Empirical Game-Theoretic Analysis of the TAC Supply Chain Game

Patrick R. Jordan, Christopher Kiekintveld, and Michael P. Wellman

University of Michigan
Computer Science & Engineering
Ann Arbor, MI 48109-2121 USA
{ prjordan, ckiekint, wellman }@umich.edu

ABSTRACT

The TAC Supply Chain Management (TAC/SCM) game presents a challenging dynamic environment for autonomous decision-making in a salient application domain. Strategic interactions complicate the analysis of games such as TAC/SCM, since the effectiveness of a given strategy depends on the strategies played by other agents on the supply chain. The TAC tournament generates results from one particular path of combinations, and success in the tournament is rightly regarded as evidence for agent quality. Such results along with post-competition controlled experiments provide useful evaluations of novel techniques employed in the game. We argue that a broader game-theoretic analysis framework can provide a firmer foundation for choice of experimental contexts. Exploiting a repository of agents from the 2005 and 2006 TAC/SCM tournaments, we demonstrate an empirical game-theoretic methodology based on extensive simulation and careful measurement. Our analysis of agents from TAC-05 reveals interesting interactions not seen in the tournament. Extending the analysis to TAC-06 enables us to measure progress from year-to-year, and generates a candidate empirical equilibrium among the best known strategies. We use this equilibrium as a stable background population for comparing relative performance of the 2006 agents, yielding insights complementing the tournament results.

General Terms

Economics, Experimentation, Measurement

Keywords

Multiagent systems, Empirical game theory, Electronic markets, Trading agent competition, Supply chain management

1. INTRODUCTION

Many if not most multiagent systems (MAS) research projects that produce new strategies for agent behavior evaluate their proposals through some experimental regimen comprising simulation of alternative behaviors in a chosen environment. The typical objective of such experimentation is to establish that the proposed strategy possesses some advantageous characteristic(s) compared to alternatives in a given setting, or to develop a model of performance as a function of environment features. Methodologies employed in experimental analyses are various, and dependent of course on the issues at hand. A key issue distinguishing MAS settings from single-agent applications of computational experiments is that the effectiveness of an agent's strategy depends pivotally on the strategies employed by other agents. Determining the configurations of agent behaviors to simulate is therefore a crucial issue in MAS experimental design.

Although it appears that much MAS research pursues this determination in an ad hoc manner, the issue is often recognized, and several approaches address it directly. In a *factorial* design, the combinations of agent strategies are simulated exhaustively. This is infeasible when there are large numbers of possible strategies, or a large population of agents. Even when feasible, in interpreting the experiments the analyst must render judgments about the degree to which the various configurations are relevant in order to draw conclusions about proposed strategies.

One appealing way to determine a relevant set of agent strategies is to generate a population iteratively through some evolutionary process. The evolutionary approach was pioneered in computational agent research by Axelrod's famous Prisoner's Dilemma tournament [2], and has become a standard method among researchers in agent-based computational economics [17]. Evolutionary search techniques provide (at least) two useful functions in MAS experimentation:

1. *Generating strategies* for exploration given a set of primitive building blocks, employing stochastic local search from an initial population. Techniques for generation are typically based on genetic algorithms or genetic programming [10].
2. *Finding stable populations* of strategies, for example using replicator dynamics [16].

Of course, there are alternative means as well to support both of these functions. Any structured search technique (employing genetic operators or not) is a candidate method for exploring a space of available strategies. And evolutionary stability is just one criterion that might be employed to evaluate the plausibility of populations. It is uniquely compelling only to the extent that the evolutionary dynamic employed is itself a plausible model of how agent strategies might be adopted over time.

Game theory is another source of stability criteria often employed in MAS research. Although evolutionary and game-theoretic stability (i.e., equilibrium) concepts sometimes coincide [5], this is

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
AAMAS'07 May 14–18 2007, Honolulu, Hawai'i, USA.
Copyright 2007 IFAAMAS .

978-81-904262-7-5 (RPS) ©2007 IFAAMAS

not always the case. Game theory tends to avoid assuming any particular dynamic model, which may be viewed as a strength or weakness depending on one's perspective and the particular issues at hand. What game theory does provide is a rigorous mathematical framework for formalizing interactions among rational agents, and a rich set of solution concepts and other formal properties useful for characterizing alternative strategic configurations.

Whereas game theory is now quite commonly employed by MAS researchers in theoretical investigations, it is less frequently applied in experimental studies.[1] The approach we pursue here, called *empirical game-theoretic analysis*, employs an experimental methodology for explicit game-theoretic treatment of MAS simulation studies.

2. TAC SUPPLY CHAIN MANAGEMENT

The annual Trading Agent Competition (TAC) series of international research tournaments was initiated to promote research and education in the technology underlying trading agents.[2] At the core of TAC are two *games*, market-based scenarios where multiple agents compete to exchange goods and services at dynamically negotiated prices. The first TAC tournament, in July 2000, introduced the TAC Travel game [22]. A second game, in the domain of supply chain management (TAC/SCM), has been played since 2003 [1, 4].

A key feature of both games is that—like most realistic market environments—they are sufficiently complicated (severely imperfect and incomplete information revealed over time throughout dynamic activity) to defy analytic solution. Thus, empirical methods appear indispensable to progress.

Our current study focuses on TAC/SCM. Although most details of the game rules are inessential to the analysis here, we establish some context by providing a capsule description. A complete specification of the game [3] is available at the web sites referenced, and further description and discussion is provided in many of the papers cited herein.

In TAC/SCM, six agents representing PC (personal computer) manufacturer agents compete to maximize their profits over a simulated year. There are 220 scenario days, and agents have approximately 14 seconds to make decisions each day. Agents participate simultaneously in markets for supplies (components) and finished PCs. There are 16 different types of PCs (divided into three market segments), defined by the compatible combinations of 10 different component types. Components fall into one of four categories: CPU, motherboard, memory, and hard disk. There are four types of CPUs and two types of all other components; one component from each category is required to produce a PC.

Agents negotiate deals with suppliers and customers through an *RFQ* (*request-for-quote*) mechanism. The suppliers and customers execute policies defined by the game specification and implemented in the server. The suppliers have limited production capacity that varies during the game according to a random walk. They make offers and set prices based on their ratio of available capacity. The customer generates requests for PCs each day. The number of requests is driven by a stochastic demand process for each market segment.

Agents face substantial uncertainty in both markets. The underlying supplier capacities, customer demand parameters, and local state of other manufacturer agents are not directly observable, so agents must estimate these from other sources of information. There is also strategic uncertainty, since agents do not know the exact strategies employed by their competitors.

Each manufacturer is endowed with an identical factory that has limited production capacity, measured in *cycles*. Each PC type requires a different number of cycles to produce. Agents pay storage costs for all components and PCs held in inventory each day, and are charged (or paid) interest on bank balances. At the end of the game agents are evaluated based on total profit, and any remaining inventory is worthless.

3. EMPIRICAL ANALYSIS METHOD

A previous application of empirical game-theoretic analysis to the TAC/SCM domain considered the issue of strategic procurement of components at the beginning ("day 0") of the simulated manufacturing period [20]. That study investigated a phenomenon observed in the 2003 tournament, employing strategies defined by varying one aspect of the University of Michigan's agent, Deep Maize. Controlled experiments varying only the degree to which agents procure components on day 0 verified that the aggressive procurement policies observed (informally) in tournament play actually represents an equilibrium of sorts—and one that is mutually destructive to manufacturing profits. The analysis further confirmed that Deep Maize's *preemptive* strategy of blocking day-0 procurement neutralized this issue, forming a new equilibrium where all agents (not just the preempting Deep Maize) were more profitable.[3]

The force of day-0 procurement in the game was considered a design flaw by the TAC/SCM community, and revisions of the game in 2004 and 2005 attempted to attenuate its influence. The 2004 redesign was unsuccessful from this perspective [8], and empirical game-theoretic analysis demonstrates that no reasonable settings of the focal storage-cost parameter would have likely been sufficient [18]. The 2005 redesign [3] included deeper changes to supplier behavior, and appears to have dramatically lessened the salience of day-0 procurement issues.

Most published research on TAC/SCM agents presents evidence from tournaments, as well as controlled experiments with variants on the agent strategy under study. These experiments typically include simulations where some subset of the agents play such variants, and the remainder play some fixed or background strategies. What strategies to assume for the background agents is a key experimental design choice. One option—employed, for example, in a recent study on SouthamptonSCM [6]—is to use the "dummy" agents provided along with the TAC/SCM game server. Another is to use agent strategies developed by other TAC/SCM participants. This has been greatly facilitated by the introduction of a TAC agent repository following TAC-05.[4] For example, Pardoe & Stone [11] run simulated games, each with two variants of their agent TacTex playing with a fixed background of four agents drawn from the repository (Mertacor, MinneTAC, GoBlueOval, and RationalSCM).

[1]Perhaps it is starting to emerge. Although we do not attempt here to identify the earliest sources (see [14, Section 3.9] for a survey), we do acknowledge that many MAS works have included elements of game-theoretic perspective in experimental studies. One recent thread along these lines is represented prominently by the work of [13]. Our claimed contribution is to systematizing and enriching the methodology, not completely originating it.

[2]See http://tradingagents.org, and http://www.sics.se/tac.

[3]Here we refer to the 2003 version of Deep Maize. Subsequent discussion applies to the 2005 or 2006 versions of Deep Maize and all other agents mentioned, as indicated explicitly or by context.

[4]Designed and implemented by Joakim Eriksson (Swedish Institute of Computer Science) and Kevin O'Malley (University of Michigan), and available at http://www.sics.se/tac/showagents.php.

Playing with real tournament agents lends realism to the simulations, but the question still remains as to what mixtures of background strategies are most relevant. This is where empirical game-theoretic analysis can provide some guidance. Our premise is that—all else equal—profiles of strategies that are more strategically stable (i.e., closest to game-theoretic equilibrium) are more plausible as background contexts. Of course, this is at best a starting point, as introduction of a new strategy may alter the strategic landscape. Therefore, one must update the analysis to reflect any promising new strategies identified during experimentation.

We have undertaken an empirical game-theoretic analysis of agent strategies from TAC/SCM 2005 and 2006. Our approach comprises the following steps, which we elaborate in turn:

1. Approximate the six-player SCM game by its three-player reduced version, SCM$\downarrow_3$ [21].
2. Run many simulations covering all distinct strategy profiles.
3. Process the simulation data by checking game validity and adjusting for stochastic demand variability.
4. Analyze the resulting empirical game by searching for equilibria and approximate equilibria.

3.1 Reducing the Game

Given a symmetric game with N players and S strategies, there are $\binom{N+S-1}{N}$ distinct pure-strategy profiles. For TAC/SCM, $N = 6$, and in our current analysis we consider $S = 6$ agent strategies. This induces a total of 462 profiles that would need to be estimated for a full-granularity analysis. We can significantly decrease this number by restricting attention to cases where strategies are assigned to *pairs* of agents rather than individuals. Specifically, the resulting 3-player game, denoted SCM$\downarrow_3$, comprises only 56 profiles over the same 6-strategy set. The payoff to a strategy in an SCM$\downarrow_3$ profile is defined as the *average* payoff to the two agents playing this strategy in the original 6-player game.

In several contexts, it has been shown experimentally and theoretically that this form of *hierarchical game reduction* produces results approximating well the original unreduced game, with great computational savings [14, 21]. Although we have not validated this specifically in TAC/SCM, intuitively we would expect this game to share the necessary property of payoffs smoothly varying with the number of other agents playing a given strategy.

3.2 Running Simulations

We have collected results from well over 12,000 sample games combined in the TAC-05 and TAC-06 environments. We performed most of our simulations using a computing cluster operated by the University of Michigan. The cluster facility provides scalable and homogeneous processing, supporting parallel simulation with a fair allocation of computational power to each agent.

Each game simulation reserves seven CPUs for a period of one hour; one for each agent and one for the TAC SCM server. We group games into sets of 3–5 to reduce the overhead cost of configuring the simulation on the assigned cluster nodes. Game results (in the form of server log files) are sent back to central repository. A central server tracks the results and submits new simulations to the cluster as each job is completed.

The simulated strategies potentially differ from the actual tournament agents in one important respect. Tournament agents can maintain state from one game instance to another, and so can *adapt* their strategy for later games based on experience in earlier games. Several agents take advantage of this opportunity, including the top-scoring agent from both 2005 and 2006 tournaments, TacTex [12]. Our simulation analysis is based on sampling from a pool of fixed strategies, so necessarily restricts the agents to versions that adapt only *within* a game instance.

3.3 Post-Processing Simulation Results

During a game simulation, much can go wrong, for example network outages or delays, or interference with one or more processors. We therefore attempt to filter our data set by removing game instances tainted in this way. We considered various procedures for identifying tainted games, ultimately settling on a very simple rule. A game is scratched if, for any agent, there are six or more days (out of 220) in which the server did not receive a message from that agent (as indicated by the game log).

Given the expense of generating samples by simulation (over 7 processor-hours per game), we seek to glean the most information we can from each data point. Toward that end, we employ statistical techniques to reduce variance. In particular, the method of *control variates* [15] improves the estimate of the mean of a random function by exploiting correlation with observable random variables. In the case of TAC/SCM, the most significant stochastic factor bearing on payoffs is the level of customer demand for PCs during the game.

As in the analysis of TAC/SCM-03 by Wellman et al. [20], we use control variates to derive a payoff measure called *demand-adjusted profit* (DAP). This adjustment considers the average level of demand (measured in total number of PCs requested) for each of the PC market segments: *low*, *mid*, and *high*.[5] We collected the demand and score data from games played in the TAC/SCM tournaments: quarter-final, semi-final, and final rounds. For TAC-05, the overall tournament comprised 96 games, of which 71 remained after applying the tainted-game filter described above. Table 1 presents summary demand statistics for TAC-05 and TAC-06 games.

	Segment	Mean ($\bar{Q}_{\text{seg}}$)	DAP Coeff. (δ_{seg})	Low (2.5%)	High (97.5%)
2005	Low	132,498	69.67	31.91	139.23
	Mid	157,481	63.80	28.98	98.62
	High	129,641	85.57	29.08	110.25
2006	Low	127,899	52.82	-1.47	107.11
	Mid	151,590	82.94	41.82	124.07
	High	130,372	132.00	85.59	178.32

Table 1: TAC/SCM tournament demand statistics with 95% confidence intervals around the DAP coefficients.

The center column of the table (DAP coefficient) presents the result of a linear regression of mean agent score on demand in the respective segments. The DAP regression R^2 statistics for TAC-05 and TAC-06 are 0.3469 and 0.4594, respectively, with p-values of 2.5e-6 and 3.1e-8. We obtain the DAP for agent i in game x by subtracting from its actual profit an adjustment based on the demand in that game.

$$\text{DAP}_i(x) = \text{Profit}_i(x) - \sum_{\text{seg}\in\{\text{low,mid,high}\}} \delta_{\text{seg}}(Q_{\text{seg}}(x) - \bar{Q}_{\text{seg}}), \tag{1}$$

where $Q_{\text{seg}}(x)$ denotes the actual demand for the specified segment in game x, and $\bar{Q}_{\text{seg}}$ the mean demand as presented in Table 1.

[5] In the original TAC/SCM-03 rules, one stochastic process governed demand for all PC types. Thus, the adjustment formula was necessarily revised from earlier analysis.

4. SCM$\downarrow_3$ ANALYSES

We apply our empirical analysis methodology to agents released as binaries after both the 2005 and 2006 TAC SCM competitions. We first discuss the 2005 and 2006 game results individually, and then consider the combined set of strategies to make comparisons across tournaments.

4.1 SCM 2005

Our TAC-05 analysis employs a dataset of 2110 validated game instances, covering a minimum of 28 samples each of 56 distinct profiles of six TAC-05 agents: TacTex (Tx) [11, 12], Mertacor (Mr) [9], Deep Maize (Dm) [7], MinneTAC (Mt), PhantAgent (Ph), and GoBlueOval (Gb).[6] As shown in Table 2, the first four of these made it to the final round of the TAC/SCM-05 tournament (the other two finalists are not currently available in the repository), PhantAgent was a semi-finalist, and GoBlueOval was a quarter-finalist.

Interaction among the strategies is one factor explaining differences in scores—and even relative rankings—between rounds of the tournament. Game-theoretic analysis serves to assess the robustness of tournament rankings to strategic interactions. Another factor that explains differences between rounds is modifications to agents made between rounds (developers are allowed to modify agents between tournament rounds, but not within a round). In one case both a semi-final and final round version of a single agent has been released, but we typically do not have access to all versions of the agent and are thus unable to investigate these variations in our empirical analysis.

Figure 1 summarizes our stability analysis of the pure strategy profiles of the game. Each node represents a profile (three strategies). The outgoing edge from a node indicates the *best deviation* from that profile—that is, the transition providing the greatest gain in payoff for one agent switching strategies. For example, the profile with all Deep Maize (DmDmDm, in Level 4 around 10 o'clock) points to profile DmDmMr, which means that switching from Deep Maize to Mertacor in this context offers the greatest increase in payoff. That the arrow signifying the edge is solid rather than dashed means that the benefit is statistically significant in this case, at the $p \leq 0.05$ level.

The magnitude of the potential benefit from deviating is represented by the node's placement in the diagram. We denote this quantity by ϵ, since a profile with maximal benefit to deviation of ϵ constitutes an ϵ-Nash equilibrium. The profiles in the innermost ellipse (Level 1) represent the most stable (closest to equilibrium), with $0.04M \leq \epsilon \leq 0.6M$. Concentric rings define levels with increasing values of ϵ. Level 4 (outermost ring) profiles are quite unstable, as a single agent (in the 3-player game) can benefit by at least 4.4M by deviating from its designated strategy. Note that the best deviation links usually, but not necessarily, connect profiles to more stable alternatives.

Since all profiles in Figure 1 have outgoing edges, we can conclude that the empirical game has no pure-strategy Nash equilibria (PSNE). Indeed, there exists a directed cycle among three relatively stable profiles, and all paths lead to this cycle.

There are, however, mixed strategy equilibria, and we have identified one symmetric Nash equilibrium, as well as several approximate equilibria. We found these mixtures using replicator dynamics (RD), and present them in Table 3. Specifically, we ran RD seven times: once with all strategies present, and once for each subset of five out of six. In all cases the initial population is distributed uniformly. The profile generated by RD with all agents present is a symmetric Nash equilibrium. GoBlueOval is not played in this equilibrium, and indeed omitting that agent leaves the RD result unchanged. Two other RD results are approximate ($\epsilon < 1.0M$) equilibria; not surprisingly these respectively omit the agents (Deep Maize and Minnetac) with lowest positive probability in the known exact equilibrium.

Our analysis reveals several striking observations. First, all agents perform quite poorly with many copies of themselves. Three out of the four most unstable profiles (MnMnMn, TxTxGb, TxTxTx, and MrMrMr, respectively) comprise a single strategy. This fact can be explained by the multiple copies all competing for the same "niche", or exploiting opportunities typically left available by other agents (but not themselves, of course). In addition, some of the problem may be simply that the agents are hardwired to procure components on certain days or with certain lead times, and these naturally interfere when more than one copy exists. Similarly, multiple copies may make the same predictions and estimates of prices and other market conditions, so may be making bidding and other decisions in an interfering manner.

Second, PhantAgent performs much better in the game-theoretic sense than might be expected from the TAC/SCM-05 tournament outcome.[7] PhantAgent is least sensitive to playing with copies of itself, and appears with substantial probability in all the profiles produced by RD in Table 3. In fact it is most probable in all but two cases: the one where it was excluded, and the one with highest ϵ value.

Third, Mertacor appears especially strong in a wide variety of contexts. Like PhantAgent, Mertacor is present with large probability in all the symmetric stable profiles identified. Most remarkable is that of the 35 profiles without Mertacor, 30 of them have a best deviation where some strategy changes to Mertacor. Of the 21 profiles with Mertacor, the best deviation changes from Mertacor in only three.

We should note that the first observation above raises some questions about our analysis approach. Presumably TAC entrants design their agents with tournament play in mind, and so may not be concerned about the performance of their agents with copies of themselves in the environment. On the other hand, one might argue that performance in self-play is important, and the tournament unduly neglects this aspect of strategy. Our reduced-game analysis is especially sensitive to this question, since all profiles have at least two copies of any strategy present. We plan to explore this issue further in ongoing development of our methodology.

4.2 SCM 2006

The agents that competed in the 2006 TAC SCM finals are listed in Table 4. Versions of five of these agents were released to the agent repository: TacTex (Tx), PhantAgent (Ph), DeepMaize (Ds and Df), Maxon, and MinneTAC (Mt). Two versions of DeepMaize were released, corresponding to versions that played in the final round and semi-final round (significant changes were made to the agent between rounds, particularly to procurement behavior). The MinneTAC agent is the version that played in the semi-final round. This agent was also changed for the final round, but the final round version has not been released. Both semi-final and final round versions of Maxon were released. We analyze five of the seven agents available, including both versions of Deep Maize but excluding both Maxon agents[8]. The full symmetric game for the

[6] The versions of these agents in the repository do not maintain state from game to game, so may differ from the actual tournament agents as noted above.

[7] We are unaware of specific problems that may have afflicted the agent in the semi-final round, but this is a possibility.

[8] Maxon was the last agent to be released, and we do not have enough simulation data for these agents to be included in our full

Agent	Affiliation	Finals	Semi-Finals	Quarter-Finals	Seeding
TacTex	U Texas	4.74	3.57 [1]	17.78 [A]	14.89
SouthamptonSCM	U Southampton	1.60	4.62 [2]	3.50 [B]	10.05
Mertacor	Aristotle U Thessaloniki	0.55	2.66 [2]	4.58 [B]	9.30
Deep Maize	U Michigan	–0.22	3.68 [1]	17.49 [D]	10.23
MinneTAC	U Minnesota	–0.31	2.27 [1]	11.91 [A]	9.86
Maxon	Xonar Inc.	–1.98	3.80 [2]	5.23 [C]	8.76
PhantAgent	Politechnica U Bucharest	n/a	–6.64 [1]	7.03 [A]	9.87
GoBlueOval	Ford Motor Co. and U Michigan	n/a	n/a	–2.60 [B]	12.60

Table 2: TAC/SCM-05 finalists, plus PhantAgent and GoBlueOval, with average scores ($M) from seeding through final rounds (semi-final and quarter-final groups in brackets).

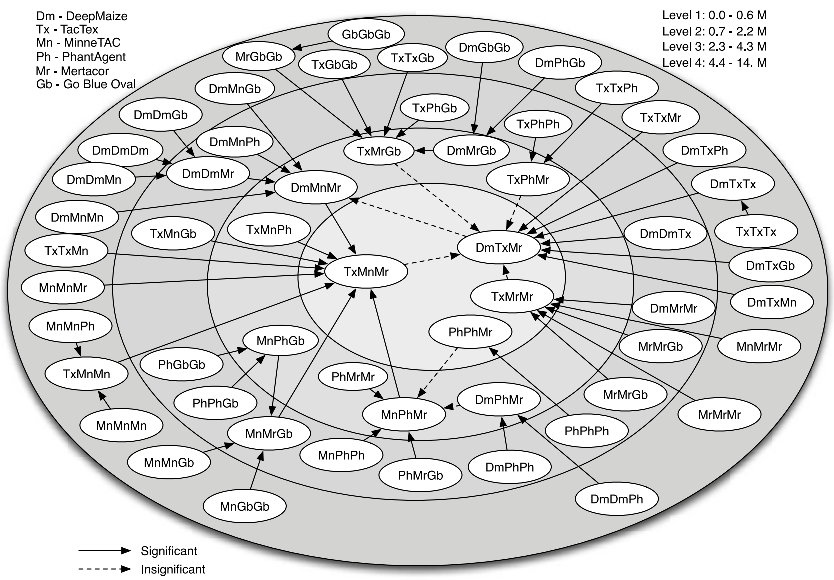

Figure 1: Deviation analysis for pure profiles of 2005 SCM$\downarrow_3$.

five agents we include in our analysis comprises 35 profiles. We have over 1100 validated game instances, with a minimum of 15 samples for each profile (typically 30 or more).

Stability results for the 2006 agent set are shown in Figure 2. This game contains a pure strategy Nash equilibrium (DsPhTx) and an approximate equilibrium (DsDsTx) that has a small, statistically insignificant, benefit of $0.09M$ for deviating to the PSNE. We also applied replicator dynamics to this game to search for symmetric mixed equilibria, starting from mixtures generated uniformly at random. In all cases, RD converged to a mixture of TacTex, PhantAgent, and DeepMaize SF. Figure 3 shows the field for replicator dynamics over the simplex of these three strategies. The fixed point corresponds to the symmetric Nash equilibrium mixture (0.254, 0.188, 0.558).

Table 5 presents several statistics about the deviations in 2006 SCM$\downarrow_3$. *Percent positive deviations* is the fraction of possible deviations to the agent that result in a net benefit. *Best Deviation* is the number of instances where deviating to the agent was the most beneficial deviation. *Mean Deviation* and *std. error* reflect the average benefit ($M) for deviating to this agent, which may be negative.

analysis.

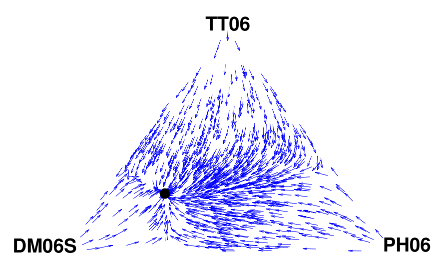

Figure 3: Replicator dynamics field for the top three agent strategies from the 2006 SCM$\downarrow_3$. The Nash equilibrium (0.254, 0.188, 0.558) is shown as a black dot.

Agent	all	(Deep Maize)	(TacTex)	(MinneTAC)	(PhantAgent)	(Mertacor)	(GoBlueOval)
Deep Maize	.055	—	.015	.035	.219	.326	.055
TacTex	.112	.137	—	.100	.210	.156	.112
MinneTAC	.057	.079	.106	—	0	.109	.057
PhantAgent	.400	.418	.533	.482	—	.271	.400
Mertacor	.376	.366	.346	.384	.559	—	.376
GoBlueOval	0	0	0	0	.012	.138	—
ϵ	0	0.49M	1.46M	0.42M	1.28M	3.50M	0

Table 3: Profiles resulting from replicator dynamics. Each column presents probabilities for a mixed profile, with associated ϵ in $\mathrm{SCM}\downarrow_3$ specified in the bottom row. The first column presents the result from RD including all agent strategies (initial proportions uniform). Subsequent columns respectively omit one strategy from the RD process.

Agent	Affiliation	Finals	Semi-Finals	Quarter-Finals	Seeding
TacTex	U Texas	5.85	7.55 [2]	7.48 [B]	13.73
PhantAgent	Politechnica U Bucharest	4.14	5.71 [2]	17.37 [C]	12.56
DeepMaize	U Michigan	3.58	6.46 [1]	9.61 [A]	16.60
Maxon	Xonar Inc.	1.75	4.08 [1]	17.74[D]	10.63
Botticelli	Brown U	0.48	1.94 [1]	0.83 [A]	4.21
MinneTAC	U Minnesota	–2.70	2.06 [2]	13.45 [C]	9.59

Table 4: TAC/SCM-06 finalists, with average scores ($M) from seeding through final rounds (semi-final and quarter-final groups in brackets).

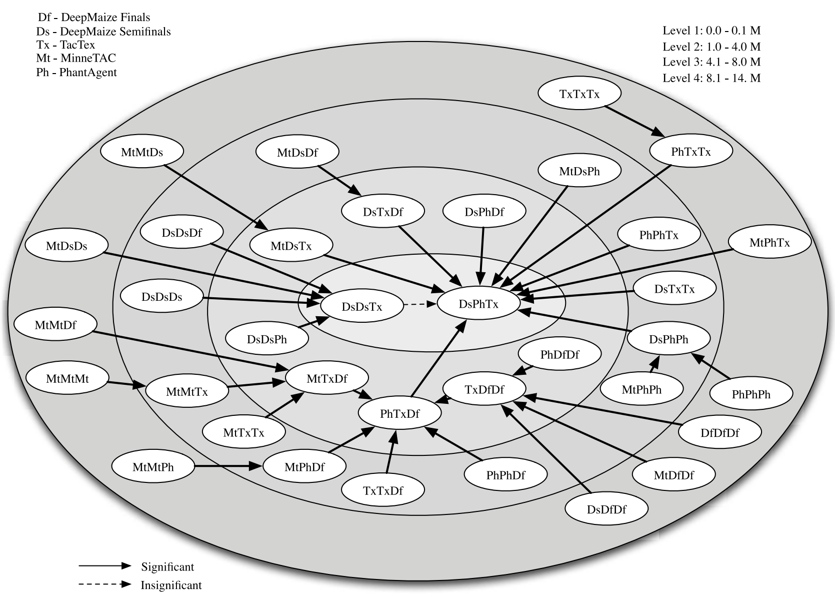

Figure 2: Deviation analysis for pure profiles of 2006 $\mathrm{SCM}\downarrow_3$.

Deviations to TacTex, PhantAgent, and DeepMaize SF are beneficial in a least 60% of the cases. The mean value for deviating is highest for TacTex and DeepMaize SF, and TacTex is the best deviation most frequently. The three agents comprising the PSNE are nearly indistinguishable in this analysis.

Perhaps the most striking result is that the semi-final version of DeepMaize clearly outperforms the finals version in this game-theoretic environment. This is another instance where an agent that did not participate in the finals shows strong performance in our game-theoretic analysis. Given these results, the selection of the weaker version of DeepMaize to play in the final round would have been prevented. In our combined analysis we discovered further evidence that given only the 2005 agent strategies and the two versions of Deep Maize, the semifinals version is a more robust choice. This suggests that the type of analysis we present here should have applications to strategy selection as well as post-

Agent	% Positive Deviations	Best Deviation	Mean Deviation	Std. Error
TacTex	61.67	18	1.45	5.91
Deep Maize SF	63.33	5	1.43	4.68
PhantAgent	60.00	8	0.89	4.77
Deep Maize F	53.33	3	0.88	4.62
MinneTAC	11.67	0	–4.67	6.41

Table 5: Deviation statistics for agent strategies of 2006 SCM$\downarrow_3$.

Background Context		Deviation Gain (ϵ)		
05 Agent	Mixture		Server Rules	
		06 Agent	2005	2006
Deep Maize	0.083	PhantAgent	5.33M	6.57M
Mertacor	0.431	TacTex	5.07M	4.73M
PhantAgent	0.314	Deep Maize SF	4.22M	4.56M
TacTex	0.172			

Table 6: Deviation gain comparison of top 2006 agents in the context of a symmetric mixed Nash equilibrium of top 2005 agents for the 2005 and 2006 server rules.

tournament analysis. Finally, we note that agents in the 2006 game show similar difficulties playing against copies of themselves to the 2005 agents; all of the profiles with six identical agents are among the least stable profiles again. This effect may partially explain why the profile with the top three agents playing is a PSNE. None of the agents has enough of an advantage over the others to overcome the penalty from playing against more copies of itself, so none of the deviations to these three is beneficial.

4.3 Combined Analysis

Sections 4.1 and 4.2 focused on analyzing sets of agents from a single year of competition. One of the exciting opportunities afforded our empirical analysis methodology is to consider combinations of agents not observed during tournament play, including agents from different years. To facilitate general comparisons between agents we introduce an alternative means to rank agents. We select a particular strategic context based on game-theoretic stability and rank agents according to the benefit of deviating to the agent from this context.

For this analysis we employ data from approximately 10000 simulations focused on the profiles containing eleven agents from 2005 and 2006 (listed in Table 7). The experiments using the top 2005 agent sets as background context were simulated using both the 2005 and 2006 server rules. The simulations using the top 2006 agents as background context were run using only the 2006 server rules, which have two modifications from the 2005 rules. Under 2006 rules, the identity of opposing agents is revealed at the start of the game, and the reputation mechanism for suppliers was slightly modified. Agents from 2005 are compatible with the 2006 server, but may be at a disadvantage because they were not designed for the new rule set.

We begin by testing the hypothesis that the strongest agents from 2006 should show substantial improvements over the strongest agents from 2005. Our first step is to select a symmetric Nash equilibrium for the 2005 agent set {DeepMaize, Mertacor, PhantAgent, TacTex}, which corresponds to the support of the mixed strategy equilibrium of the full 2005 game less MinneTAC.[9] Using this 2005 equilibrium as the background context, we test possible deviations to three of the top 2006 agents. The results are given in Table 6, along with the background context. Each of the 2006 agents is a beneficial deviation from the 2005 equilibrium using both the 2005 and 2006 server rules, offering strong support for the hypothesis of improvements from 2005 to 2006. In the subgame that includes the background agents and DeepMaize SF, DeepMaize SF is the sole survivor of iterated elimination of dominated strategies, providing even stronger evidence for improvement in this agent.

In Table 7 we present a ranking of eleven agents from 2005 and 2006 in the context of the symmetric mixed Nash equilibrium given in Figure 3.[10] This equilibrium is robust to the addition of the 2005 agents into the strategy pool. All agents are ranked based on the benefit of deviating to the agent from the equilibrium context. This ranking is interesting particularly because it spans agents that have never faced one another directly in tournament competition. Table 7 also gives the tournament results, where applicable.

Agent	Deviation Gain	Tournament Scores Finals 05	Finals 06
TacTex 06	0	n/a	5.85
PhantAgent 06	0	n/a	4.15
Deep Maize 06 SF	0	n/a	n/a
Mertacor 05	–0.57	0.55	n/a
Deep Maize 06 F	–0.95	n/a	3.58
Maxon 06 S[11]	–1.03	n/a	n/a
MinneTAC 05[11]	–1.23	–0.31	n/a
PhantAgent 05	–1.51	n/a	n/a
Deep Maize 05	–3.18	–0.22	n/a
MinneTAC 06	–3.48	n/a	–2.70
TacTex 05	–5.96	4.74	n/a

Table 7: Ranking of eleven TAC SCM agents based on deviations from an equilibrium context, along with tournament results (in \$M).

This ranking supports the case for substantially improved agent performance in the 2006 competition. Deviating to a 2005 agent from the 2006 equilibrium typically incurs a large loss. The exception is Mertacor, which shows a relatively small loss–smaller than two of the 2006 agents, including DeepMaize F which placed third overall. This agent continues to show strong performance in the game-theoretic analysis. We also note that this ranking generally corresponds to the ranking based on tournament results. The exception is TacTex-05, which ranks lower than one might expect based on tournament performance.

5. DISCUSSION AND FUTURE WORK

Our case study of the TAC/SCM market illustrates some of the methods we have found useful in applying empirical game-theoretic analysis to scenarios of interest. Through the use of the agent repository we were able to compare successive years of agent strategies. The deviation and equilibrium analysis revealed that the 2005 agent PH05 which, by tournament standards, was relatively weak compared to the final's agents, displayed strong behavior with respect to the extended strategy pool.

The subsequent year's analysis again revealed that an agent from the semi-final rounds of the tournament (DeepMaize SF) had strong support in the game equilibrium. In this case the Deep Maize team

[9]This omission does not change the context substantially, and requires much less data.

[10]This is the only symmetric equilibrium we have found after extensive search, but we cannot guarantee that it is unique.

[11]These strategies had a lower minimum number of samples, 13 & 18, respectively, than the remaining nine strategies.

had a strategic choice of which agent to play in the final rounds. Our analysis reveals (with hindsight) that playing DeepMaize F was the wrong choice given the strategic context.

Finally, we provide a method of ranking agent strategies alternative to the traditional tournament rankings. Our method is relative to a background context, which in our analysis was a sample Nash equilibrium.[12] The resulting NE-response ranking is consistent with the TAC-06 tournament, in that the top two tournament agents are tied for first in the new ranking, and the ordering of two other tournament agents is preserved. Our method allows us to extend the ranking beyond the tournament finals, including one semi-finals agent (tied with the top two tournament finals for first), and agent strategies from the prior year. This analysis has also yielded empirical evidence that agents are increasing in competency, in particular that they are improving responses to the previous year's equilibrium mixture.

Further development of these and other techniques with experience will lead to a rich set of tools bridging simulation and game-theoretic approaches to understanding complex multiagent systems. In particular, given the rate of growth of the profile space, uniform exploration over future strategies would be quite infeasible, so we require some guidance to focus on the parts of the profile space most relevant to strategic analysis. We suspect that some existing methods can be employed to improve and automate our sampling process. For example, the information-theoretic criteria proposed by Walsh et al. [19], designed to allocate additional samples given a completely evaluated empirical game, could perhaps be extended to cases with missing profiles. We intend that future work be addressed to principled methods for introducing new profiles and strategies as well.

Acknowledgments

We thank the TAC community, especially the creators of the agent repository (Joakim Eriksson and Kevin O'Malley), and entrants who have shared their agents for experimentation. Jason Miller contributed significantly to the simulation testbed setup. This work was supported in part by the National Science Foundation grant IIS-0205435 and the National Science Foundation IGERT grant 0114368.

6. REFERENCES

[1] R. Arunachalam and N. M. Sadeh. The supply chain trading agent competition. *Electronic Commerce Research and Applications*, 4:63–81, 2005.

[2] R. Axelrod. *The Evolution of Cooperation*. Basic Books, 1984.

[3] J. Collins, R. Arunachalam, N. Sadeh, J. Eriksson, N. Finne, and S. Janson. The Supply Chain Management Game for the 2005 Trading Agent Competition. Technical Report CMU-ISRI-04-139, Carnegie Mellon University, 2004.

[4] J. Eriksson, N. Finne, and S. Janson. Evolution of a supply chain management game for the trading agent competition. *AI Communications*, 19:1–12, 2006.

[5] D. Friedman. Evolutionary games in economics. *Econometrica*, 59:637–666, 1991.

[6] M. He, A. Rogers, X. Luo, and N. R. Jennings. Designing a successful trading agent for supply chain management. In *Fifth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 1159–1166, Hakodate, 2006.

[7] C. Kiekintveld, J. Miller, P. R. Jordan, and M. P. Wellman. Controlling a supply chain agent using value-based decomposition. In *Seventh ACM Conference on Electronic Commerce*, pages 208–217, Ann Arbor, MI, 2006.

[8] C. Kiekintveld, Y. Vorobeychik, and M. P. Wellman. An Analysis of the 2004 Supply Chain Management Trading Agent Competition. In *IJCAI-05 Workshop on Trading Agent Design and Analysis*, 2005.

[9] I. Kontogounis, K. C. Chatzidimitriou, A. L. Symeonidis, and P. A. Mitkas. A robust agent design for dynamic SCM environments. In *Fourth Hellenic Conference on Artificial Intelligence*, Heraklion, Greece, 2006.

[10] M. Mitchell. *An Introduction to Genetic Algorithms*. MIT Press, 1996.

[11] D. Pardoe and P. Stone. Predictive planning for supply chain management. In *Sixteenth International Conference on Automated Planning and Scheduling*, Cumbria, UK, 2006.

[12] D. Pardoe and P. Stone. TacTex-05: A champion supply chain management agent. In *Twenty-First National Conference on Artificial Intelligence*, pages 1489–1494, Boston, 2006.

[13] S. Phelps, M. Marcinkiewicz, S. Parsons, and P. McBurney. A novel method for automatic strategy acquisition in n-player non-zero-sum games. In *Fifth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 705–712, Hakodate, 2006.

[14] D. M. Reeves. *Generating Trading Agent Strategies: Analytic and Empirical Methods for Infinite and Large Games*. PhD thesis, University of Michigan, 2005.

[15] S. M. Ross. *Simulation*. Academic Press, third edition, 2002.

[16] P. Taylor and L. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156, 1978.

[17] L. Tesfatsion and K. L. Judd, editors. *Handbook of Agent-Based Computational Economics*. Elsevier, 2006.

[18] Y. Vorobeychik, C. Kiekintveld, and M. P. Wellman. Empirical mechanism design: Methods, with application to a supply-chain scenario. In *Seventh ACM Conference on Electronic Commerce*, pages 306–315, Ann Arbor, MI, 2006.

[19] W. Walsh, D. Parkes, and R. Das. Choosing samples to compute heuristic-strategy Nash equilibrium. In *Fifth Workshop on Agent-Mediated Electronic Commerce*, 2003.

[20] M. P. Wellman, J. Estelle, S. Singh, Y. Vorobeychik, C. Kiekintveld, and V. Soni. Strategic Interactions in a Supply Chain Game. *Computational Intelligence*, 21:1–26, 2005.

[21] M. P. Wellman, D. M. Reeves, K. M. Lochner, S.-F. Chen, and R. Suri. Approximate Strategic Reasoning through Hierarchical Reduction of Large Symmetric Games. *Twentieth National Conference on Artificial Intelligence*, pages 502–508, 2005.

[22] M. P. Wellman, P. R. Wurman, K. O'Malley, R. Bangera, S. de Lin, D. M. Reeves, and W. E. Walsh. Designing the market game for a trading agent competition. *IEEE Internet Computing*, 5(2):43–51, 2001.

[12]We found only a single symmetric equilibrium for the 2006 $\text{SCM}{\downarrow_3}$ empirical game. For cases of multiple equilibria, we propose taking a weighted average of the response scores to the various equilibria.