Automated Agents in Financial Markets and the Problem of Ad Hoc Coordination in Heterogeneous Competitive Agent Populations

Stefano V. Albrecht  
s.v.albrecht@sms.ed.ac.uk

Subramanian Ramamoorthy  
s.ramamoorthy@ed.ac.uk
Structure of the Talk

1. Autonomous Agents in Financial Markets

2. The Ad Hoc Coordination Problem

3. Harsanyi-Bellman Ad Hoc Coordination
Autonomous Agents in Financial Markets
Stock Prices

- Stock Price
- Supply & Demand
- Market Participants
Stock Prices

Stock Price

Supply & Demand

Market Participants
Market Participants

- Human Agents
- Automated Agents
Market Participants

Facts:

- More than 70% of US stock trading volume in 2009 and approx. 50% in 2012
- Market crashes of 1987 and 2010 believed to be due to automated agents
Market Participants

Facts:

- More than 70% of US stock trading volume in 2009 and approx. 50% in 2012
- Market crashes of 1987 and 2010 believed to be due to automated agents
Market Participants

Facts:
- More than 70% of US stock trading volume in 2009 and approx. 50% in 2012
- Market crashes of 1987 and 2010 believed to be due to automated agents
What is an Autonomous Agent?

Lives in an **environment**
  - e.g. financial market

Can **see** things in environment
  - e.g. stock prices

Can **do** things in environment
  - e.g. buy/sell stocks

Accomplish tasks **without human intervention**
  - e.g. maximise profits
Research in Autonomous Trading Agents

- Artificial Intelligence
- Game Theory
- Computer Science
- Economics

Autonomous Agents
Research in Autonomous Trading Agents

Trading Agent Competition (TAC)
- Annual competition held since 2000
- Produced wealth of papers on autonomous trading agents
- http://tradingagents.eecs.umich.edu

Penn-Lehman Automated Trading Project (PLAT)
- Developed the “Penn Exchange Simulator”
- Merges automated agent orders with real-world market data
- www.cis.upenn.edu/~mkearns/projects/plat.html
Research in Autonomous Trading Agents

Human-machine experiment at IBM (Das et al., 2001)
- Tested 2 agent algorithms against non-professional humans
- In markets with 6 automated agents and 6 human agents
- Agents achieved consistently higher profits than humans

Remake of IBM experiment (De Luca and Cliff, 2011)
- Tested 4 agent algorithms against non-professional humans
- Same experimental setup as IBM experiment
- Agents again consistently higher profits than humans
The Ad Hoc Coordination Problem
The Other Market Participants

Markets populated by very heterogeneous crowd:
- Could be humans or automated agents
- May have very different strategies
- Other distinguishing features

Problem:
- I don’t know who the others are and how they behave
- Hence, I cannot *a priori coordinate* my strategy

What is needed is an autonomous agent which is able to deal with this problem in a principled way.
Motivation (2 Examples)

“Jump & Dump” (Savani and Veal, 2005)

- Winning agent of the 2005 PLAT competition
- Took control of the market by buying large shares
- Other agents were not prepared for such a strategy

Markets with varying populations (Kim et al., 2012)

- Experiments in the “Social Ultimatum Game”
- 10 populations based on 5 agent types (including human data)
- Agents did well in own/similar populations but failed elsewhere
The Ad Hoc Coordination Problem (informal)

Design agent (the ad hoc agent) which is able to achieve optimal flexibility and efficiency in a multiagent system that admits no prior coordination between the ad hoc agent and the other agents.

- **Flexibility**: Ad hoc agent’s ability to solve its task with a variety of other agents in the system
- **Efficiency**: Relation between the ad hoc agent’s payoffs and time needed to solve its task
- **No prior coordination**: Ad hoc agent does not a priori know who the other agents are and how they behave
Motivation (outside finance)

Ad hoc coordination is relevant for many applications
  - Financial markets are an example

Human-machine interaction cast as ad hoc coordination:
  - Humans have variable behaviour (*flexibility*)
  - Expect agent to perform quickly (*efficiency*)
  - Often no description of behaviour (*no prior coordination*)
  - Examples: robots in homes or shops, video games

Thus, useful to study ad hoc coordination in general setting
Formal Model

Related problem known as **incomplete information game**
- Each player has “private information” hidden from others
- Private information relevant for player’s decision making

Can be modelled as **Bayesian game** (Harsanyi, 1967)
- **Types** determine players’ payoffs and strategies
- Allows reduction to complete information game

But: Solutions not directly applicable to ad hoc coordination
- Focus on equilibrium attainment, not flexibility and efficiency

⇒ Still, can use notion of private information for modelling
Stochastic Bayesian Games

A stochastic Bayesian game (SBG) consists of:

- state space $S$ with initial and terminal states $S^0, \bar{S} \subset S$
- players $N = \{1, \ldots, n\}$ and for each $i \in N$:
  - set of actions $A_i$ (where $A = \times_i A_i$)
  - type space $\Theta_i$ (where $\Theta = \times_i \Theta_i$)
  - payoff function $u_i : S \times A \times \Theta_i \to \mathbb{R}$
  - strategy $\pi_i : \mathbb{H} \times A_i \times \Theta_i \to [0, 1]$ where $\mathbb{H}$ is set of histories $H^t = \langle s^0, a^0, \ldots, s^t \rangle$ s.t. $s^\tau \in S, a^\tau \in A$
- state transition function $T : S \times A \times S \to [0, 1]$
- type distribution $\Delta : \mathbb{N}_0 \times \Theta \to [0, 1]$
Definitions

In the following, let

- $\Gamma$ be a SBG
- $\mathbb{D}$ be a set of type distributions $\Delta$ for $\Gamma$
- $\rho = \langle s^0_{\rho}, \theta^0_{\rho}, a^0_{\rho}, s^1_{\rho}, \theta^1_{\rho}, a^1_{\rho}, \ldots, s^{t_{\rho}}_{\rho} \rangle$ be a path in $\Gamma$
  - with $s^\tau_{\rho} \in S$, $\theta^\tau_{\rho} \in \Theta$, $a^\tau_{\rho} \in A$, and $s^0_{\rho} \in S^0$
- $\Phi$ be the set of all terminating paths $\rho$ in $\Gamma$ (i.e. $s^{t_{\rho}}_{\rho} \in \bar{S}$)
- $\Pr(\rho|\Gamma, \Delta)$ be the probability of $\rho$ in $\Gamma$ with type distribution $\Delta$

$$\Pr(\rho|\Gamma, \Delta) = \prod_{\tau=0}^{t_{\rho}-1} \Delta(\tau, \theta^\tau_{\rho}) \cdot T(s^\tau_{\rho}, a^\tau_{\rho}, s^{\tau+1}_{\rho}) \prod_{k \in \mathbb{N}} \pi_k(H^\tau_{\rho}, (a^\tau_{\rho})_k, (\theta^\tau_{\rho})_k)$$
Flexibility & Efficiency

The **flexibility** $F(\alpha|\Gamma, \mathcal{D})$ and **efficiency** $E(\alpha|\Gamma, \mathcal{D})$ of ad hoc agent $\alpha$ (controlling player $i$) in $\Gamma$, with respect to $\mathcal{D}$, are defined as

$$F(\alpha|\Gamma, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{\Delta \in \mathcal{D}} \sum_{\rho \in \Phi} \Pr(\rho|\Gamma, \Delta)$$

$$E(\alpha|\Gamma, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{\Delta \in \mathcal{D}} \sum_{\rho \in \Phi} \overline{\Pr}(\rho|\Gamma, \Delta) \left( \sum_{\tau=0}^{t_{\rho}-1} u_i(s^\tau_{\rho}, a^\tau_{\rho}, \alpha) \right)^{r_1} \overline{(t_{\rho})}^{r_2}$$

where $\overline{\Pr}(\rho|\Gamma, \Delta) = \frac{\Pr(\rho|\Gamma, \Delta)}{\sum_{\rho' \in \Phi} \Pr(\rho'|\Gamma, \Delta)}$, and $r_1, r_2 \geq 1$ specify the relative importance between payoff and time.
The ad hoc coordination problem is to optimise

- flexibility $F(\alpha|\Gamma, \mathcal{D})$
- efficiency $E(\alpha|\Gamma, \mathcal{D})$

subject to the constraint that

- $\alpha$ does not know the type spaces $\Theta_j$
- and, therefore, the type distribution $\Delta$ used in $\Gamma$
Harsanyi-Bellman Ad Hoc Coordination
Bayesian Nash Equilibrium

- Problem of incomplete information solved in Bayesian games by assuming that type distribution $\Delta$ is common knowledge

A **Bayesian Nash equilibrium** (BNE) in state $s^t$ is a strategy profile $(\pi_1, ..., \pi_n)$ in which, for all $i \in N$ and $\theta_i \in \Theta_i$, $\pi_i$ maximises

$$
\sum_{\hat{\theta}_-i \in \Theta_-i} \Delta(t, \hat{\theta}_-i | \theta_i) \sum_{a \in A} u_i(s^t, a, \theta_i) \pi(H^t, a, (\theta_i, \hat{\theta}_-i))
$$

where

$$
\Delta(t, \theta_-i | \theta_i) = \frac{\Delta(t, (\theta_i, \theta_-i))}{\sum_{\hat{\theta}_-i} \Delta(t, (\theta_i, \hat{\theta}_-i))}
$$

$$
\pi(H^t, a, \theta) = \prod_{k \in N} \pi_k(H^t, a_k, \theta_k)
$$
Posterior

- However, in ad hoc coordination the type distribution $\Delta$ is unknown.

We can substitute $\Delta(t, \hat{\theta}_-|\theta_i)$ for the posterior

$$
Pr(\theta_{-i}|H^t) = \prod_{j \neq i} Pr(\theta_j|H^t)
$$

where

$$
Pr(\theta_j|H^t) = \frac{L(H^t|\theta_j) P(\theta_j)}{\sum_{\hat{\theta}_j \in \Theta_j} L(H^t|\hat{\theta}_j) P(\hat{\theta}_j)}
$$

$$
L(H^t|\theta_j) = \prod_{\tau=0}^{t-1} \pi_j(H^\tau, a^\tau_j, \theta_j)
$$
Is NE solution suitable?

- Process known to converge to NE (Kalai and Lehrer, 1993)

- But:
  - There may be multiple equilibria
  - Strategies not specified for off-equilibriums paths
  - If posteriors unequal, may not be NE (Dekel et al., 2004)

- Most importantly: Ad hoc coordination means that other agents may not reason like this

- Still, can use as best-response method
User-defined types

Problem:
- Ad hoc coordination means we don’t know the type spaces $\Theta_i$.

Idea:
- We can instead hypothesise type spaces $\Theta_i^*$.
- Each $\theta_i^* \in \Theta_i^*$ is an hypothesis of how agent $i$ might behave.
- Therefore, we call $\theta_i^*$ a user-defined type.
Harsanyi-Bellman Ad Hoc Coordination

Harsanyi-Bellman Ad Hoc Coordination (HBA) is defined as

\[ a_i^t \sim \arg \max_{a_i \in A_i} E_{st}^{a_i}(H^t) \]

where

\[ E_{st}^{a_i}(\hat{H}) = \sum_{\theta^*_i \in \Theta^*_i} \Pr(\theta^*_i | H^t) \sum_{a_{-i} \in A_{-i}} Q_{s}^{a_{i,-i}}(\hat{H}) \prod_{j \neq i} \pi_j(\hat{H}, a_j, \theta^*_j) \]

with

\[ Q_{s}^{a}(\hat{H}) = \sum_{s' \in S} T(s, a, s') \left[ u_i(s, a, \alpha) + \gamma \max_{a_i} E_{s'}^{a_i} (\langle \hat{H}, a, s' \rangle) \right] \]
Properties

Property:
Let $\Gamma$ be a SBG with static pure type distribution $\Delta$. If all players $i \in N$ are controlled by an HBA agent $\alpha_i$ with user-defined types $\Theta^*_i$, and if $\Theta_j \subseteq \Theta^*_i$ for all $i, j \in N$ with $i \neq j$, then play will converge to NE.

Property:
Let $\Gamma$ be a SBG with static pure type distribution $\Delta$ where $\alpha$ controls player $i$, and let $\Theta^D$ be the class of deterministic learners. If, for all $j \neq i$, $\Theta_j \subseteq \Theta^D$ and $\Theta_j \subseteq \Theta^*_j$, then $\alpha$ will be optimally efficient.
Temporally Reweighted Posteriors

Product posterior eliminates wrong/inaccurate types

- What if types change over time?
- What if user-defined types are best approximation we have?

A temporally reweighted posterior redefines

$$L(H^t | \theta_j) = \sum_{\tau=0}^{t-1} f(t - \tau) \pi_j(H^\tau, a_j^{\tau}, \theta_j)$$

with time weight $f$ where $f(\xi) \geq 0$ and $f(\xi) \geq f(\xi + 1)$ for all $\xi \in \mathbb{N}^+$. 

General time weight:

$$f(\xi) = \max(0, a-b(\xi-1)^c)$$
Conceptual Types

Can include methods for opponent modelling in type spaces

- Conceptual types use hypothesised **world conceptualisations**

A conceptual type $\theta_j^c$ for player $j$ consists of a symmetric distance function $d_j : S \times S \rightarrow \mathbb{R}_0^+$, a radius $r \in \mathbb{R}_+$, and a time weight $f$, with

$$
\pi_j(H^t, a_j, \theta_j^c) = \begin{cases} 
|A_j|^{-1} \text{ if } \exists \tau < t : g(s^t, s^\tau) > 0 \\
\eta \sum_{a^\tau \in H^t : a^\tau_j = a_j} f(t - \tau) g(s^t, s^\tau) 
\end{cases}
$$

where $g(s_1, s_2) = \max(0, 1 - d_j(s_1, s_2) r^{-1})$ and $
\eta$ is a normalisation constant s.t. $\sum_{a_j} \pi_j(H^t, a_j, \theta_j^c) = 1.$
Simulated Experiments

Level-based foraging domain:

- Grid domain with $n$ players (circles) and $m$ foods (squares)
- Each player and food has a level
- Players can load food if sum of their levels $\geq$ food’s level
- Payoffs: When loading a food, the food’s level; else -1

Goal: Collect all foods in minimal time, maximise own payoff
Results

▶ a–c: 8 × 8 grid, 2 players, 5 foods
▶ d: 10 × 10 grid, 3 players, 8 foods
▶ Gtw, Unl, Lim = Different HBA configurations
▶ Cor = HBA with correct types
▶ Human = Best human performance
Human-Machine Experiment

At **Royal Society Summer Science Exhibition 2012** in London

- Experiment with a total of **427 participants**
- Lowest age: 9; Highest age: 72; Average age: 17
- Excellent test environment since visitors vary widely
### Experimental Setup

#### Prisoner’s Dilemma

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>3,3</td>
<td>0,5</td>
</tr>
<tr>
<td>D</td>
<td>5,0</td>
<td>1,1</td>
</tr>
</tbody>
</table>

#### Rock-Paper-Scissors

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>P</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>0,0</td>
<td>-1,1</td>
<td>1,-1</td>
</tr>
<tr>
<td>P</td>
<td>1,-1</td>
<td>0,0</td>
<td>-1,1</td>
</tr>
<tr>
<td>S</td>
<td>-1,1</td>
<td>1,-1</td>
<td>0,0</td>
</tr>
</tbody>
</table>
Results

Participants played 2 matches of one game:

- One match against HBA, one against standard solution
- Each match lasted for 20 rounds of one-shot game
- Humans did not know against who they played

Results presented this week at AAMAS 2013 in St. Paul
Thank you
Further Reading

Detailed descriptions of this work can be found in:


Further Reading

References (in order of appearance):

- R. Das, J. Hanson, J. Kephart, G. Tesauro. **Agent-human interactions in the continuous double auction.** International Joint Conferences on Artificial Intelligence, 2001.


Further Reading

- E. Kim, Y. Chang, R. Maheswaran, Y. Ning, L. Chi. **Agent adaptation across non-ideal markets and societies.** Joint Workshop on Trading Agent Design and Analysis (TADA) and Agent-Mediated Electronic Commerce (AMEC), 2012.

