **NOT** aimed for a deployable strategy, but...

  designed a consistently profitable trading agent

Future Work

- RL agent
  - More focused use of RL
- TF, MM agents
  - Improved trend detection model
  - (MM) On-line adjustment of trade size, profit margin