Empirical Game-Theoretic Analysis of the TAC Market Games

Michael P. Wellman, Patrick R. Jordan, Christopher Kiekintveld, Jason Miller, and Daniel M. Reeves

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

ABSTRACT

We illustrate developing techniques for empirical game-theoretic analysis by application to two challenging market games employed in an annual Trading Agent Competition. These games exemplify relevant environments beyond analytic tractability, yet which can be investigated experimentally through simulation and careful measurement. Our analysis of agents from the 2005 TAC Supply Chain Management game reveals interesting interactions not seen in the tournament, and demonstrates the value of a post-competition agent repository. Our ongoing analysis of the TAC Travel game illustrates methods for scaling up the strategy and profile space, and demonstrates the value of empirical game-theoretic analysis for strategy selection. Interesting open issues remain, particularly in regard to controlling the experimental search process.

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 [Axelrod, 1984], and has become a standard method among researchers in agent-based computational economics [Tesfatsion and Judd, 2006]. 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 [Mitchell, 1996].
2. *Finding stable populations* of strategies, for example using replicator dynamics [Taylor and Jonker, 1978].

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 [Friedman, 1991], this is not always the case. Game theory generally 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] Over the past few years, our research group has been developing an experimental methodology for explicit game-theoretic treatment of MAS simulation studies. We refer to the approach as *empirical game-theoretic analysis*. Along the way we have accumulated a body of tricks and techniques that enhance the empirical analysis approach. We illustrate the application of some of these by way of presenting analyses of two scenarios from an MAS research competition.

2. TAC MARKET GAMES

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 [Wellman et al., 2001]. A second game, in the domain of supply chain management, has been played since 2003 [Arunachalam and Sadeh, 2005].

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.

Although most details of the game rules are inessential to our analyses here, we establish some context by providing capsule descriptions. Complete specifications of the games are available at the web sites referenced, and further description and discussion is provided in many of the papers cited herein.

2.1 Supply Chain Management

In the Trading Agent Competition Supply Chain Management game (TAC/SCM), six agents representing PC (personal computer) manufacturer agents compete to maximize their profits over a simulated year. There are 220 simulation 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.

2.2 Travel Shopping

In the TAC Travel game, agents assemble flights, hotels, and entertainment into trips for a set of eight probabilistically generated clients. Clients are described by their preferred arrival and departure days, the premium they are willing to pay to stay at the nicer hotel, and their respective values for three different types of entertainment events. The agents' objective is to maximize the value of trips for their clients, net of expenditures in the markets for travel goods. The three categories of goods are exchanged through distinct market mechanisms.

Flights. A feasible trip includes air transportation both ways, comprising an inflight day i and outflight day j, $1 \leq i < j \leq 5$. Flights in and out each day are sold independently, at prices determined by a stochastic process. The initial price for each flight is uniformly distributed, and follows a random walk thereafter with an increasingly upward bias.

Hotels. Feasible trips must also include a room in one of the two hotels for each night of the client's stay. There are 16 rooms available in each hotel each night, and these are sold through ascending 16th-price auctions. Agents submit bids for various quantities, specifying the price offered for each additional unit. When the auction closes, the units are allocated to the 16 highest offers, with all bidders paying the price of the lowest winning offer. Each minute, the hotel auctions issue *quotes*, indicating the 16th- and 17th-highest prices among the currently active unit offers. Each minute, one of the hotel auctions is selected at random to close, with the others remaining active and open for bids.

Entertainment. Agents receive an initial random allocation of entertainment tickets (indexed by type and day), which they may allocate to their own clients or sell to other agents through continuous double auctions. The entertainment auctions issue quotes representing the highest outstanding buy and lowest sell offer, and remain open for buying and selling throughout the 9-minute game duration.

At the end of a game instance, the TAC server calculates the optimal allocation of trips to clients for each agent, given final holdings of flights, hotels, and entertainment. The agent's game score is its total client trip utility, minus net expenditures in the TAC auctions.

3. TAC/SCM 2005

Our 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 [Wellman et al., 2005a]. In that study, we

[1] Perhaps it is starting to emerge. Although we do not attempt here to identify the earliest sources (see Reeves [2005, Section 3.9] for a survey), we do acknowledge that many MAS works have included elements of game-theoretic perspective in experimental studies. 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`.

investigated a phenomenon observed in the 2003 tournament, employing strategies defined by varying one aspect of the University of Michigan's agent (our own), Deep Maize. Controlled experiments varying only the degree to which agents procure components on day 0 verified that the agressive procurement policies observed (informally) in tournament play actually represents an equilibrium of sorts—and one that is mutually destructive to manufacturing profits. Our 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 [Kiekintveld et al., 2005], and empirical game-theoretic analysis demonstrates that no reasonable settings of the focal storage-cost parameter would have likely been sufficient [Vorobeychik et al., 2006]. The 2005 redesign [Collins et al., 2004] included deeper changes to supplier behavior, and appears to have dramatically lessened the salience of day-0 procurement issues.

Like our own reports, 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 [He et al., 2006]—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 and Stone [2006] 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).

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. Our approach comprises the following steps:

1. Approximate the six-player SCM game by its three-player reduced version, SCM$\downarrow_3$ [Wellman et al., 2005b].
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.

We elaborate each of these steps in turn.

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, we have found experimentally and theoretically that this form of *hierarchical game reduction* produces results approximating well the original unreduced game, with great computational savings [Reeves, 2005, Wellman et al., 2005b]. Although we have not validated this specifically in TAC/SCM, intuitively we would expect that payoffs vary smoothly with the number of other agents playing a given strategy.

3.2 Running Simulations

We have collected results from well over 2000 sample games, covering 56 distinct profiles of six agents: TacTex [Pardoe and Stone, 2006], Mertacor [Kontogounis et al., 2006], Deep Maize [Kiekintveld et al., 2006], MinneTAC, PhantAgent, and GoBlueOval.[5] As shown in Table 1, 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. The game-theoretic analysis here thus also serves to assess the robustness of tournament ranking results.

We performed most of our simulations using a computing cluster operated by the Center for Advanced Computing at 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. Our basic computational package called for seven CPUs (the six agents plus a game server) for a period of three and a half hours. When the cluster's schedule grants our request, we distribute the agent and server binaries to the allocated nodes, updating configuration files dynamically to point agents to the corresponding game server. We then run three games, and afterwards copy the game logs back to a central

[3] Here we refer to the 2003 version of Deep Maize. Subsequent discussion applies to the 2005 versions of Deep Maize and all other agents mentioned.

[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`.

[5] At least two of these, TacTex and Deep Maize, employed facilities for adapting behavior between games during the TAC-05 tournament. Our analysis here restricts the agents to versions that adapt only within a game instance.

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 1: 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).

repository. By packaging games into groups of three, we achieve a balance of amortizing configuration time with limiting recomputation necessary in case of failure.

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). The dataset analyzed in this paper comprises 2110 validated games, with a minimum of 28 games for each of the 56 distinct strategy profiles.

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* [Ross, 2002] 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 our analysis of TAC/SCM-03 [Wellman et al., 2005a], we use control variates to derive a payoff measure we call *demand-adjusted profit* (DAP). Our 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*.[6] We collected the demand and score data from games played in the TAC/SCM-05 tournament: quarter-final, semi-final, and final rounds. The overall tournament comprised 96 games, of which 71 remained after applying the tainted-game filter described above. Table 2 presents summary demand statistics for these games.

Segment	Mean (Q_{seg})	Std. Dev.	DAP Coeff. (δ_{seg})
Low	132,498	29,102	69.67
Mid	157,481	33,698	63.80
High	129,641	21,870	85.57

Table 2: TAC/SCM-05 tournament demand statistics.

The rightmost column of the table (DAP coefficient) presents the result of a linear regression of score on demand in the respective segments. The R^2 statistic for the DAP regression is 0.3469 with a p-value of 2.5e-6. We then obtain the DAP for agent i in game x by subtracting from its actual profit an adjustment based on the demand in that game.

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

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 2.

3.4 Game 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

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

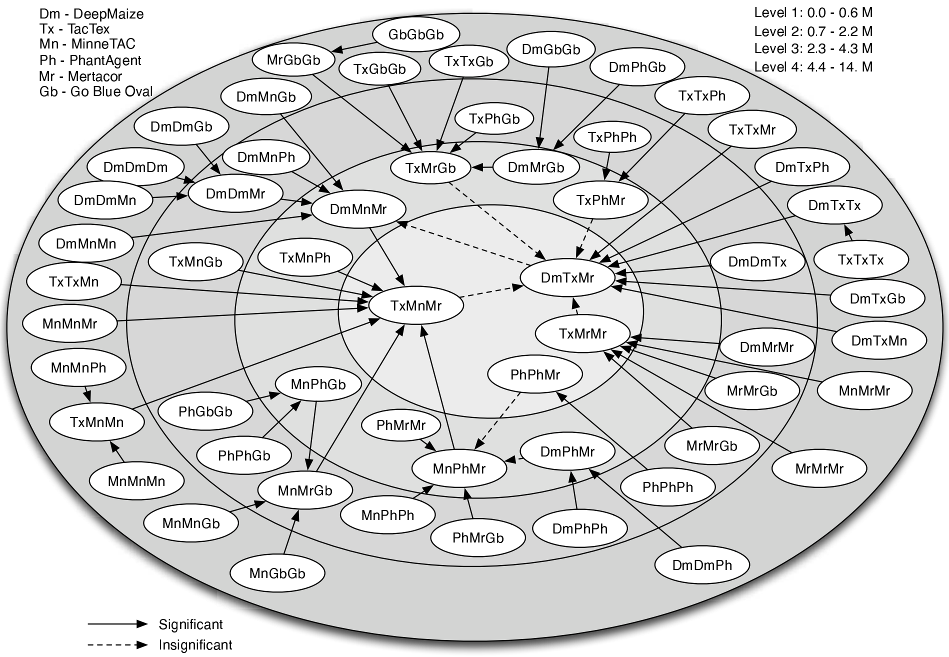

Figure 1: Deviation analysis of pure profiles of SCM$\downarrow_3$.

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 is most probable in all but the case 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. TAC TRAVEL

In contrast with the TAC/SCM analysis presented above, our empirical game-theoretic analysis of TAC Travel focuses on variations of our own agent, Walverine [Cheng et al., 2005]. The database of games comprises simulations from a dedicated testbed running virtually continuously since mid-2004. Our primary objective in this experiment has been to inform our choice of strategy parameter settings to play in the tournament [Wellman et al., 2005c]. In fact, this turned out successfully, as Walverine placed third in the 2005 TAC Travel tournament, and actually scored highest if we ignore games tainted by faulty behavior of one of the agents.

Pieces of our TAC Travel analysis have appeared in previous reports on Walverine and our methodology [Wellman et al., 2005b,c]. Here we provide an update based on additional simulations (over 25,000 more games, for a total of 72,971), as well as some new results and discussion not included elsewhere.

As in the TAC/SCM analysis, we employ control variates to reduce

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

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 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.

the variance due to stochastic inputs in our simulation process. In TAC Travel, the important stochastic influences are agent-specific, namely the client preferences assigned to agents at the beginning of the game. Accordingly, we adjust observed scores based on the potential premium available to clients from hotel and entertainment, a measure of the conflict of their demand with other agents' clients, and one agent-independent feature: the initial flight prices [Wellman et al., 2005c].

Another technique shared in the two analyses is our exploitation of game reduction to limit the explosion of strategy profile space. TAC Travel is an 8-player game, and thus enables reduction across a hierarchy of levels. Table 4 shows how our dataset is apportioned among the 1-, 2-, and 4-player reduced games. We are able to exhaustively cover the 1-player game, of course. We could also have exhausted the 2-player profiles, but chose to skip some of the less promising ones (around one-quarter) in favor of devoting more samples elsewhere. The available number of samples could not cover the 4-player games, but as we see below, even 2.4% is sufficient to draw conclusions about the possible equilibria of the game. Spread over the 8-player game, however, 73,000 instances would be insufficient to explore much, and so we refrain from any sampling of the unreduced game.

p	Profiles			Samples/Profile	
	total	evaluated	%	min	mean
4	123,410	2967	2.4	15	24.6
2	840	624	74.3	18	36.4
1	40	40	100.0	30	89.1

Table 4: Profiles evaluated, reduced TAC games (TAC$\downarrow_p$).

4.1 TAC$\downarrow_4$ Analysis

Although our coverage of TAC$\downarrow_4$ is far from exhaustive, for particular subsets of strategies (which we call *cliques*), we do have samples for all possible profiles. The largest TAC$\downarrow_4$ cliques contain five strategies, and we currently have seven of these (which have overlapping strategy sets). We used replicator dynamics to derive a symmetric mixed equilibrium for each of these clique games, and for five of these the profile is a candidate equilibrium ($\epsilon < 1$) with respect to the entire dataset (i.e., no beneficial deviations are found among the evaluated profiles). These candidates are presented in Table 5.

Strategies in Table 5 are described by index number, each corresponding to a vector of parameter settings. The actual values are not relevant to present purposes, except to note that the Walverine played in the 2004 tournament finals was strategy 17, and that strategy 37 played in 2005. We used equilibrium analysis of the sort displayed here to suggest a subset of plausible strategies (e.g., those appearing with significant probability in some equilibrium), and then selected among these based on play against the actual TAC field in preliminary tournament rounds.

Strategy	mixed profiles				
3	0	.125	—	—	—
4	0	.687	—	—	—
5	—	—	0	0	—
7	—	—	.273	—	—
9	—	—	—	.250	—
16	.744	.071	.601	.418	.085
17	.225	0	—	—	0
18	—	—	—	—	—
21	—	—	0	0	—
23	.031	—	—	—	—
24	—	.117	.126	.332	—
37	—	—	—	—	.225
39	—	—	—	—	0
40	—	—	—	—	.690

Table 5: Candidate equilibria in TAC $\downarrow_4$, as determined by replicator dynamics applied to clique subgames. Each column represents a symmetric mixed profile, with dashes indicating that a strategy was not included in the corresponding clique.

Another form of analysis considers the maximum benefit from deviation (ϵ bounds) established for the various TAC$\downarrow_4$ profiles. As indicated in Table 4, we have evaluated 2967 TAC$\downarrow_4$ profiles. Of these, 201 are TAC$\downarrow_2$ profiles with no evaluated neighbors in TAC$\downarrow_4$ (i.e., no deviations tested). Although these are technically PSNE candidates, we distinguish them from PSNE candidates that have actually survived some challenge. The remaining 2766 evaluated profiles are of course too many to diagram as in Figure 1. Instead, we plot the distribution of ϵ bounds, in Figure 2.

Figure 2 also shows, inset, the distribution of epsilon bounds over the 182 strategy pairs for which we have evaluated all combinations in TAC $\downarrow_4$ (i.e., the 2-cliques). Among these are one confirmed equilibrium at $\epsilon = 2.5$, with all other pairs refuted at $\epsilon > 9$.

Recall from our discussion of the SCM empirical game the observation that Mertacor appears versatile, as indicated by the frequency with which the best deviation from a profile is to that agent. We can

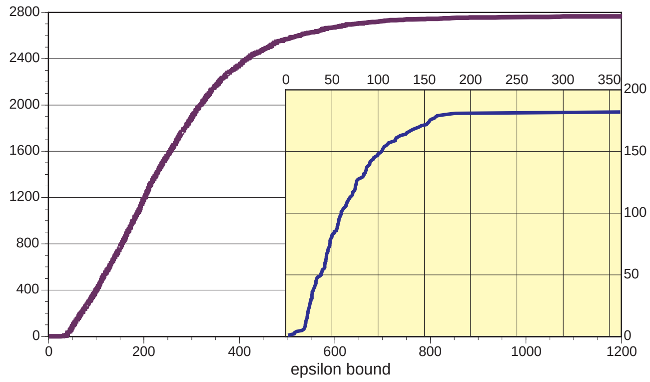

Figure 2: Cumulative distribution of ϵ bounds in TAC$\downarrow_4$. Main graph: pure profiles. Inset: 2-strategy mixtures.

examine deviations from our TAC$\downarrow_4$ profiles analogously, and find that of the 2766 evaluated profiles with deviations, in 354 of them (12.8%), the best deviation is to strategy 37 (Walverine-2005). The next "most deviated to" strategy covers a like amount (350), and next two after that cover 262 and 259 profiles, respectively. Note that one must interpret such numbers with much caution, since the potential deviations are not evaluated with equal frequency. Moreover, any figure of merit that aggregates across profiles is suspect, as the distribution of available profiles is not uniform, and profiles are not equally relevant to such an evaluation anyway. Nevertheless, the breadth of contexts in which 37 performs well provides further support for its choice from among the set of strategies considered.

4.2 Search through Profile Space

Our empirical game for TAC Travel is based on a large dataset of game instances, accumulated over a great deal of time with much computation. One important part of the experimental procedure not described thus far is how we chose which agent strategy configurations to sample, and to what extent. Indeed, the process is manually controlled, informed by standard analysis routines but somewhat ad hoc. Nevertheless, we describe our basic approach, and raise the issue as an important area for future research in empirical game-theoretic methodology.

The relevant question at any point in the sampling process is: "What profile to sample next?" We can choose to generate (1) an additional sample of a profile already evaluated, (2) a first sample for an unevaluated profile comprised of existing strategies, or (3) a first sample for a profile including a new strategy.

Our approach to introducing new strategies in TAC Travel was entirely manual and admittedly arbitrary. Since the profile space explodes in the number of strategies, we are generally conservative, becoming more amenable as the existing strategy base appears to us relatively well understood. In many cases, we introduced new strategies based on discovering new ideas for agent components, or problems with some of the existing elements.

We followed a much more structured process for introducing new profiles of existing strategies. In general, profiles are introduced with a view toward refuting candidate equilibria. Specifically, we tend to seek profiles that represent deviations from an existing pure profile or 2-strategy mixture with small ϵ bound. By interleaving game analysis with sampling, we can identify prospective profiles routinely. Since there will generally be many choices of how to deviate, we require secondary criteria as well. For instance, we prefer profiles that deviate from multiple candidates, or have many evaluated neighbors (e.g., will contribute to forming cliques) already in the dataset.

Note that the foregoing selection can be applied with respect to the game at any level of reduction. We have interleaved consideration of TAC$\downarrow_1$, TAC$\downarrow_2$, and TAC$\downarrow_4$, devoting more effort toward the finer-grained games as the coarser levels become better defined (i.e., once deviations from candidates of more severely reduced games have been thoroughly explored).

With respect to profiles already sampled, our highest priority is to maintain a minimum number of samples (see Table 4) for any evaluated profile. Next, whenever we explore new deviations from a candidate, we also allocate some samples to the candidate itself, and profiles that currently seem to be the best deviations.

Although the process described here is clearly informal and could benefit from analysis and optimization, we believe it contains several important qualitative features. Given the size of the search space, uniform exploration would be quite infeasible, and so we require some guidance to focus on the parts of profile space most relevant to strategic analysis. The criteria we adopted aim to balance exploration of new directions with better understanding of areas with established promise. 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. [2003], 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.

5. CONCLUSION

Our two case studies illustrate some of the methods we have found useful in applying empirical game-theoretic analysis to scenarios of interest. 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.

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. Members of the Walverine team over the years have contributed much to our TAC Travel experiments. This work was supported in part by the National Science Foundation under grants IIS-0205435 and IIS-0414710.

References

Raghu Arunachalam and Norman M. Sadeh. The supply chain trading agent competition. *Electronic Commerce Research and Applications*, 4:63–81, 2005.

Robert Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.

Shih-Fen Cheng, Evan Leung, Kevin M. Lochner, Kevin O'Malley, Daniel M. Reeves, and Michael P. Wellman. Walverine: A Walrasian trading agent. *Decision Support Systems*, 39:169–184, 2005.

John Collins, Raghu Arunachalam, Norman Sadeh, Joakim Eriksson, Niclas Finne, and Sverker Janson. The supply chain management game for the 2005 trading agent competition. Technical report, Carnegie Mellon University, 2004.

Daniel Friedman. Evolutionary games in economics. *Econometrica*, 59:637–666, 1991.

Minghua He, Alex Rogers, Xudong Luo, and Nicholas R. Jennings. Designing a successful trading agent for supply chain management. In *Fifth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, Hakodate, 2006.

Christopher Kiekintveld, Yevgeniy Vorobeychik, and Michael P. Wellman. An analysis of the 2004 supply chain management trading agent competition. In *IJCAI-05 Workshop on Trading Agent Design and Analysis*, Edinburgh, 2005.

Christopher Kiekintveld, Jason Miller, Patrick R. Jordan, and Michael P. Wellman. Controlling a supply chain agent using value-based decomposition. In *Seventh ACM Conference on Electronic Commerce*, Ann Arbor, MI, 2006.

Ioannis Kontogounis, Kyriakos C. Chatzidimitriou, Andreas L. Symeonidis, and Pericles A. Mitkas. A robust agent design for dynamic SCM environments. In *Fourth Hellenic Conference on Artificial Intelligence*, Heraklion, Greece, 2006.

Melanie Mitchell. *An Introduction to Genetic Algorithms*. MIT Press, 1996.

David Pardoe and Peter Stone. Predictive planning for supply chain management. In *Sixteenth International Conference on Automated Planning and Scheduling*, Cumbria, UK, 2006.

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

Sheldon M. Ross. *Simulation*. Academic Press, third edition, 2002.

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

Leigh Tesfatsion and Kenneth L. Judd, editors. *Handbook of Agent-Based Computational Economics*. Elsevier, 2006.

Yevgeniy Vorobeychik, Christopher Kiekintveld, and Michael P. Wellman. Empirical mechanism design: Methods, with application to a supply chain scenario. In *Seventh ACM Conference on Electronic Commerce*, Ann Arbor, MI, 2006.

William E. Walsh, David Parkes, and Rajarshi Das. Choosing samples to compute heuristic-strategy Nash equilibrium. In *AAMAS-03 Workshop on Agent-Mediated Electronic Commerce*, Melbourne, 2003.

Michael P. Wellman, Peter R. Wurman, Kevin O'Malley, Roshan Bangera, Shou-de Lin, Daniel Reeves, and William E. Walsh. Designing the market game for a trading agent competition. *IEEE Internet Computing*, 5(2):43–51, 2001.

Michael P. Wellman, Joshua Estelle, Satinder Singh, Yevgeniy Vorobeychik, Christopher Kiekintveld, and Vishal Soni. Strategic interactions in a supply chain game. *Computational Intelligence*, 21:1–26, 2005a.

Michael P. Wellman, Daniel M. Reeves, Kevin M. Lochner, Shih-Fen Cheng, and Rahul Suri. Approximate strategic reasoning through hierarchical reduction of large symmetric games. In *Twentieth National Conference on Artificial Intelligence*, Pittsburgh, 2005b.

Michael P. Wellman, Daniel M. Reeves, Kevin M. Lochner, and Rahul Suri. Searching for Walverine 2005. In *IJCAI-05 Workshop on Trading Agent Design and Analysis*, Edinburgh, 2005c.