Performance Evaluation Methods for the Trading Agent Competition *

Brett Borghetti and **Eric Sodomka**
Dept of Computer Science and Engineering, University of Minnesota, 200 Union St SE, Minneapolis, MN 55455
borg@cs.umn.edu and sodo0002@umn.edu
http://www.cs.umn.edu/~borg

Abstract

This paper proposes a novel method to characterize the performance of autonomous agents in the Trading Agent Competition for Supply Chain Management (TAC-SCM). We create benchmarking tools that manipulate market environments to control the conditions and provide guidelines to test trading agents. Using these tools, we show how developers can inspect their agents and unveil behaviors that might otherwise have gone undiscovered.

Introduction

One of the most prominent proving grounds for current research in autonomous trading agents is the Trading Agent Competition for Supply Chain Management (TAC-SCM). In this yearly international event, autonomous agents battle for supremacy in a simulation where the highest profit-earning agent wins. In TAC-SCM, agents make all the decisions to run a virtual computer-manufacturing operation. They negotiate to purchase parts from suppliers, optimize their assembly lines, and sell by reverse auction their products to customers. They manage parts and product inventories, minimize costs, and attempt to maximize profit.

Agents compete against five other adversaries in each game. Different combinations of competitors, in addition to the randomness in the game, cause different market conditions to arise. An agent must perform well under many different market conditions to succeed.

To evaluate an agent's performance under various market conditions, we developed a pair of benchmark agents that can manipulate market conditions to simulate various levels of supply and demand in the market. Any team wishing to evaluate their agent can use these stand-alone market manipulator agents to create a configureable level of pressure in the marketplace. The benchmark agents do not require alteration of the game server or the agent being examined. They control the supply and demand characteristics of a game by buying parts and selling computers at prices that generate the desired demand and supply in the marketplace.

*We would like to thank Professor Maria Gini for all her help and support with this effort. Eric Sodomka is an undergraduate student.

Related Work

TAC-SCM allows research teams to develop trading agents and compare their relative performance in a complex, standardized environment (Collins *et al.* 2005). Several teams developed methods of analyzing agent performance. The University of Southampton team examined running variations of their competitive agent with different risk strategies for pricing (Minghua He & Jennings 2005) to show that their competitive agent made the highest profit among the variations. The University of Michigan team analyzed post-competition performance of the TAC-SCM winning agents and explored relationships between total profit and other measurements of performance (Kiekintveld, Vorobeychik, & Wellman 2005). The team at the University of Texas at Austin focused on comparing multiple types of learning agents competing against each other (Pardoe & Stone 2004). While each of these teams characterized agent's performance against other agents, no one developed a stable, universal benchmarking environment in which to characterize agent performance. We felt in order for the field to advance, we must explore this region of agent analysis.

Approach

There are two key challenges in designing a useful benchmarking environment for TAC. The first is how to manipulate the market conditions for observing agent behavior. The second is to manage the randomness and repeatability of the games. While we developed an experimental system to address both challenges, in this paper we focus on the first.

To manipulate the market conditions for benchmarking performance, we developed two new agents, the *Market Relief* or "do nothing" agent, and the *Market Pressure* agent. The Market Relief Agent occupies one or more of the 6 slots in a TAC simulation without making any financial transactions. This agent reduces demand on the suppliers which leads to a lowering of supplier prices; and it reduces available supply for the customers which leads to an increase in what customers were willing to pay for computers from other agents.

Conversely, the configurable Market Pressure Agent does the opposite: it increases available supply to customers, allowing them to pay less for computers while simultaneously putting more demand on suppliers, encouraging them to increase their prices. Our Market Pressure Agent operates by continually adjusting its customer offer prices to achieve a desired market share. Since it is a build-to-order agent it purchases parts to build the computers ordered and the market share achieved on the customer side is reflected on the supplier side. This agent can capture any market share from 0% to 100% because it has an unlimited line of credit and has no concern for its own profit earning capability. When our agent captures the desired market share, no other agent(s) can use that portion of the market.

Using combinations of Market Relief and Market Pressure Agents, developers can control the market environment. When a developer makes alterations to a competitive agent they wish to observe, they can use these tools to benchmark and compare the change in performance of their alteration.

Experiments

We examine what happens to an agent when we alter product supply and demand for parts in the marketplace. As shown in Figure 1, our MinneTAC (Ketter *et al.* 2004) agent earns a relatively constant profit (shallow slope) until the Market Pressure Agent absorbs significant market share near the *Market Saturation Threshold*: 83% (5/6) of the available market. Market Saturation represents the point at which there is little unmet customer demand remaining: if one agent wishes to sell more computers and gain market share from another agent, price wars occur. Notice the profit slope changes abruptly around 70% market pressure. The point at which the slope of agent profit changes in this region is relevant: a better performing agent should have a shallower slope through a higher market pressure. Thus, if we make adjustments to our agent, we should strive to find adjustments that increase the total area under this curve. Additionally, changes which reduce the slope while maximizing the area under the curve should lead to agents that perform well in a wide variety of market conditions.

Notice the point at which our agent is forced to lose money: when the market pressure is 92%. At this point our agent would be better off scaling back on any financial transactions: since material costs exceed revenue, any further transactions lead to a deeper negative profit.

We are likely to see similar high market pressures exhibited by competitors in actual competitions. Using market pressure tools to discover and guide the correction of undesirable agent behavior can help us develop an agent that performs better in the competition.

Conclusion and Future Work

We developed a framework of tools to control the simulation environment and measure the performance of trading agents under various market conditions. With these methods in place, research teams are poised to better understand how individual changes to their agents affect overall performance. Additionally, teams can now evaluate their agent in a variety of market environments that were previously unable to be simulated. We believe that agents optimized with these methods will perform better in actual competition.

Although these tools provide a guide for the developer wishing to optimize an agent, there is room for improvement of the tools. One problem is that when a new customer order arrives, the Market Pressure Agent purchases parts immediately in high quantities with an as-soon-as-possible delivery date. But because other agents compete for parts in 3 dimensions (quantity, price, and delivery date), and the Market Pressure Agent only operates in the quantity and price dimensions, the Market Pressure Agent is unable to create pressure against future purchases of parts the other agents make. As a result, our agent tightly controls market share in the product market, but it only loosely influences the parts market. Resolving this problem is a topic for future research.

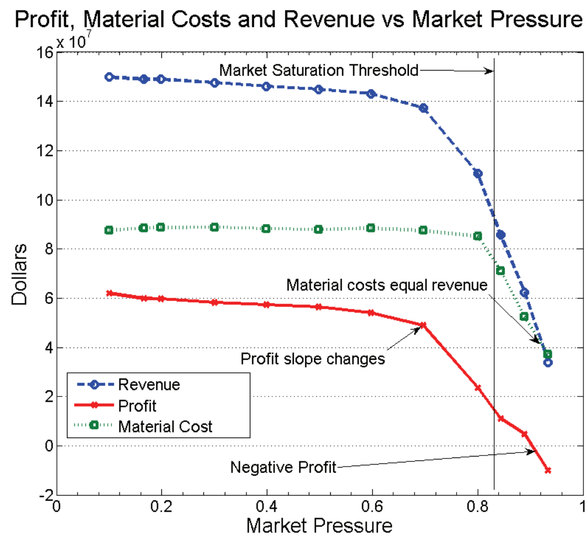

Figure 1: Market Pressure Effects on MinneTAC

References

Collins, J.; Arunachalam, R.; Sadeh, N.; Ericsson, J.; Finne, N.; and Janson, S. 2005. The supply chain management game for the 2006 trading agent competition. Technical Report CMU-ISRI-05-132, Carnegie Mellon University, Pittsburgh, PA.

Ketter, W.; Kryzhnyaya, E.; Damer, S.; McMillen, C.; Agovic, A.; Collins, J.; and Gini, M. 2004. Analysis and design of supply-driven strategies in TAC-SCM. In *AAMAS-04 Workshop on Trading Agent Design and Analysis*, 44–51.

Kiekintveld, C.; Vorobeychik, Y.; and Wellman, M. P. 2005. An analysis of the 2004 supply chain management trading agent competition. In *IJCAI-05 Workshop on Trading Agent Design and Analysis TADA-05*.

Minghua He, Alex Rogers, D. E., and Jennings, N. R. 2005. Designing and evaluating an adaptive trading agent for supply chain management applications. In *IJCAI-05 Workshop on Trading Agent Design and Analysis TADA-05*.

Pardoe, D., and Stone, P. 2004. Bidding for customer orders in TAC SCM. In *AAMAS 2004 Workshop on Agent Mediated Electronic Commerce VI*.