A GAME THEORETIC BIDDING AGENT FOR THE AD AUCTION GAME

Yevgeniy Vorobeychik
Sandia National Laboratories, Livermore, CA, U.S.A.

Keywords: Bidding agents, Keyword auctions, Game theory.

Abstract: TAC/AA (ad auction game) provides a forum for research into strategic bidding in keyword auctions to try out their ideas in an independently simulated setting. We describe an agent that successfully competed in the TAC/AA game, showing in the process how to operationalize game theoretic analysis to develop a very simple, yet highly competent agent. Specifically, we use simulation-based game theory to approximate equilibria in a restricted bidding strategy space, assess their robustness in a normative sense, and argue for relative plausibility of equilibria based on an analogy to a common agent design methodology. Finally, we offer some evidence for the efficacy of equilibrium predictions based on TAC/AA tournament data.

1 INTRODUCTION

Trading Agent Competition (TAC) is a successful forum for research into competitive agent design in an independent, highly complex, simulation environment. The ad auction game was recently introduced with a specific focus on several key strategic aspects of the keyword auction environment, carefully stylized into a TAC/AA simulation. We developed our agent to compete in TAC/AA, focusing primarily on a simulation-based game theoretic approach to enlighten bidding strategy.

There has been much discussion about the normative and descriptive value of Nash equilibria in actual strategic settings such as the one faced by a TAC/AA agent. Historically, the use of game theory has been relatively rare in agent design, even in the TAC tournaments (see (Wellman et al., 2006) for an exception). One reason that agent designers often eschew game theoretic techniques is that in general there may be many equilibria, and the problem of equilibrium selection requires coordination among the agents. Additionally, any asymmetric equilibrium requires coordination on roles. Finally, other agents may be imperfectly rational in a variety of ways (for example, buggy). These are valid issues which reveal considerable methodological uncertainty in operationalizing game theoretic techniques even if we believed them to be reasonable in a particular setting (i.e., when opponent agents are rational and attempt to maximize their payoffs). Our main contribution is to offer some general guidance to agent designers in operationalizing game theory, which we illustrate in the context of TAC/AA bidding strategy.

Our bidding strategy analysis restricts the consideration set to discretized linear strategies that compute a fraction of the myopic value per click to bid. We perform simulation-based game theoretic analysis in this restricted strategy space to (a) identify equilibria, (b) suggest equilibrium selection techniques, and (c) evaluate robustness of various possible strategies. We find, for example, that a particularly appealing equilibrium, one reached by iterative best response seeded with truthful bidding, is also very robust and is actually a best response to a range of reasonable opponent strategies.

Finally, we assess predictive value of equilibrium bidding policies derived using simulations based on actual tournament data, finding that predictions progressively improve over the span of the tournament, becoming relatively accurate on some measures.

2 THE TAC/AA GAME

The TAC/AA game features eight autonomous software agents representing advertisers in a simulated keyword (ad) auction. The advertisers interact with the environment by submitting bids and ads to be shown for a set of keywords over a sequence of 60 simulated days, each lasting 10 seconds. The environment itself is comprised of the *publisher* (search engine) agent, who collects the bids and ads from the

Vorobeychik Y..
A GAME THEORETIC BIDDING AGENT FOR THE AD AUCTION GAME.
DOI: 10.5220/0003140900350044
In *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence* (ICAART-2011), pages 35-44
ISBN: 978-989-8425-41-6

advertisers and displays the ads on the search page ordered by a ranking rule, as well as 90000 *users* who search for keywords, click on ads, and make purchases (conversions) from the advertisers. This interaction scheme between the agents is depicted visually in Figure 1. Next we describe in some detail the agent tasks and TAC/AA simulator implementation.[1]

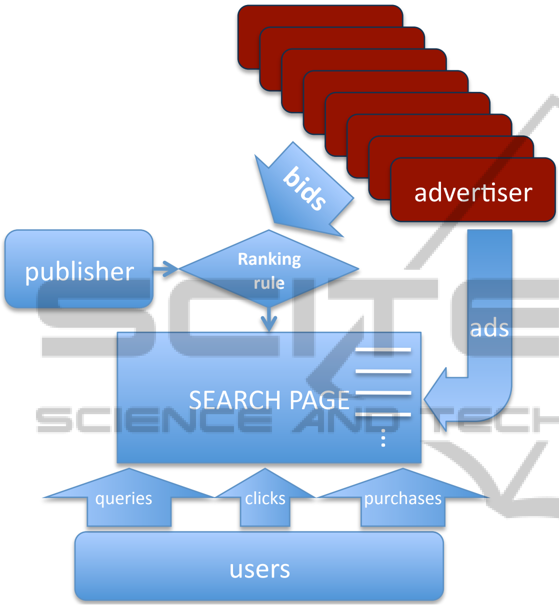

Figure 1: Schematic of the TAC/AA game.

2.1 Advertiser Agents

A TAC/AA advertiser agent plays a role of a retailer of home entertainment products. Each product is a combination of a *manufacturer* and a *component* (e.g., Lioneer TV). The game features three manufacturers and three components, for a total of nine products. While all advertisers are able to sell all products, every advertiser specializes in a manufacturer and a component. The manufacturer specialization yields a 1.5 factor increase in profits from sales, while component specialization results in a boost (roughly a factor of 1.5) in conversion rates.

An advertiser may submit a bid and an ad for any keyword and on any simulation day, to take effect on the following day. In addition, he may specify a *budget constraint* that limits spending for each keyword individually, as well as for an entire day. Only two ad types are allowed: *generic* (default) and *targeted*, which specifies a product. Advertiser's total payoff is the sum of his revenues from product sales less all per-click costs over the span of a simulation.

[1]For more details, see (Jordan and Wellman, 2009).

2.2 Publisher

The publisher has two tasks: ranking advertisers for each keyword and computing advertiser payments per click. An advertiser a is endowed at the beginning of a game (simulation) with a baseline click-through-rate (CTR) e_q^a for each keyword q, which is only revealed to the publisher. Given a collection of bids b_q^a, advertisers are ranked by a score $b_q^a(e_q^a)^\chi$, where $\chi \in [0,1]$ is chosen and revealed to advertisers at the beginning of each game.[2] The payments per click are determined according to the *generalized second-price (GSP)* scheme (Lahaie and Pennock, 2008). Specifically, suppose that advertisers are indexed by their rank (i.e., advertiser with $a = 1$ is ranked in the first slot). Then the payment of advertiser a is

$$p_a = \frac{b_q^{a+1}(e_q^{a+1})^\chi}{(e_q^a)^\chi},$$

that is, the score of the advertiser ranked immediately below, divided by his click-through-rate to the power χ. An exception to this payment rule arises when the reserve price r_a of a slot a (the slot in which a is placed) is higher than p_a, in which case the advertiser simply pays r_a. When an advertiser drops out due to saturating a budget constraint, rank and payments per click are recomputed for the remaining ads.

2.3 Search Users

Each of 90000 users has a specific product preference and will only purchase his preferred product. User preferences are distributed evenly across all products.

A user may submit three kinds of queries (keywords): *F0*, *F1*, and *F2*. A unique *F0* query specifies neither the manufacturer nor the component of the user's preferred product. Six *F1* queries partially reveal a user's preference: three specify only the manufacturer and three only the component of the desired product. Finally, nine *F2* queries completely reveal the user's preferred product (specify both the manufacturer and the component).

A user's behavior is determined by his "state". Indeed, a user may not even submit search queries, or may submit queries and click on ads with no intent to purchase. In the latter case, such "informational" users select uniformly among the three queries (*F0*, *F1*, or *F2*) to submit to the publisher. Finally, a "focused shopper" submits a query depending on his "focus level" (0, 1, or 2) corresponding to the three keyword classes above (thus, for example, a user in focus level 1 submits a *F1* query). A user in a focused

[2]See (Lahaie and Pennock, 2008) for a discussion of this class of ranking rules.

state makes a purchase (given a click) with conversion probability described below. Transitions between user states are Markovian, but non-stationary, as any user who actually makes a purchase is effectively “reset”. Consequently, user state distribution is affected by advertiser decisions.

After a keyword search, a user proceeds down the list of ads in a Markovian fashion, clicking on an ad he currently “views” with probability determined by the baseline CTR e_q^a of that ad, as well as the *targeting effect*, which is negative if the advertised product does not match a user’s preference, positive if it does, and neutral if the ad is generic.

Upon clicking an ad, the probability that a user subsequently makes a purchase depends on three factors: user’s state, advertiser’s specialty, and advertiser’s *capacity*. Users in an “informational” state may click on ads, but never make a purchase. A focused shopper will purchase with probability $\eta(I_d \pi_q, f_c)$, where

$$\eta(p,x) = px/(px + (1-p)),$$

π_q is a baseline conversion rate that depends on the keyword q, f_c is a factor that is 1.5 if the advertiser specializes in the component preferred by the user and 1 otherwise. Finally,

$$I_d = 0.995^{\left(\Sigma_{i=d-4}^{d} c_i - C^a\right)^+},$$

with C^a a capacity constraint of the advertiser, d current day, and c_i advertiser’s sales on day i. Note that the value of I_d on day d changes dynamically with each sale on that day (i.e., as c_d changes).

3 TOURNAMENT

15 participants registered for the TAC/AA tournament, which proceeded in three rounds: qualifying rounds, semifinals, and finals. No agents were eliminated in the qualifying rounds, as all were deemed competent enough to proceed. The eight top scoring agents from the semifinal round competed in the finals. The final ranking of the top agents is shown in Table 1. Our agent, QuakTAC, finished with the fourth highest score, a mere 1.25% below the third-place finisher and 2.38% below the second place.

Table 1: Final ranking for the TAC/AA tournament.

Rank	Agent	Average Score
1	TacTex	79,886
2	AstonTAC	76,281
3	Schlemazl	75,408
4	QuakTAC	74,462
5	munsey	71,777
6	epflagent	71,693
7	MetroClick	70,632
8	UMTac09	66,933

4 AGENT DESIGN

The decision environment in which a TAC/AA agent acts is very complex, with much uncertainty and decision interdependence between keywords and days. Thus, the process of designing and building an agent must of necessity involve two aspects: an analysis based on high-level strategic abstraction, as well as low-level implementation details. Our design of agent strategy (high level) had simulation-based game theoretic analysis at its core. To understand this analysis, however, we must first weave together some low-level details, as well as abstraction steps that were undergone before the corresponding game theoretic problem was appropriately defined.

First, we made a grand simplification in agent design by focusing almost exclusively on bidding strategy. As such, our budget was left always entirely unconstrained. Furthermore, we fixed the ad selection policy before any strategic analysis of bidding, hopeful that the specific ad choice has relatively low payoff impact (we revisit this assumption below).

4.1 Ad Selection

We choose a generic ad for a *F0* keyword and a targeted ad for all others. For a *F1* keyword, we choose the product in the ad to match the manufacturer/component in the keyword, while the missing product element of the keyword is filled with the advertiser’s specialty. The ad for the *F2* keyword matches the product in the keyword.

4.2 Bidding Policy

The problem of developing an effective bidding strategy in keyword auctions has received much attention in the literature, but there is relatively little practical evidence of efficacy of any of the proposed techniques. TAC/AA gives us an arguably objective, highly complex, yet still stylized forum to test bidding strategy development.

Perhaps the most natural approach to bidding in a complex multiagent setting like TAC/AA is via a combination of optimization and machine learning. Indeed, machine learning has enjoyed considerable success in TAC games historically (see, for example, (Pardoe and Stone, 2006)). Additionally, (Kitts and Leblanc, 2004) suggested computing a myopic (one-shot) profit maximizing bid given learned regression models of expected position and payment per click. One problem with learning-based approaches is that they do not prescribe what should be done in the absence of any information about the adversaries. Additionally, they assume that adversary behavior is stationary and, thus, past behavior is a good predictor of future behavior. In fact, learning may take some time before its prescriptions are effective, and the opponents will often be learning themselves, creating complex interactions between the learning algorithms, with policies that are unlikely to be stationary.

We steer away from learning-based approaches entirely, with our bidding policy determined by a *simulation-based equilibrium estimate*. We do so not to suggest that learning is a lost cause; rather, we follow a precise research agenda: *developing an agent that plays an equilibrium strategy alone allows us to directly measure the efficacy of a pure game theoretic approach*. Success of our approach will, thus, make a good case for equilibrium as *initial* prediction and strategic prescription, while further online exploration may or may not lead an agent to play other, more promising strategies.

In order to apply simulation-based game theoretic techniques to bidding, we need to first abstract the complex environment of TAC/AA into a computationally tractable restricted bidding strategy class. To this end, we make a dramatic simplification in considering bidding strategies which are linear in an estimate of an advertiser's value per click v, i.e., $b(v) = \alpha v$. The motivation for such a restriction comes from the literature on the theory of one-item auctions (Krishna, 2002), which often exhibits equilibria that are linear in bidder valuations, as well as other game theoretic treatments of far simpler models of keyword auctions (Vorobeychik, 2009). Note that this bidding function is entirely myopic, as it contains no temporal dependence (or any other state information about the game that may be available). On the other hand, it is very simple to implement and highly intuitive: an agent is asked to determine what fraction of his value he wishes to bid. Indeed, particularly due to the similarity of the *GSP* price mechanism to Vickrey auction, a very natural strategy would be to bid one's value, setting $\alpha = 1$. As we demonstrate below, this "truthful bidding" turns out to be a very poor strategy in our context.

While we have now a concrete class of bidding strategies to focus on, we have yet another question to answer before we can proceed to the actual analysis stage: as value per click is not directly given, how do we derive it from the TAC/AA specification and/or game experience? We devote the next section to this question.

4.3 Estimating Value per Click

A value per click of an advertiser a for a keyword q is the expected revenue from a click,

$$v^a = \Pr\{\text{conversion}|\text{click}\}E[R_q^a|\text{conversion}].$$

Revenue from a conversion depends entirely on whether the manufacturer in the keyword (user preference) matches the advertiser's specialty. If the manufacturer is specified in the keyword, the revenue is \$15 if it matches the specialty and \$10 otherwise. If not, the expected revenue is $15 \times \frac{1}{3} + 10 \times \frac{2}{3} = \frac{35}{3}$, as there is a 1/3 chance of a specialty match.

To compute the conversion probability, we need to estimate two things: the proportion of focused shoppers and the (expected) value of I_d. We begin with the former, assuming that an estimate of I_d is available. Since the proportion of focused shoppers actually depends on agent policies, we obtain an initial estimate using an arbitrary fixed policy, use the result to estimate bidding equilibria, and then refine the estimate using equilibrium bidding policies.[3] If we fix agent policies, the proportion of focused shoppers on a given day for a keyword q can be computed as the ratio of the empirical fraction of clicks that result in purchases and the estimate of conversion probability of a focused shopper. We average such empirical proportions for every simulation day over 100-130 simulations to obtain a daily estimate of expected proportion of focused shoppers for each keyword. We further average the resulting empirical proportions of focused shoppers over keyword classes (that is, over 6 *F1* keywords in one case and over 9 *F2* keywords in another). Thus, we have in the end empirical proportions of focused shoppers for the three classes of keywords, shown in Figure 2. Two features of this plot are worthy of note. First, the proportions are essentially the same for all keyword classes. This is not very surprising: there isn't a very strong a priori reason to believe that they would of necessity be different. Second, proportions follow a damped harmonic oscillation pattern. These oscillations are caused by the nonstationarity in the state transition process: a higher proportion of focused shoppers yield a higher conversion probability and, therefore, more sales, which result in the drop of conversion probability due to exhausted capacity

[3] In practice, it turned out that our estimates of focused shopper proportions were not very sensitive to the specifics of a bidding policy in our linear strategy space.

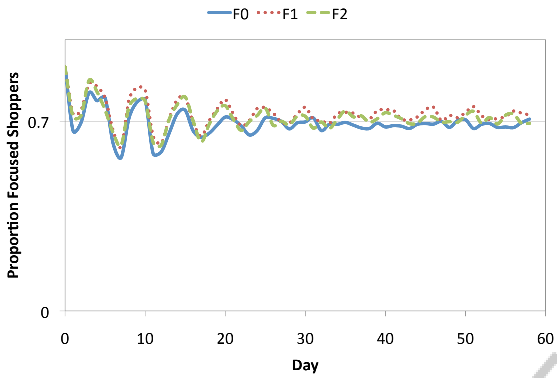

Figure 2: User proportions in the focused shopping state.

and also a drop in the fraction of focused shoppers. When the conversion probabilities are low, however, few transactions occur, increasing the proportion of focused shoppers. Interestingly, this process reaches a near steady-state at around the midway point of a game.

Suppose that we are estimating value per click on day d for tomorrow (day $d+1$). In order to compute the value of I_{d+1}, we need sales information for the three days that precede day d, as well as total sales for day d. Additionally, the value of I_{d+1} (and, hence, value per click) is not actually fixed but will change with every additional sale on day $d+1$. On day d we have exact information about sales on $d-1$, $d-2$, and $d-3$ based on advertiser sales reports that are provided at the beginning of each simulated day. Furthermore, we can estimate the expected sales on day d as the product of CTR, today's conversion rate, and the total number of impressions. While we know none of these exactly, we can estimate each with reasonable accuracy. First, we crudely estimate CTR as the average observed CTR throughout the game. Today's conversion rate requires estimating I_d, for which we have data from all days except current. We obtain a conservative (high) estimate for today's conversion rate by only using those "known" sales in computing I_d (which in the end underestimates value per click for $d+1$). The total number of impressions for each day of every keyword is estimated by running 100-130 simulations offline and averaging the number of observed impressions, using a fixed agent policy vector, just as in estimating focused shopper proportions. Next, we project total sales on day $d+1$ using again a conservative estimate of the conversion rate that would be effective at the beginning of that day. Finally, since value is roughly linear in I_{d+1}, we compute average I_{d+1} over each incremental sale made on day $d+1$.

5 SIMULATION-BASED GAME THEORETIC ANALYSIS

5.1 Equilibrium and Best Response Analysis

Having restricted our bidding strategies to be of the form $b(v) = \alpha v$, we use simulation-based game theoretic analysis to estimate an equilibrium in this strategy space. We note that an equilibrium estimated in the analysis actually plays a dual role, one predictive, describing what other agents will do, and one prescriptive, telling us how to optimally respond to that prediction.

In order to operationalize an equilibrium solution in the prescriptive context, we make a substantial further restriction and focus only on *symmetric* strategy profiles, that is, restrict all agents to follow the same bidding strategy $b(v)$. Hence, we use α to refer both to a specific bidding strategy and to a symmetric profile of these. There are two key reasons for restricting attention to symmetric profiles. First, an asymmetric equilibrium is difficult to operationalize, since it is not clear (when agents are ex-ante symmetric) which *role* our agent should play. Second, even if we pick a role for our agent, we still must assume that others coordinate on their respective roles just as we predict (at the minimum, no other agent may chooses our agent's role). Furthermore, we do not necessarily lose much by the restriction from the descriptive standpoint, since the agent ultimate cares about other players' choices only in the *aggregate*, insofar as they impact CTR and payments, and it seems reasonable that this is sufficiently captured by a symmetric equilibrium profile.

Since bids should be strictly positive to ensure any profit and, myopically, there is no reason to bid above value per click, we restrict α to the $(0,1]$ interval. Furthermore, to enable a more detailed analysis, we limit our equilibrium search to a discrete grid $\{0.1, 0.2, \ldots, 1\}$ (we also performed analysis "between" some of these grid points, but found that we do not lose very much due to our particular choice of discretization).

One major hurdle in equilibrium-based agent design is the issue of equilibrium selection. Since in our case equilibrium would offer both a prediction of opponent play and a best response to it, the goal, if we are to choose an equilibrium, is to choose one that yields the most plausible such prediction.

A common and highly effective technique employed in designing computational agents to compete against others is *self-play* (for example, Tesauro's

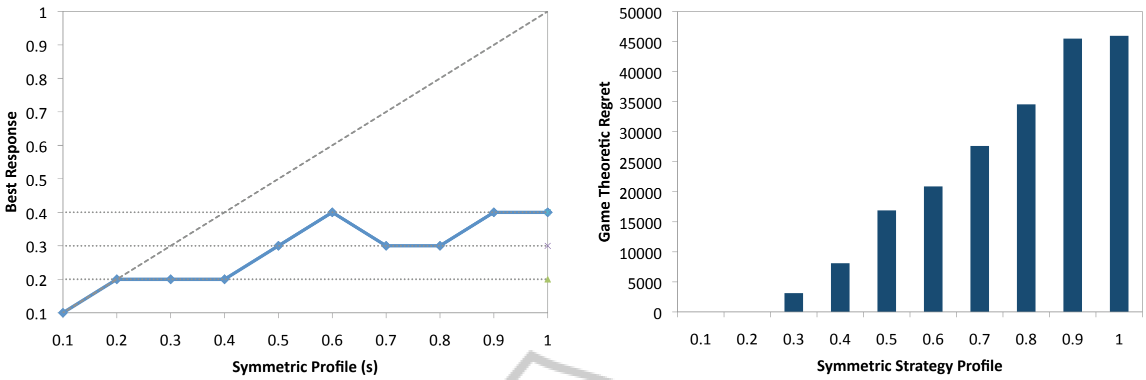

Figure 3: Best response function (left) and game theoretic regret (right) for all symmetric strategy profiles on the grid.

TD-Gammon agent was developed in such a way (Tesauro, 1995)). While this approach is usually applied at the level of individual game decisions when opponents move sequentially, we can detect a rough correspondence between self-play and a well-known *iterative best response* dynamic, where a player computes a maximizing action in each iteration assuming stationary opponents. In our case, iterative best response would proceed by first selecting a *starting* (seed) symmetric profile α^0, approximating a single-agent best response strategy $\hat{\alpha}^0$ to it, then setting the symmetric profile in the next iteration to be $\alpha^1 = \hat{\alpha}^0$. If this process converges and best responses are truly optimal, it necessarily converges to a Nash equilibrium α^*. The fact that the process can be viewed as roughly analogous to self-play suggests that equilibria found in such a manner may have good predictive properties, at least regarding the most competent of opponents. However, the dynamic itself is not sufficient: even if we believe other agents to follow a similar process, all need to agree on a starting point α^0. The choice of a starting point would, in general, be informed by whatever conventions govern typical algorithmic design in specific domains. In the context of auctions with one-dimensional valuations (such as our case), a rather focal starting point is truthful bidding, particularly so since *GSP* is reminiscent of Vickrey auctions which are, in fact, truthful. Hence, setting $\alpha^0 = 1$ seems a very reasonable way to seed a best response dynamic in a way that would lead to good predictions.

Following this approach, we obtained the equilibrium strategy for the purposes of the tournament via several iterations of best response dynamics starting at $\alpha = 1$. A look at Figure 3 (left) shows that a best response to a symmetric strategy profile with $\alpha = 1$ is $\alpha = 0.4$, and a best response to a symmetric profile with $\alpha = 0.4$ is $\alpha = 0.2$, which happens to be a symmetric equilibrium in our restricted policy space. Consequently, we were able to obtain a symmetric equilibrium for the restricted discrete bidding strategy space after only two best response iterations.

Based on the rapid convergence of iterative best response in our setting, we can make another conjecture: the equilibrium that we thus locate is relatively *robust* in the sense that the equilibrium strategy is a best response (or nearly so) for a number of other opponent strategies besides equilibrium. We suggest that this is another positive side-effect of considering best response dynamics in some settings. Another example of this phenomenon is a first-price sealed-bid auction with private valuations uniformly distributed on a unit interval, where the best response to truthful bidding in the linear strategy space is also a symmetric equilibrium strategy.[4] Figure 3 (left) demonstrates this robustness in our case: $\alpha = 0.2$ is a best response to 0.2, 0.3, and 0.4. Indeed, this figure additionally reveals another equilibrium at $\alpha = 0.1$, but it is only a best response to itself.

After the tournament we ran additional simulations to paint a more complete picture of the best response function in our discrete strategy space, which is depicted in Figure 3 (left), with payoffs for any configuration of agent strategies computed based on 15-30 simulation runs.[5] Several items are noteworthy from Figure 3. First, we may note that none of $\alpha > 0.4$ are ever a best response. This does not necessarily imply that these are poor strategies to play: it may be that an agent gains little by deviating from such a strategy, if all others jointly also play it. The corresponding measure of strategic stability, *game theoretic regret*, evaluates, for any strategy profile α the amount of gain an agent can obtain by unilaterally deviating when all others play the prescribed symmetric

[4] Of course, we make no general claims here, just offer some empirically motivated intuition and conjecture.

[5] The total number of runs we could execute was limited due to our experimental environment and the non-trivial running time of each simulation.

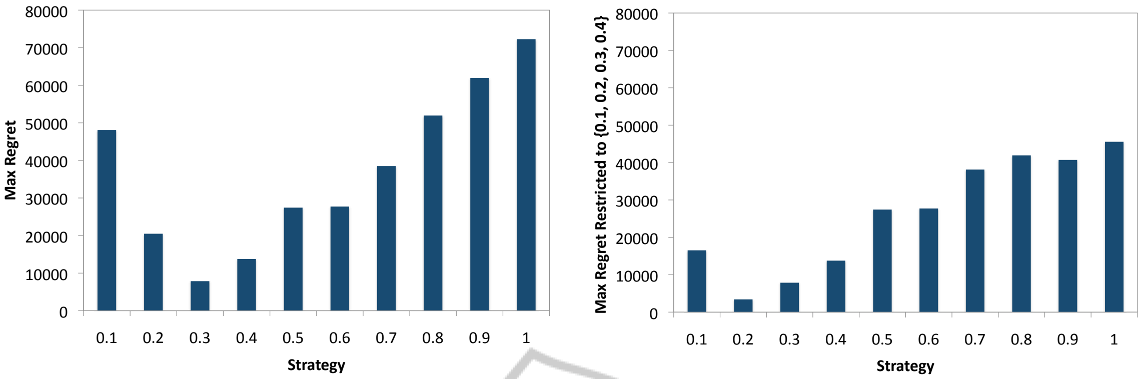

Figure 4: Max regret of each strategy on the grid against all strategies (left) and against a limited subset of "reasonable" opponents (right).

strategy:

$$\varepsilon(\alpha) = \max_{\alpha' \in [0,1]} u(\alpha', \alpha_{-1}) - u(\alpha),$$

where $u()$ is the symmetric utility function of each bidder (in our case, we estimate it by running simulations) and α_{-1} indicates that all players besides bidder 1 play the prescribed α, while α' denotes a deviation by bidder 1 (we pick bidder 1 arbitrarily here since every bidder is equivalent in this context by symmetry). Figure 3 (right) plots game theoretic regret of all symmetric strategy profiles in our discrete strategy space. This figure further, and rather soundly, confirms that any $\alpha > 0.4$ makes a very poor strategic choice, one unlikely to be adopted by even somewhat competent opponents. This observation alone already dramatically restricts our consideration set, and one may well use such knowledge derived from a game theoretic analysis to proceed with a more traditional machine learning approach. This is a rather important and general point: *game theoretic techniques may often be quite useful in restricting the number of options one needs to consider in learning, resulting, perhaps, in significant improvement in learning performance.* Another interesting observation is that the equilibrium $\alpha = 0.2$ is actually a best response to *nearly every reasonable strategy* (i.e., $\alpha < 0.4$, with $\alpha = 0.1$ being the lone exception) in our restricted space.

5.2 Robustness Analysis

We now turn to offer a methodology for a largely *prescriptive* game theoretic analysis, which is particularly salient in practical agent design settings like TAC/AA. This analysis is complementary to more traditional equilibrium analysis above, as it allows us (among other things) *to assess alternative equilibrium strategies*.

When deciding on a strategy for an agent in a multiagent system, an important consideration is robustness to uncertainty about opponent decisions. A common way to measure robustness of a particular strategy is via maximum regret, or the most that an agent would have gained by switching to another strategy, maximized over all opponent policies in a specific consideration set. Figure 4 (left) shows max regret of every strategy in response to our entire restricted consideration set, while Figure 4 (right) plots max regret when we restrict opponents to play only "reasonable" strategies. We can observe that $\alpha = 0.2$ fairs reasonably well in both cases; although $\alpha = 0.3$ and $\alpha = 0.4$ are more robust to very aggressive opponents (left), if we assume that all opponents are reasonable, $\alpha = 0.2$ has the smallest regret. In fact, good robustness property of $\alpha = 0.3$ in the unrestricted opponent setting actually prompted us to use that strategy, rather than 0.2, in the semifinal rounds, due to the risk that some of the agents competing at that stage are still rather unpolished (a prediction that proved correct). In contrast, the $\alpha = 0.1$ equilibrium has relatively poor regret properties in both settings. *The upshot of this discussion is that we can augment standard simulation-based game theoretic analysis with an analysis of max regret, as well as game theoretic regret, to allow us to best balance the risks from poor opponent strategy assessment with benefits of optimally responding to our predictions in a given setting.*

6 SOME ENHANCEMENTS

Our discussion above centered around an assumption that the same bidding policy (parametrized by α) is used for any keyword. If we use a state abstraction that captures all relevant strategic aspects of the environment, then there is no loss in utilizing a single bidding policy for all keywords. However, as we focus on policies that only use a myopic value per click, a relatively simple way to compensate for our restriction is to use different policies for different *keyword classes*. Thus, we may wish to use a bidding strategy that is a vector $<\alpha_{F0}, \alpha_{F1}, \alpha_{F2}>$, where each component prescribes the bidding strategy for the corresponding keyword. Another natural generalization is to contemplate *quadratic* bid functions. We implement the extension to quadratic bidding policies by specifying a value of α_{low} to use when $I_d = 0$ (and, hence, $v = 0$), and take the specified α (as above) to be applicable when $I_d = 1$ (value is maximal), with the restriction that $\alpha_{low} \geq \alpha$; the actual strategy is then a linear interpolation between these two extremes. This allows us to add only a single parameter, even while allowing different α for different keywords. The intuition for our special restricted class of quadratic bidding functions is that a higher fraction of value is submitted as a bid when value is low. This is motivated by the equilibrium structure of multiunit auctions (Krishna, 2002).

The equilibrium analysis that we had performed above had used a one-dimensional strategy space, and so estimating a best response did not require very much computation. By considering, instead, a four-dimensional strategy space, we make the problem of exhaustive sampling of the entire strategy space intractable due to the considerable simulation time required by each ad auction game. As a result, we can no longer implement iterative best response precisely as would be prescribed in an ideal setting. Rather, we simplify search the process by iterating one-dimensional best response dynamics sequentially over strategic parameters. Specifically, we proceed as follows. First, we ignore the strategy space extension and estimate an equilibrium in the one-dimensional strategy space as described above. This gives us α^*. We thereby set $\alpha_{F0} = \alpha_{F1} = \alpha_{F2} = \alpha_{low} = \alpha^*$ and proceed to sequentially explore individual strategic parameters, starting with α_{low}. More formally and generally, let $s = \{s_1, \ldots, s_L\}$ be a strategy factored into L parameters and suppose that s is initialized to s^0. We suggest the following algorithm for approximating an equilibrium in this factored strategy space:

1. Fix $s_l = s^0$, and perform best response dynamics *only allowing* s_1 *to vary*. Assume that best response dynamics converges (if not, we can terminate it after a finite number of iterations and select an equilibrium approximation based on some criterion from all the explored candidates) to $s_1 = s_1^*$. This gives us a new $s = \{s_1^*, s^0, \ldots, s^0\}$
2. Fix all strategic features at these values except s_2, and vary s_2 in another sequence of best response iterations
3. Repeat step 2 sequentially for all strategic parameters.

When we have completed the procedure above for all strategic parameters, we thereby have obtained $s^* = \{s_1^*, \ldots, s_L^*\}$. Note that s^* is not guaranteed to be an equilibrium, since we only vary a single strategic parameter at a time. Validating that such a procedure actually yields good equilibrium approximation is a subject of future work; for now, suffice it to say that its performance was quite satisfactory in the actual tournament, where we used $< 0.1, 0.2, 0.2 >$ and $\alpha_{low} = 0.3$, all obtained in this fashion.

7 ALTERNATIVE AD SELECTION POLICIES

Having assumed until now that our choice of ad policy is reasonable (and, moreover, that a specific ad policy has relatively little profit impact), we consider two simple alternative ad selection policies. The first, *Generic Ad Selection*, always chooses a generic ad. The second, *Specialty Ad Selection*, always chooses the ad to match the product to the advertiser's manufacturer and component specialty. Figure 5 shows that we were incorrect on one account: ad selection does make a significant impact on profits. Fortunately, the policy we actually used proved sensible, as it is significantly better than generic at the 0.9 confidence level and empirically (though not statistically significantly) better than the specialty ad selection policy.[6]

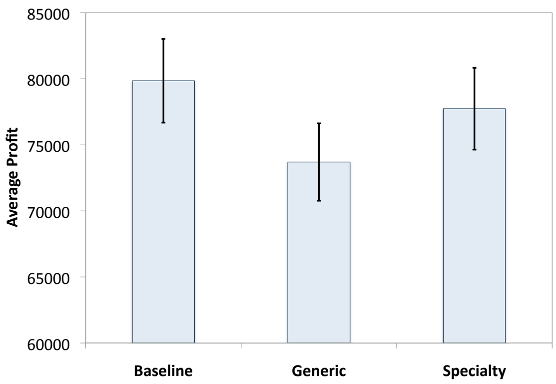

Figure 5: Payoffs of three ad policies (baseline is the one actually used in tournament). Error bars are confidence intervals at the 0.9 level.

[6]Indeed, since we evaluated the selection policies for an agent by fixing the policies of others to be our "baseline" described above, the baseline policy is shown to be an approximate equilibrium if we restrict the ad policy space to only these three options.

8 PREDICTIVE VALUE OF EQUILIBRIUM: EVIDENCE FROM TOURNAMENT

We close our discussion with some evidence about the descriptive quality of our approximate equilibrium policies from the TAC/AA tournament. In an ad auction, key determinants of an agent's profits are the distributions of CTRs and payments per click as functions of submitted bids. We use the data from tournament qualifying rounds, semifinals, and finals to see whether these distributions appear to converge to equilibrium predictions. We evaluate the error of an equilibrium prediction with respect to the tournament evidence about the distribution of some measure (say, payments per click) as follows. First, we bin all bids from simulated equilibrium and tournament experience of our agent into 50 intervals. For each bid interval, we compute the maximum error between the tournament and equilibrium distributions of the measure of interest (essentially, we use the Kolmogorov-Smirnov test statistic), and then compute the weighted average error over all bid intervals, with weights corresponding to the number of bids that fall into each interval.

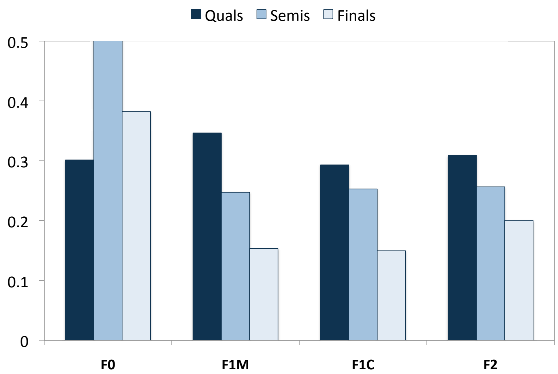

Figure 6: Average distance between equilibrium and realized (tournament) click-though-rate distributions.

The results, shown in Figures 6 and 7, show a clear downward trend in error as the tournament progresses: as agents become more competent on average, equilibrium prediction becomes increasingly accurate. At the minimum, this suggests that using equilibrium predictions as initial policies in the absence of other information can be quite effective.[7]

A similar, and much rosier picture for equilibrium

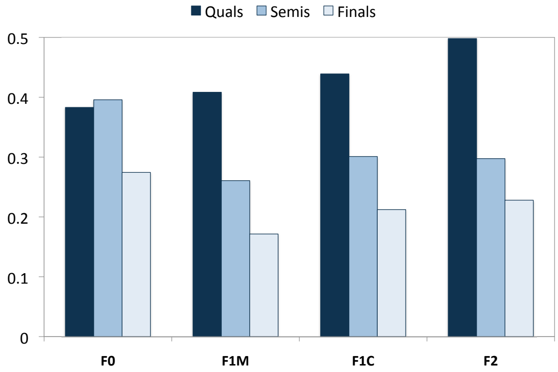

Figure 7: Average distance between equilibrium and realized (tournament) payment per click distributions.

prediction, is shown in Figure 8, where we look at average realized values of α observed in the tournament: there is a clear downward trend as tournament progresses, and the strategies in the final rounds are extremely close to equilibrium predictions.[8] We also note that average profits exhibit a similar trend, starting rather low (high α) and growing to near the levels predicted by the symmetric equilibrium at $\alpha = 0.2$ as the tournament progresses. This reveals, indirectly, that as agent pool becomes more competent, submitted bids are lower (on average), allowing bidders to realize higher profits.

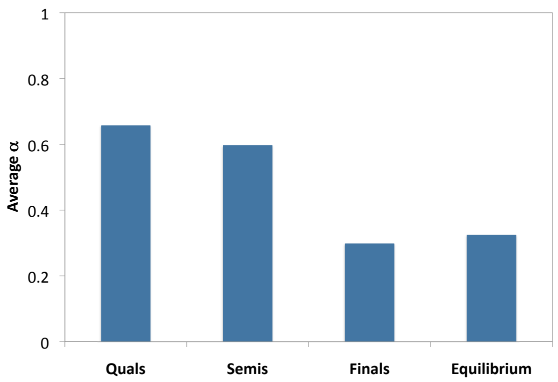

Figure 8: Average α values used by bidders throughout the tournament.

9 CONCLUSIONS

We demonstrate in concrete terms how to operationalize a pure game theoretic bidding strategy in a complex simulated keyword auction game, combining equilibrium analysis (which offers a combination of

[7]On the other hand, final errors are still non-trivial, so augmenting this approach with learning seems quite desirable.

[8]Of course, these aren't actual policies used, just our abstraction of them into the linear strategy space.

descriptive and prescriptive insights) with a purely prescriptive analysis based on robustness. All the analysis is done using simulations, as compared to more traditional game-theoretic analyses which usually involve mathematical treatments. Furthermore, in spite of the approximate nature of the resulting equilibria, we find that they offer very valuable predictions about the actual ad auction tournament bidding (as captured by submitted bids, as well as observed distribution of CTRs and prices). Finally, we offer an algorithm for equilibrium approximation when strategies are multi-dimensional, based on a sequence of single-dimensional analyses. In the process, we offer numerous general insights about operationalizing game theoretic approaches in practical agent design in multiagent systems.

REFERENCES

Jordan, P. R. and Wellman, M. P. (2009). Designing the ad auctions game for the trading agent competition. In *IJCAI-09 Workshop on Trading Agent Design and Analysis*.

Kitts, B. and Leblanc, B. (2004). Optimal bidding on keyword auctions. *Electronic Markets*, 14(3):186–201.

Krishna, V. (2002). *Auction Theory*. Academic Press.

Lahaie, S. and Pennock, D. M. (2008). Revenue analysis of a family of ranking rules for keyword auctions. In *Eighth ACM Conference on Electronic Commerce*, pages 50–56.

Pardoe, D. and Stone, P. (2006). Tactex-05: A champion supply chain management agent. In *Twenty-First National Conference on Artificial Intelligence*, pages 1489–1494.

Tesauro, G. (1995). Temporal difference learning and td-gammon. *Communications of the ACM*, 38(3):58–68.

Vorobeychik, Y. (2009). Simulation-based game theoretic analysis of keyword auctions with low-dimensional bidding strategies. In *Twenty-Fifth Conference on Uncertainty in Articial Intelligence*.

Wellman, M. P., Reeves, D. M., Lochner, K. M., and Suri, R. (2006). Searching for walverine-05. In *Agent-Mediated Electronic Commerce: Designing Trading Agents and Mechanisms*, LNAI 3937, pages 157–170. Springer-Verlag.