Generalization Risk Minimization in Empirical Game Models

Patrick R. Jordan and Michael P. Wellman
University of Michigan
Computer Science & Engineering
Ann Arbor, MI 48109-2121 USA
{prjordan,wellman}@umich.edu

ABSTRACT

Experimental analysis of agent strategies in multiagent systems presents a tradeoff between granularity and statistical confidence. Collecting a large amount of data about each strategy profile improves confidence, but restricts the range of strategies and profiles that can be explored. We propose a flexible approach, where multiple game-theoretic formulations can be constructed to model the same underlying scenario (observation dataset). The prospect of incorrectly selecting an empirical model is termed *generalization risk*, and the generalization risk framework we describe provides a general criterion for empirical modeling choices, such as adoption of factored strategies or other structured representations of a game model. We propose a principled method of managing generalization risk to derive the optimal game-theoretic model for the observed data in a restricted class of models. Application to a large dataset generated from a trading agent scenario validates the method.

General Terms

Economics, Experimentation

Keywords

Empirical game, Generalization risk minimization

1. INTRODUCTION

Multiagent systems (MAS) research has long relied on both theoretical and experimental analysis to understand the implications of alternative agent behaviors, or strategies. For situations amenable to analytic modeling, researchers often appeal to the framework of game theory, attracted by its normative force and generality, as well as its rich mathematical structure. For scenarios that are too complex or lack directly specified game forms (i.e., payoff or utility functions), analytic game theory is not immediately applicable. In such situations, *empirical game models* [15], where observations or simulations of agent play are used to construct estimates of their utility, can support game-theoretic analysis despite lack of explicit game descriptions.

Empirical game-theoretic models are founded on an underlying game simulator, which generates outcomes from agent behaviors taken from an underlying strategy space. The simulator defines the system in which the agents participate. Configuration parameters may include the number of agents, the allowable actions for each, possible type realizations (i.e., settings of private information for the agents), and other factors. Though the full strategy space allowed by the game simulator may be large or infinite, due to computational constraints empirical game models usually restrict the strategy space to a small number of heuristically defined strategies. Even within this restricted space, modeling accuracy is limited by statistical sampling, given the inherent stochastic behavior of typical simulation environments. Given a limited computational budget, modelers must choose carefully which strategy profiles to simulate and how many samples of those profiles to take.

One benefit of these empirical models is that game representations of a MAS can be made with varying degrees of fidelity to the underlying scenario. For instance, a stock market scenario may involve thousands of participants with a potentially infinite number of available trading strategies. A simulation of the scenario may be reduced to a manageable number of representative participants each selecting from a handful of promising strategies. Depending on the computational budget, an analyst may increase or decrease the population size or available trading strategies in the simulator.

Empirical game-theoretic techniques have expanded researchers' capabilities of analyzing complex strategic scenarios by enabling increasingly faithful models of interaction. However, with the increased complexity comes an increase in modeling risk. Increasingly complex simulations, all else being equal, are more costly in terms of computation requirements. Moreover, increasing the complexity of the simulation generally entails increasing the size of the potential strategy space used in simulating the scenario. Because the strategy space grows while the number of heuristic strategies that can be feasibly analyzed is at best held constant, there is an increasing chance that a useful strategy is omitted from the heuristic strategies in analysis. This risk is endemic in modeling MAS scenarios and can be mitigated only partially by increments in computational budget.

A second source of risk arises from the analysis of the available observations. Once the simulator and strategy sets are defined, then observations can be collected and an empirical game model estimated from the observation data. The empirical game model provides an estimate for the utility of playing a profile within its profile space. In much of the previous research on empirical games, the strategy space of the empirical game is assumed identical to that of underlying simulation. We relax this constraint, allowing models where the strategy sets or even players do not correspond precisely to the base notions defined by the simulator. For instance, an empirical game model may treat two strategies that are distinct for the simulator as interchangeable in its own strategy space. This coarsens the model the empirical game uses to predict payoffs, re-

Cite as: Generalization Risk Minimization in Empirical Game Models, Patrick R. Jordan, Michael P. Wellman, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009),* Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 553 – 560

ducing the model's complexity compared to the finer-grained strategy space of the simulator.

Entertaining multiple candidate models of varying complexity provides useful flexibility. A more complex model may capture observations better than a simpler one, however it may also be more susceptible to fitting spurious information in the observations. We term this the *generalization risk* associated with an empirical game model. In the game-theoretic context, the consequence of incorrect generalization could be that profiles that appear stable are actually unstable, or vice versa. Thus, we develop a framework that allows us to compare candidate models and make an appropriate selection based on the model's predictive power. To illustrate the flexibility of this approach we describe two modeling choice scenarios for empirical games: the *equivalent strategy model* and the *factored model*. These models take an existing strategy space and transform it in different ways. In our experimental analysis, Section 4, we construct a hierarchy of equivalent strategy models which are used to analyze a game of interest.

Our motivating problem for this analysis is the Trading Agent Competition's Supply Chain Management game (TAC/SCM). The annual Trading Agent Competition (TAC) series of international research tournaments was initiated to promote research and education in the technology underlying trading agents.[1] At the core of TAC are several *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 [16]. A second game, in the domain of supply chain management, has been played since 2003 [3]. CAT, the third game in the series, is a market design game in which agents play the role of market specialists [9]. A new game in the domain of advertising auctions (TAC/AA) is scheduled to debut in 2009.

A key feature of all TAC 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. In such situations, empirical game-theoretic methods can often provide a useful basis for strategic reasoning. Typically in this approach, a game structure (player and profile space) is fixed and then simulation is undertaken to provide estimates for the utility function of the underlying game. Once a sufficient number of observations have been made, a suite of empirical game-theoretic analysis methods can be employed. Previous research has developed equilibrium-inspired methods for ranking strategies and visualizing deviations for empirical games such as TAC [6]. Also useful in empirical modeling are techniques to reduce the size of the game space through clustering profiles [5, 17], construct compact game factorizations [4], and exploit strategic independence [8]. These approaches define reduced or simplified game structures which, when applicable, require fewer observations to support accurate estimates of the utility function. Complementary to the modeling options are search methods that focus on identifying stable strategies under a limiting computational budget [7, 11, 12, 13]. To date, there has been little research on expanding the fixed game form methodologies to test multiple game form hypotheses on the same observation set. In this paper we develop a general framework for comparing the *fit* of multiple game forms on observational MAS data.

The subsequent portions of this analysis are arranged as follows. In Section 2, we introduce our framework for empirical games based on observations from a simulator. In Section 3, we present our approach to managing generalization risk and discuss two applicable classes of empirical game models, equivalent strategy models and factored models. In Section 4, we apply this approach to the case of equivalent strategy modeling, using the domain of TAC/SCM. Our experiments employ data taken from years of observations on this domain. We then finish with a small discussion of extensions of this technique to other game-theoretic analysis methods.

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

2. EMPIRICAL GAME FRAMEWORK

The concept of empirical games has antecedents in both the MAS and economics literatures. The heuristic-strategy approach of Walsh et al. [14] explicitly constructs empirical games, and much experimental work in MAS effectively produces statistical estimates of payoff functions. In economics, Armantier et al. [2] defined the concept of *constrained strategic equilibrium* (CSE) as an approximation of Bayesian Nash equilibrium (BNE) induced by constraints on the strategy space. They went on to show that any sequence of CSEs has a subsequence that converges toward a BNE when the strategy space is compact. Experimental MAS studies will generally not conform to Armantier et al.'s compactness requirement; nevertheless, their result provides theoretical support for our expectation that equilibria in empirical games will improve in approximation to equilibria of the true game as more strategies are evaluated.

To describe our modeling framework, we extend notation from Armantier et al. [1, 2]. Let Ω be the set of states of nature and for each player i, let A_i denote i's available actions. In a scenario of N players the *joint action space* is $A = \prod_{i \in N} A_i$ and an element $a \in A$ is a *joint action*. Each player receives a *type* $\xi_i \in \Xi_i$ and the *joint type space* is defined as $\Xi = \prod_{i \in N} \Xi_i$. A player's strategy is a measurable function $\rho_i : \Xi_i \to A_i$ where $a_i = \rho_i(\xi_i)$. $P_i = \{\rho_i\}$ is player i's strategy space and $P = \prod_{i \in N} P_i$ is the *joint strategy space*. A *joint strategy* or *profile* is thus $\rho \in P$, where $a = \rho(\xi)$ and ξ is the joint type. The *joint opponent strategy space* is denoted $P_{-i} = \prod_{j \in N \setminus \{i\}} P_j$. A *mixed strategy* σ_i is a probability distribution over strategies in P_i, with $\sigma_i(\rho_i)$ denoting the probability player i will play strategy ρ_i. The *mixed strategy space* for player i is given by $\Delta(P_i)$. Similarly, $\Delta(P) = \prod_{i \in N} \Delta(P_i)$ is the *mixed profile space*.

In complex multiagent environments, strategies are often best described procedurally, in effect, as computer programs that take type information as input and return the selected action or sequence of environment interactions as their output. For this reason, we often can do no better than treat these strategies as black boxes, analyzing them in terms of input-output but not internal structure. We define a *simulator* as a function that maps joint strategies to outcomes, in the form of payoff vectors. Typically, a simulator will itself be realized as a program that generates type information and implements the interaction among the participating agents and the environment.

DEFINITION 1 (SIMULATOR). *A simulator is a measurable function $\mathcal{S} : \Omega \times P \to \mathbb{R}^{|N|}$, where P is the profile space and N is the set of players.*

An individual run of the simulator, or *simulation*, produces an *observation* $\theta = (\rho, \pi)$, where ρ is the strategy profile simulated and $\pi : N \to \mathbb{R}$ is the *joint payoff* received by the agents. Let $\Theta = \{\theta_k\}$ be the *observation set*. A simulator $\mathcal{S}$ implicitly defines a specific game over the set of players N and a profile space P, but without an explicit utility function as would normally be specified in a game description. Instead, the observation set generated by the simulator provides the basis for an estimated game model, which is what we call the *empirical game*.

DEFINITION 2 (EMPIRICAL GAME). *An empirical game is a tuple $\mathcal{E} = \langle N^{\mathcal{E}},\ P^{\mathcal{E}},\ u^{\mathcal{E}}, \phi, \mu\rangle$ where $N^{\mathcal{E}}$ is the set of players, $P^{\mathcal{E}}$ is the joint strategy space, and $u^{\mathcal{E}}$ is the utility function. The mappings $\phi : P \to P^{\mathcal{E}}$ and $\mu : P \times \mathbb{R}^{|N^{\mathcal{E}}|} \to \mathbb{R}^{|N|}$ relate the empirical game structure to that of the underlying simulator.*

Note that as defined here, the empirical game need not employ the same player and profile space as the simulator. Indeed, the observation set will generally not span the full profile space as the underlying game. Moreover, the most *useful* game-theoretic model may not correspond exactly with the profiles simulated. Because our evidence is statistical, we must consider that the simulator's framing of the profile space may contain extraneous information that can decrease our model's predictive power. Our formulation of this tradeoff and proposed criterion to resolve it is in fact the central contribution of this work.

Accordingly, we use the superscript $\mathcal{E}$ notation for $N^{\mathcal{E}}$, $P^{\mathcal{E}}$, and $u^{\mathcal{E}}$ to distinguish the empirical game's player set, strategy space, and utility from those of the game underlying the given simulator. The relationship between these is characterized by the two mappings provided as part of the empirical game model. First, the function $\phi : P \to P^{\mathcal{E}}$ maps profiles in the simulation space to profiles in the empirical game space, preserving the neighborhood relation (defined below in terms of deviation) in the simulation space. Second, the function $\mu : P \times \mathbb{R}^{|N^{\mathcal{E}}|} \to \mathbb{R}^{|N|}$ maps a profile in the simulator and a payoff in the empirical game to a payoff in the simulation space. Using these functions we can define an *empirical utility model* $\widetilde{u}^{\mathcal{E}}(\rho) = \mu(\rho,\ u^{\mathcal{E}}(\ \phi(\rho)\ )\ )$, which applies the induced model over the simulation profile space. Note that $\widetilde{u}^{\mathcal{E}}$ may be of different dimension than $u^{\mathcal{E}}$. This can occur, for instance, if the number of players in the empirical game differs from the number of players in the simulator. For convenience, we also denote $\rho^{\mathcal{E}}$ by its equivalent, $\phi(\rho)$. Conceptually, the functions ϕ and μ allow us to move between the base simulation space and the embedded space given by the empirical game model, while preserving game-theoretic interpretations of the scenario. In Section 3, we propose a measure of generalization risk based on squared prediction error of the model in question with respect to the observed payoffs. In principle, we could choose from a multitude of estimators to minimize this error, such as neural networks or other complex functional forms, that have no explicit game-theoretic formulation. The justification for ϕ and μ in the definition is to ensure that the estimates can be meaningfully interpreted with respect to the underlying game.

3. GENERALIZATION RISK AND MODEL SELECTION

Now that we have defined our empirical game framework, we can address the problem of *fitting* an empirical model to an observation set Θ generated by simulation. Because we may be able to fit multiple models, we require some criterion for selecting among them. In particular, we would like to be able to evaluate and compare the *goodness of fit* for different models so that we may identify which is most useful for analysis. In doing so, we build upon the rich history of statistical analysis, treating game models as forms of statistical hypotheses.

A standard measure of loss for a statistical model is the mean of squared errors with respect to the data. In our context, we define the loss function $\mathscr{L}$ of a candidate empirical game model by

$$\mathscr{L}(\Theta,\ \mathcal{E}) = \frac{1}{|\Theta|}\sum_{\theta\,\in\,\Theta}\left[\widetilde{u}^{\mathcal{E}}(\theta.\rho) - \theta.\pi\right]^T\left[\widetilde{u}^{\mathcal{E}}(\theta.\rho) - \theta.\pi\right],$$

where $\theta.\rho$ and $\theta.\pi$ are the joint strategy and payoffs comprising the observation θ. We endeavor to find an empirical game model which minimizes the expected loss $\mathbb{E}[\mathscr{L}(\Theta,\ \mathcal{E})]$, where Θ is the random observation set generated from a simulator. Note that because we do not know the true distribution of Θ, we must estimate the expected loss using an existing observation set. In our experiments, we use cross-validation on the observation set to construct this estimate. We outline a k-fold cross-validation procedure in Section 3.3.

After a selected game model $\mathcal{E}$ has been fit to the observation sequence, we can analyze its game-theoretic properties. In particular, we are interested in determining the stability of profiles in terms of *regret*, the potential benefit to some player of deviating to a different strategy. In order to calculate the regret of a profile, we construct the deviation set for a particular profile in our game model. In words, the *unilateral deviation set* is simply the set of profiles in which a single player has changed its strategy with respect to the original profile. We define the construction using set notation as follows.

DEFINITION 3 (UNILATERAL DEVIATION SET). *For some $\mathcal{E}$, the* unilateral deviation set *for player $i \in N^{\mathcal{E}}$ and profile $\rho \in P^{\mathcal{E}}$ is*

$$\mathcal{D}_i^{\mathcal{E}}(\rho) = \{(\hat{\rho}_i, \rho_{-i}) : \hat{\rho}_i \in P_i^{\mathcal{E}} \setminus \{\rho_i\}\},$$

and the corresponding set, unspecified by player, is

$$\mathcal{D}^{\mathcal{E}}(\rho) = \bigcup_{i \in N^{\mathcal{E}}} \mathcal{D}_i^{\mathcal{E}}(s).$$

Once the profile structure of $\mathcal{E}$ is defined and the model induced from observations, we have an estimate for the utility function. Note that once an empirical game model has been selected, the deviations and utility calculations are performed in the embedded space, not in the original simulation space. We would like a metric that conveys the loss or regret a player incurs for playing a specific strategy given its alternatives. The measure of regret should be minimized at a Nash equilibrium. Alternative forms of this measure have been proposed in the literature, however we consider the following notion of regret to be most representative for our purposes.

DEFINITION 4 (REGRET). *For some $\mathcal{E}$, the regret of strategy profile ρ, $\epsilon^{\mathcal{E}}(\rho)$, is the maximum gain from deviation from ρ by any player. Formally,*

$$\epsilon^{\mathcal{E}}(\rho) = \max_{i\,\in\,N^{\mathcal{E}},\ \hat{\rho}\,\in\,\mathcal{D}_i^{\mathcal{E}}(\rho)\cup\{\rho\}} u_i^{\mathcal{E}}(\hat{\rho}) - u_i^{\mathcal{E}}(\rho).$$

It should be noted that loss calculations used in selecting an empirical game model are not affected by which specific measure of regret is used. Because of the generality of the loss function, we can compare many classes of candidate game models for a given simulation. For the purpose of exposition, we highlight two different modeling choice scenarios that we have identified as useful in experimental settings. Following our description of each, we conclude the section with a discussion of our cross-validation technique and iterative model selection algorithm.

3.1 Equivalent strategy models

In order to motivate the introduction of *equivalent strategy models*, we point out the strategy evaluation methodology proposed by Wellman et al. [18]. The authors systematically explored parameterized variations of their TAC travel shopping agent, Walverine.

The behavior of the agent is defined by a large number of parameters. The effects of a parameter setting are potentially dependent on the value of other parameters. For instance, two parameters determine the agent's bid shading behavior. The first parameter turns the bid shading *on* or *off* and the second parameter determines *by how much* the bids are shaded. Clearly, if the first parameter is set to *off*, then the second parameter is behaviorally irrelevant. In a factorial design methodology all possible settings would be tested, and if conducted naively, a great deal of effort will be spent analyzing duplicate strategies. In this case, it is relatively simple to identify the behaviorally equivalent strategies, but in many cases it may not be.

This highlights the general issue that, many times when designing or evaluating strategies, researchers have a set of strategies, some of which may be copies of other strategies that have simply been *labeled* differently. On the other hand, a strategy may be defined by a very large parameter space in which the behavioral effects of modifying the parameters are slight or imperceptible. In either case, researchers may not know *a priori* if there are *equivalent strategies* in the behavioral sense. Not identifying these equivalent strategies increases the generalization risk when the observation sequence has a relatively small number of observations compared to the complexity of the empirical game model's class. We introduce a game model which uses the concept of equivalent strategies to form a reduced game.

DEFINITION 5 (EQUIVALENT STRATEGY MODEL). *An equivalent strategy empirical game model $ESG = \langle N, P, \{\sim_i\}, u^{\mathcal{E}} \rangle$ constitutes a model for an empirical game where $\sim_i$ is an equivalence relation on P_i which forms equivalence classes $[\rho_i]$ with*

- $N^{\mathcal{E}} = N$
- *Player i's strategy set is the set of equivalence classes, i.e. $[\rho_i] \in P_i^{\mathcal{E}}$*
- $\phi(\rho) = ([\rho_1], [\rho_2], \ldots, [\rho_{|N|}])$
- $\mu(\rho, u^{\mathcal{E}}(\phi(\rho))) = u^{\mathcal{E}}(\phi(\rho))$.

The *equivalent strategy model* forms equivalence class $[\rho_i]$ from each relation $\sim_i$ where player i may select any element of the equivalence class with the same result. Thus all elements in $[\rho_i]$ are considered different labelings of the same underlying strategy. In other words, whenever a strategy ρ_i is observed in a simulation profile ρ, we replace it with the representative strategy $[\rho_i]$. Given some observation set Θ, we can estimate the empirical utility over $\mathcal{E}$ for a profile $u_i^{\mathcal{E}}(\rho^{\mathcal{E}})$ as the sample mean of the observed payoff set $\{\pi_i | (\pi, \rho) \in \Theta$ and $\rho^{\mathcal{E}} = \phi(\rho)\}$. By dropping the player parameterization of the equivalent relation in Definition 5, we can define a similar notion of an *equivalent strategy symmetric game*. We use this construction in our analysis of TAC/SCM in Section 4.

3.2 Factored models

In their paper on factoring games [4], Davis et al. introduce an algebra over games. In this algebra, an empirical game $\mathcal{E}$ may be composed of multiple factor games. For instance, consider a two factor game model, i.e., an empirical game $\mathcal{E}$ that is composed of two factor games $\mathcal{E}^A$ and $\mathcal{E}^B$, written $\mathcal{E} = \mathcal{E}^A \otimes \mathcal{E}^B$. The strategy sets in $\mathcal{E}$ are constructed using the cross product of the strategies in each of the factor games $\mathcal{E}^A$ and $\mathcal{E}^B$. Let P^A be the strategy set for game $\mathcal{E}^A$ and P^B for game $\mathcal{E}^B$, with respective utility functions u^A and u^B. Each composite strategy ρ in $\mathcal{E}$ is composed of two factor strategies ρ^A and ρ^B, i.e, $\rho = (\rho^A, \rho^B)$. The factors are additive in the utility, so that $u^{\mathcal{E}}(\rho) = u^A(\rho^A) + u^B(\rho^B)$. If a factoring exists for a game, it can greatly reduce the number of observations required to sufficiently observe game. Once a proposed factorization is defined, the utility function is simply a linear system of equations over the entries in the factors' payoff matrices.

Consider an agent participating in a scenario such as TAC/SCM. Agents are typically constructed in a modular way, exposing behavioral parameters. For example, the agent Deep Maize, which is further discussed in the experiments section, maintains separate parameters sets for sales and procurement decisions. The TAC/SCM scenario could be modeled as two separate factor games: one for sales and the other for procurement. If sales and procurement factorization *fit* well, this would yield substantial computational savings in the subsequent sampling analysis.

3.3 Cross-validating Empirical Game Models

In order to compare the generalization risk of differing empirical game models, we construct an estimate for the expected loss $\mathbb{E}\mathscr{L}$. This estimate is calculated using a cross-validation technique known as k-fold cross-validation. Below we provide a description of our extension to k-fold cross-validation that restricts how observations are partitioned into distinct observation sets.

First, we separate the observation set Θ into k distinct partitions as follows. Let $\Theta^{\hat{\rho}} = \{\theta \in \Theta \,|\rho(\theta) = \hat{\rho}\}$, that is all of the observations which correspond to the profile $\hat{\rho}$. For each $\rho \in P$, we randomly partition Θ^{ρ} into k equally sized groups $\Theta_1^{\rho}, \ldots, \Theta_k^{\rho}$. Each one of the k groups is assigned to one of the new observation sets such that

$$\Theta_i = \bigcup_{\rho \in P} \Theta_i^{\rho}.$$

For a given set Θ_i we define $\Theta_{-i} = \Theta \setminus \Theta_i$. We denote $\hat{\mathscr{L}}(\Theta, \mathcal{E})$ as the estimate for the expected loss given some game model $\mathcal{E}$ and define it as follows

$$\hat{\mathscr{L}}(\Theta, \mathcal{E}) = \frac{1}{k} \sum_{i=1}^{k} \mathscr{L}(\Theta_i, \mathcal{E}(\Theta_{-i}))$$

where $\mathcal{E}(\Theta_{-i})$ is the result when a game model $\mathcal{E}$ is *fit* from the observation set Θ_{-i}. We use $\hat{\mathscr{L}}(\mathcal{E})$ without the Θ parameter when the context is clear.

3.4 Iterative Model Selection in ESMs

In this section, we define an iterative procedure for selecting an equivalent strategy model (ESM) in an ESM hierarchy. We use an iterative procedure to select an equivalent strategy model. We propose this algorithm for symmetric games, but the extension to the non-symmetric case is straight forward. The algorithm works by greedily selecting the best pairwise merger of the current partition's equivalence classes, until the candidate merger increases expected loss.

First, note that, due to symmetry, the player indexing on the strategy equivalence relation $\sim$ is dropped. The algorithm proceeds in an iterative fashion. In each iteration we keep a set of hypothetical equivalence relations. We use $\mathcal{H}^{(i)}$ to denote the set of equivalence relations compared in iteration i. We start with a trivial partitioning of each strategy into its own equivalent class. We denote this $\sim^{(0)}$ which induces $\mathcal{H}^{(0)} = \{\sim^{(0)}\}$. In each iteration i, we select the partitioning $\sim^{(i)} \in \mathcal{H}^{(i)}$ which minimizes $\hat{\mathscr{L}}$.

Let $P/\sim^{(i)}$ denote the set of equivalence classes induced by the i^{th} round equivalence relation $\sim^{(i)}$ on P. In the i^{th} round we construct a candidate equivalence relation $\hat{\sim} \in \mathcal{H}^{(i)}$ for every distinct $\rho_a, \rho_b \in P/\sim^{(i-1)}$ by merging the two equivalence classes ρ_a and ρ_b.

Let $\mathcal{E}^{(i)}$ be the ESM constructed from $\sim^{(i)}$, the equivalence relation selected in the i^{th} round. If $\hat{\mathscr{L}}(\mathcal{E}^{(i-1)}) > \hat{\mathscr{L}}(\mathcal{E}^{(i)})$, then the process continues to the next iteration. Otherwise the algorithm is terminated and $\mathcal{E}^{(i-1)}$ is selected. Pseudo-code for ITERATIVE-ESM-SELECTION is given in Algorithm 1.

Algorithm 1 Iterative ESM Selection algorithm.

ITERATIVE-ESM-SELECTION($\mathcal{S}$, Θ)

$\mathcal{H}^{(0)} \leftarrow \{\sim^{(0)}\}$
Fit $\mathcal{E}^{(0)}$ from $\sim^{(0)}$ and Θ
$i \leftarrow 0$
repeat $i \leftarrow i+1$
  Construct $\mathcal{H}^{(i)}$ from $P/\sim^{(i-1)}$
  Select $\sim^{(i)} = \arg\min_{\sim \in \mathcal{H}^{(i)}} \hat{\mathscr{L}}(P/\sim)$
  Fit $\mathcal{E}^{(i)}$ from $\sim^{(i)}$ and Θ
until $\hat{\mathscr{L}}(\mathcal{E}^{(i-1)}) > \hat{\mathscr{L}}(\mathcal{E}^{(i)})$
return $\mathcal{E}^{(i-1)}$

4. EXPERIMENTS

In the following experiments, we consider game models over observations taken from the TAC Supply Chain Management scenario. In previous work [6], Jordan et al. describe the construction of the data for the TAC/SCM experiments. We repeat relevant portions of the descriptions in this section, where appropriate, to give a sense of the experimental process used.

We experimentally evaluate the ITERATIVE-ESM-SELECTION algorithm on two TAC/SCM data sets. The first test is designed to confirm identification of a relabeled strategy in the original data set. The second explores a full strategy set from the Deep Maize 2008 candidate set. Once a partitioning is found for the full strategy set, we subsequently test reduced strategy sets to confirm that the full strategy set partitioning holds even in the reduced space. Before launching into analysis of TAC/SCM strategies, we illustrate a simple example.

4.1 Simple Equivalent Strategy Game

Consider the following two-player matrix game whose players are specified in Table 1. The column player has a single action, C, and the row player has three actions: A, B, and $\hat{B}$.

	C
A	$\alpha, 0$
B	$0, 0$
$\hat{B}$	$0, 0$

Table 1: Simple duplicate game

Say, for instance, that a simulator modeling this scenario adds zero-mean, unit-variance Gaussian noise to the row player's score. Discounting noise, the strategies B and $\hat{B}$ are equivalent. Additionally, when α is small A, B, and $\hat{B}$ are approximately the same. Using our partitioning scheme we have five basic strategy partitions: $A, B, \hat{B}$; $A, \{B, \hat{B}\}$; $\{A, B\}, \hat{B}$; $\{A, \hat{B}\}, B$; and $\{A, B, \hat{B}\}$. Clearly, we would like to discover that B and $\hat{B}$ are equivalent. Additionally, for some level of noise and setting of the α parameter, we would like A, B, and $\hat{B}$ to be considered equivalent to avoid over-generalization. Consider ESG models which are fit from a single observation of each profile in the Table 1. Because we know the noise distribution, we can compute $\mathbb{E}\mathscr{L}$ analytically. Each payoff observation is a χ-distributed random variable. The expected loss for each partitioning is given in Table 2.

Partition	$\alpha = 1$	$\alpha = 2$
$A, B, \hat{B}$	3	3
$A, \{B, \hat{B}\}$	2	2
$\{A, B\}, \hat{B}$	5/2	4
$\{A, \hat{B}\}, B$	5/2	4
$\{A, B, \hat{B}\}$	4/3	11/3

Table 2: The expected loss, $\mathbb{E}\mathscr{L}$, for various partitions and settings of α for the game in Table 1.

Given the expected loss, the optimal model for the game should consider A, B, and $\hat{B}$ equivalent strategies when α is small relative to the noise variance and only B and $\hat{B}$ equivalent otherwise. Note that in all cases, B and $\hat{B}$ are equivalent in the selected empirical game model, as desired, and when α varies we have a strict decision criterion which allows us to decide what magnitude of payoff differences should distinguish strategies.

4.2 TAC/SCM Analysis

In designing the Deep Maize 2008 strategy, a number of proposed variants underwent a large empirical game-theoretic analysis. The construction of the data set, as well as the methods used to evaluate each strategy after the observations, are reported in prior literature [6]. The TAC/SCM scenario defines a six-player simulation with programatically defined strategy sets. After each tournament, teams are encouraged to submit binary versions of their agent to the agent repository[2]. After initial analysis, five variants of Deep Maize 2008 remained as candidates for the final screening of strategies. These five strategies combined with top strategies from the 2007 tournament analysis comprise the strategy set used to select the Deep Maize 2008 tournament strategy. Table 3 lists the strategies used in this analysis.

Label	Description
PH	PhantAgent 2007
TT	TacTex 2007
DM6	Deep Maize 2008 variant 6
DM20	Deep Maize 2008 variant 20
DM24	Deep Maize 2008 variant 24
DM25	Deep Maize 2008 variant 25
DM28	Deep Maize 2008 variant 28

Table 3: The descriptions of the strategies used in the TAC/SCM data set.

Each simulation of a specific profile requires 7 processor hours on a cluster of computers (1 hour running simultaneously on 7 different processors). Observations are screened for defects such as network outages and other external factors which introduce nonstandard agent behavior. A full six-player, seven-strategy symmetric game has 924 distinct strategy profiles. Even using *demand adjusted profits* [6] to reduce variance in the observations' profit

[2]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`.

vectors, 30 samples per profile are still used for statistical analysis of the deviations and regret. This would require at minimum 194,040 processor hours to complete. On a typical day prior to the tournament, there were 700 processor hours available. Analysis was usually undertaken from April through the end of June. The full six-player analysis would require approximately three times the available amount of processing power assuming a zero failure rate. However, using a three-player hierarchical reduction [17], the number of distinct profiles is reduced to 84, a feasible size for analysis.

The Deep Maize test strategies were candidate strategies for the Deep Maize 2008 agent run in the 2008 TAC/SCM tournament. DM6 was a slight modification of the strategy used by the Deep Maize 2007 agent. DM6 replaced the 2007 procurement strategy by mimicking the PH long term procurement strategy. DM20 was a departure from DM6 that changed the procurement policy of the agent in the early portion of the game. DM24 and DM25 varied the mid-game procurement levels from that of DM6 and DM20. DM28 used a modified component price prediction algorithm, but was otherwise identical to DM20. The DM28 component price prediction algorithm decreased the prediction error of the DM20 predictor by about 1% RMS error [10]. While a 1% improvement is appreciable, it was not known whether the improvement would be expressed behaviorally in the relative score of the agent. Using these candidate strategies and a background set of agents that had strong support in equilibrium analysis [6], our team needed to make a decision about which candidate to play in the 2008 tournament.

The first experiment attempts to validate the equivalent strategies model for a reduced set of strategies and a duplicate. A reduced observation set was constructed from the observations which had profiles supported by strategies: PH, TT, and DM20. In order to replicate a duplicate strategy, we introduced a new strategy label: DM20C. For all observations involving strategy DM20, we changed the label to DM20C with probability 0.5. Using the resultant observation set, we ran the ITERATIVE-ESM-SELECTION algorithm. We expect the algorithm to identify DM20 and DM20C as equivalent. Table 4 verifies that the algorithm does indeed equate the two strategies. The table illustrates the algorithm as it progresses through various iterations. For each ESM model, we report the expected loss in millions. The single star highlights the best model in each round and the double star highlights the final model selected.

Round	Best	Strategy Space	RMSE
0		PH, TT, DM20, DM20C	3.64
		{PH, TT}, DM20C, DM20	12.93
		{PH, DM20}, TT, DM20C	11.27
		{PH, DM20C}, TT, DM20	10.68
		PH, DM20C, {TT, DM20}	4.30
		PH, {TT, DM20C}, DM20	4.24
	⋆	PH, {DM20, DM20C}, TT	3.57
1		PH, {DM20, DM20C}, TT	3.57
		TT, {PH, DM20, DM20C}	14.74
		{DM20, DM20C}, {PH, TT}	12.92
		PH, {DM20, DM20C, TT}	4.73
FINAL	⋆⋆	PH, {DM20, DM20C}, TT	3.57

Table 4: Iterative model selection on the TAC/SCM duplicate data set.

Building off of the validation in the previous experiment, we consider the full strategy set from Table 3. Because the ESM selected by the ITERATIVE-ESM-SELECTION algorithm depends on the observation set partition used by k-fold validation, different ESMs may be returned on different runs of the algorithm. Over thirty runs, five distinct ESMs were selected by the algorithm. Table 5 gives the frequency of these ESMs. The modal ESM contained the strategy equivalence classes: PH, TT, DM28, and {DM6, DM20, DM24, DM25}.

Equivalent Strategy Game	Freq
PH, TT, {DM6, DM20, DM24, DM25}, DM28	17
PH, TT, {DM20, DM28}, {DM6, DM24, DM25}	7
PH, TT, DM6, {DM20, DM24, DM25}, DM28	5
PH, TT, DM28, {DM20, DM24}, {DM6, DM25}	1

Table 5: ESM frequency table for TAC/SCM data set.

Figure 1 displays the modal ESM with the four distinct regions separated by the gray barrier. Three regions contain exactly one strategy, respectively. The fourth central region contains the four base strategies which the ESM identifies as equivalent: DM6, DM20, DM24, and DM25. The solid black lines interconnecting the four strategies represent the relative strengths of the pairwise equivalences as determined by their frequency in Table 5. For instance, strategies DM24 and DM25 appeared as equivalent in 29 of the 30 selected ESMs and the relationship is drawn with a thick line. Comparatively, strategies DM6 and DM20 appeared as equivalent in 17 of the 30 selected ESMs and the relationship is drawn with a slim line. The second most frequent ESM equated strategies DM20 and DM28 in 7 of the 30 selected ESMs. We denote the weaker equivalence tendency by the slim, dotted line.

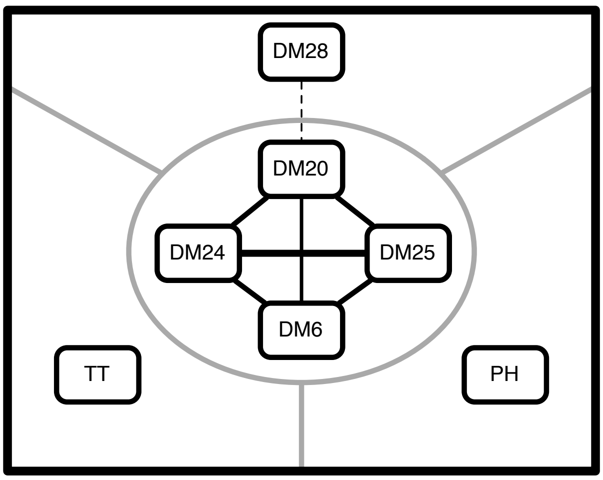

Figure 1: Modal ESP for TAC/SCM data set.

For the following experiments, we denote the equivalence class {DM6, DM20, DM24, and DM25} by DM*. We treat the modal ESM as the true ESM of the TAC/SCM simulation. We would like to know how restricting observations to a subset of the strategies affects analysis on the full strategy set. For instance, in designing and analyzing the strategies for a tournament, it often infeasible to sample the entire space of 20 or so candidate strategies, even using the hierarchical reduction technique. Often design and analysis proceed iteratively, adding strategies to a set of small background strategies. Bad strategies are pruned, while ones in the set of support of a sample Nash equilibrium are retained. New candidates strategies are added and the analysis process starts a new iteration.

There is no guarantee that strategies previously pruned would not be in the set of support of the new equilibrium.

Similarly, we would like to know if in a restricted strategy space, the ESM returned by the algorithm would hold against the full strategy set. There are two types of errors which may occur: a *false positive* (Type I) and a *false negative* (Type II). A Type I error occurs when strategies that should not be equivalent in the full strategy space are selected as equivalent in the restricted space. Type II errors occur when strategies that should be equivalent in the full strategy space are selected as distinct in the restricted space. For instance, if strategies DM6 and DM28 are in the same equivalence class, the restricted strategy ESM would contain a Type I error. If instead strategies DM24 and DM25 are in different equivalence classes, the restricted strategy ESM would contain a Type II error.

Size		Type I Error Rate			Type II
		PH	DM28	DM*	Error Rate
3	TT	0/5	0/5	0/14	
	PH		0/5	0/14	
	DM28			3/14	
	DM*				1/22
4	TT	0/10	0/10	0/20	
	PH		0/10	0/20	
	DM28			6/20	
	DM*				5/31
5	TT	0/10	0/10	0/15	
	PH		0/10	0/15	
	DM28			3/15	
	DM*				5/21
6	TT	0/5	0/5	0/6	
	PH		0/5	0/6	
	DM28			2/6	
	DM*				3/7

Table 6: Type I and II error rates on TAC/SCM data set for different strategy set sizes.

We design an experiment to explore the Type II and Type II error rates of the ITERATIVE-ESM-SELECTION in the TAC/SCM domain. For all of the strategies listed in Table 3, we create restricted strategy spaces of sizes 3 – 6. For each restricted strategy space of size n, we construct observation sets for each of the $\binom{7}{n}$ cases. For each of these observation sets, we run the ITERATIVE-ESM-SELECTION algorithm. We report the rate of each type of error in Table 6, using the modal ESM in Figure 1 as the standard for comparison. Note that since the equivalence classes for TT, PH, and DM28 are singletons, there can be no Type II errors for those respective classes. Because DM* is composed of multiple underlying strategies, we can observe Type II errors. This occurs with varying rates across the restricted strategy set sizes. For instance, in the restricted size 3 group, of the 22 simulations in which a Type II error could occur when an ESM is returned from the selection algorithm only one contained a false negative. There were no instances in any of the restricted simulations where TT or PH were incorrectly equated with any of the other strategies. Conversely, DM28 was incorrectly equated with at least one of the DM* strategies 21%, 30%, 20%, and 33% of the time for sizes 3, 4, 5, and 6, respectively. The values seem relatively large, however consider that even in the full case, 23% of the time DM20 and DM28 were in the same equivalence class.

5. DISCUSSION

We have proposed a formal model for evaluating the generalization risk incurred when modeling empirical games. We use an underlying simulation to provide the base game form and observation set. Our proposed model differs from previous models in that the empirical game modeling the underlying simulation may depart from the simulation's exact player and strategy sets. In particular, we discuss two forms of strategy set transformations:

- *Equivalent strategy models*: duplicate (equivalent) copies of strategies may exist in the simulation space. These strategies are identified and are equated in model space.
- *Factored models*: Strategies in the game model are composed of factors. These factors form strategically independent factor games whose utilities are an additive decomposition of the composite game.

For equivalent strategy models we propose an iterative algorithm, ITERATIVE-ESM-SELECTION, which is used to heuristically select an ESM for modeling a simulation. We experimentally evaluate the properties of the ITERATIVE-ESM-SELECTION algorithm on a data set of TAC/SCM observations. Using the algorithm, we select an ESM to model TAC/SCM given the 2008 tournament candidate strategies for our agent. Our first test experimentally confirmed that the ITERATIVE-ESM-SELECTION algorithm identifies identical, relabeled strategies as equivalent. Our second test discovered an equivalence relation involving four out of the five DM candidate strategies. Additionally, our reduced strategy error rate tests confirm the equivalence classes are identified even in the reduced strategy case.

This result is particularly useful due to its implications in reducing search efforts required in TAC/SCM game-theoretic analysis. If ESMs hold in reduced strategy spaces, we can analyze small cliques of strategies for strategy equivalences. If some equivalences are found, they are likely to hold in the full strategy set, thereby reducing the need to test multiple strategies in the same equivalence class. Identifying equivalences can yield a substantial reduction in the required number of observations to fully analyze the scenario. Leading into the 2008 competition, DM20 was identified as a promising strategy. DM28 was a relatively late breaking addition to the candidate set. Most of the available cycles nearing the start of the tournament were devoted to comparing the difference between DM20 and DM28, which had similar support in a sample Nash equilibrium for the empirical game. DM20 was chosen over DM28 since it had been tested thoroughly, whereas DM28, being a relative last minute update, had undergone only minimal testing outside of pairwise comparisons. Had the results of the ITERATIVE-ESM-SELECTION algorithm been available, those cycles could have been devoted to testing other promising candidates.

Davis et al. [4] describe an algorithm for finding an additive type of independence structure (factoring) in games when the general game form is not known *a priori*. In addition, they show that an approximate factoring can be found in polynomial time. The algorithm assumes knowledge of the actual payoffs and does not account for noise which occurs in empirical analysis setting. Our complementary methodology allows for the discovered factored structure to be tested for generalization risk. Thus, there is potential for using the generalization risk calculations to guide the learning algorithm which proposes the factored structure.

In this analysis we have primarily discussed empirical game models for transforming the strategy sets. Another potentially interesting transformation would involve modifying the number of players in the game. Hierarchical game reduction [17] does precisely this, and has been shown to be useful, especially in TAC related settings [6]. Applying this generalization risk minimization technique to the problem of finding the optimal reduction could be fruitful.

Finally, search techniques for selecting observation sequences [7, 12, 13, 11] in empirical games have increased the size of the strategy space researchers can feasibly analyze. Jordan et al. and Walsh et al. try to estimate the value of an additional sample of a profile and select the profile which optimizes the value in expectation. These algorithms search over the profiles in the profile space to identify stable, low regret, profiles. Because researchers are primarily interested in stable profiles, these search algorithms will more frequently observe "worthwhile" regions of the profile space saving the observational cost of regions that are viewed as less fruitful. Another potential search technique would search over the space of empirical game models instead of directly searching the profile space, selecting profiles to refute the current candidate using generalization risk as the decision criterion. One benefit of this type of search is that if a compact structure exists for the simulation and it is discovered, payoff estimates are directly formed for regions of the space which are unobserved unlike with the existing search algorithms.

6. REFERENCES

[1] O. Armantier, J.-P. Florens, and J.-F. Richard. Empirical game theoretic models : Constrained equilibrium & simulation. Technical report, SUNY Stony Brook, Stony Brook, NY 11794-4384, March 2003.

[2] O. Armantier, J.-P. Florens, and J.-F. Richard. Approximation of Bayesian Nash equilibrium. *Journal of Applied Econometrics*, 23(7):965–981, November 2008.

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

[4] G. B. Davis, M. Benisch, K. M. Carley, and N. M. Sadeh. Factoring games to isolate strategic interactions. In *Sixth International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 431–437, Honolulu, 2007.

[5] S. G. Ficici, D. C. Parkes, and A. Pfeffer. Learning and solving many-player games through a cluster-based representation. In *Twenty-Fourth Conference on Uncertainty in Artificial Intelligence*, pages 187–195, Helsinki, 2008.

[6] P. R. Jordan, C. Kiekintveld, and M. P. Wellman. Empirical game-theoretic analysis of the TAC supply chain game. *Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 1188–1195, 2007.

[7] P. R. Jordan, Y. Vorobeychik, and M. P. Wellman. Search for approximate equilibria in empirical games. *Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 1063–1070, 2008.

[8] C. Kiekintveld, M. P. Wellman, and S. Singh. Empirical game-theoretic analysis of chaturanga. *AAMAS-06 Workshop on Game-Theoretic and Decision-Theoretic Agents*, 2006.

[9] J. Niu, K. Cai, S. Parsons, E. Gerding, and P. McBurney. Characterizing effective auction mechanisms: Insights from the 2007 TAC market design competition. In *Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 1079–1086, Estoril, Portugal, 2008.

[10] D. Pardoe and P. Stone. The 2007 TAC SCM Prediction Challenge. In *AAAI-08 Workshop on Trading Agent Design and Analysis*, July 2008.

[11] A. Sureka and P. R. Wurman. Using tabu best-response search to find pure strategy Nash equilibria in normal form games. In *Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 1023–1029, Utrecht, 2005.

[12] Y. Vorobeychik and M. P. Wellman. Stochastic search methods for Nash equilibrium approximation in simulation-based games. In *Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 1055–1062, Estoril, Portugal, 2008.

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

[14] W. E. Walsh, R. Das, G. Tesauro, and J. O. Kephart. Analyzing complex strategic interactions in multi-agent systems. In *AAAI-02 Workshop on Game-Theoretic and Decision-Theoretic Agents*, Edmonton, 2002.

[15] M. P. Wellman. Methods for empirical game-theoretic analysis (abstract). In *Twenty-First National Conference on Artificial Intelligence*, pages 1152–1155, Boston, 2006.

[16] M. P. Wellman, A. Greenwald, and P. Stone. *Autonomous Bidding Agents: Strategies and Lessons from the Trading Agent Competition*. MIT Press, 2007.

[17] 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.

[18] M. P. Wellman, D. M. Reeves, K. M. Lochner, and R. Suri. Searching for Walverine 2005. In H. La Poutré, N. Sadeh, and S. Janson, editors, *Agent-Mediated Electronic Commerce: Designing Trading Agents and Mechanisms*, LNAI 3937, pages 157–170. Springer, 2006.