CAT 2008 Post-Tournament Evaluation: The Mertacor's Perspective

Lampros C. Stavrogiannis and **Pericles A. Mitkas**
Aristotle University of Thessaloniki, Faculty of Engineering, Department of Electrical and Computer Engineering
GR541 24 Thessaloniki, Greece
mertacor@olympus.ee.auth.gr

Abstract

TAC Market Design (or CAT) tournament is an effort to study the competition among modern stock exchanges trying to attract potential traders while maximizing their profit. This paper shortly presents Mertacor, our entrant for 2008, and makes an attempt to evaluate its performance. We compare Mertacor with the other available entries for the setting of CAT 2008 as well as beyond the tournament. What's more, we introduce a simple yet effective way of computing the global competitive equilibrium that Mertacor utilizes and discuss about its importance for the game.

Introduction

The invasion of the Internet in our daily life has given rise to new, innovative applications of electronic commerce. Common examples include electronic marketplaces and online trading systems that most of the major stock exchanges use to trade futures, options, equities as well as their derivatives.

The *double auction* (DA) is an auction where multiple buyers and sellers are able to make committed offers to buy and sell goods and then accept similar offers. Besides its prevalence in financial and commodities markets, many variants of the DA have been successfully applied as a solution to a multitude of resource allocation problems (Dash et al. 2007; Gomoluch and Schroeder 2003; Nisan et al. 1998), where different stakeholders compete against each other to obtain units of a scarce resource. The importance of DAs lies in the fact that they manage to exhibit a high allocative efficiency (ratio of traders' actual profit to their theoretical maximum profit) with the implementation of very simple rules.

Mathematicians, economists and computer scientists have long used game theory to analyze simple forms of this mechanism (Chatterjee and Samuelson 1983; Kagel and Vogt 1993; Satterthwaite and Williams 1993) but their findings have been criticized for being of scant relevance to practical scenarios due to their strict assumptions, like the independence of private values and traders' full rationality. Moreover, the dynamics of the CDA presents an important obstacle in any pure theoretical approach. This led to the adoption of simulation techniques where human subjects at first (Smith 1962) and software agents afterwards (Gode and Sunder 1993; Rust, Miller, and Palmer 1993) trade to verify the effectiveness of the mechanism. The use of multi-agent systems in this kind of experiments introduced a new scientific field, known as *Agent-based Computational Economics* (Tesfatsion 2002).

Each DA consists of two distinct aspects: its *structure* and its *behavior* (Vytelingum 2006). The latter is mainly occupied with the bidding strategies of the traders and has dominated related research. However, scientists have recently turned their attention to the structure of the DA, that is, to the rules and the protocols that govern every such auction. The majority of the relevant literature until now deals with isolated markets which operate free of charge. Nevertheless, in today's global economy, each country's market institutions compete with each other as well as with the remainder stock exchanges worldwide. Having recognized this, scientists from the universities of Liverpool and Southampton, and Brooklyn College introduced TAC Market Design (or CAT) tournament in 2007, in a joint effort to study the impact of dynamically changing mechanisms on trading.

This paper is organized as follows: Section 2 provides a short description of the CAT tournament. The notion of the global competitive equilibrium and how Mertacor manages to accurately estimate it are given in Section 3. Section 4 shortly discusses the strategies implemented by Mertacor for the games of 2008. Section 5 presents the results of our experiments, comparing Mertacor with its opponents. A brief summary concludes the paper in Section 6.

CAT Tournament 2008

The CAT game consists of two principal entities: *trading agents* (or *traders*) and *specialists*. Each trader may be either a buyer or a seller willing to exchange goods, whereas each specialist represents a DA market where these traders will trade. Trading agents are provided by the organizers and specialists are designed by the competition entrants. The platform of the tournament is JCAT, a client-server implementation of the Java Auction Simulator API (JASA), providing additional support for the operation of multiple markets (Niu et al. 2008b).

Traders are equipped with a *trading strategy* and a *market selection strategy*. The first determines their bidding behavior, the decision making process of selecting their offers (or shouts) in the market, and follows one of the four extensively studied strategies in the DA literature, namely ZI-C (Gode and Sunder 1993), ZIP (Cliff and Bruten 1997), RE (Roth and Erev 1995) and GD (Gjerstad and Dickhaut 1998). The market selection strategy specifies the specialist to register for their trades and is typically based on their profit from the market. Implemented market selection strategies in JCAT treat the selection as an n-armed bandit problem (Sutton and Barto 1998). Every trader is endowed with a set of goods to trade and a private value (the maximum amount willing to purchase or the minimum accepted sale's price for buyers and sellers respectively) for each of them. Both strategies and private values constitute personal information which is not revealed to the competitors during the game.

Each entrant owns a single exchange market and must effectively set its rules so as to meet his design objectives. Common questions to answer are: Which offers to accept in the market? How to match accepted offers? What should the price of each transaction be? How much to charge for every service provided?

Each game of CAT comprises several virtual trading days, each of which is further divided in trading rounds of fixed duration. At the beginning of the day, specialists announce their fees and traders must decide upon which market to select for the rest of this day. Traders' shouts are single-unit and persistent, meaning that every offer expresses the desire to trade one unit of the commodity and, once accepted, remains active until a transaction is executed or the end of the day is reached. Traders' private values are drawn from an unknown distribution at the start of the game and remain constant for the rest of it.

The daily evaluation of the entrants consists of three parts: the *market-share*, which is the percentage of the total traders' population registered in the market, the *profit-share*, which is the ratio of the daily profit a specialist obtains to the profit of all specialists, and, finally, the *transaction success rate* (TSR), which is the percentage of the shouts accepted that result in transactions. The daily score of each specialist is the mean value of the above metrics. Assessment commences and terminates in randomly selected days and total score is the sum of the scores across these days (Gerding et al. 2007).

The Global Competitive Equilibrium

The *global competitive equilibrium* is the competitive equilibrium of the equivalent single global market where all the buyers and sellers would trade had it not been their splitting due to the existence of multiple specialists. In an efficient global allocation only globally intra-marginal traders (buyers and sellers with private values above and below the price of the global competitive equilibrium respectively) transact.

However, the diffusion of the traders in the various specialist markets provides the opportunity for the globally extra-marginal traders to conduct transactions either for the reason that they might be intra-marginal ones for the market registered with or because of the inability of the specific specialist's rules to prevent extra-marginal trades, thus leading to a drop in both global and market's (for the latter case) allocative efficiency. It is therefore to the entrant's interest to identify this equilibrium and coordinate its transaction prices with it.

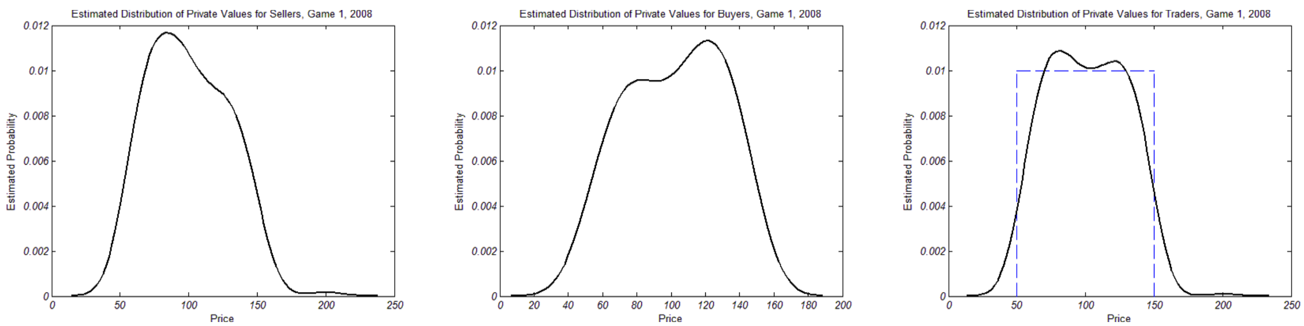
To estimate this point Mertacor continually keeps track of the highest bids (buy offers) and lowest asks (sell offers) submitted in its market. These prices constitute the closest available estimation of traders' private values. Moreover, the number of goods traded every day in the past consists a very accurate estimation of their daily endowment. When a sufficient number of trading agents have been explored, Mertacor forms the global cumulative demand and supply curves and computes the desired competitive equilibrium pair of price and quantity. This threshold was set at 80% of the total traders' population for the games of 2008. In addition, Mertacor exploits the possibility of subscription provided by CAT, gaining access to the shouts placed in the opponent markets, thus accelerating the process.

Table 1 illustrates the expected and estimated values of the global competitive equilibrium price for the three final games of CAT 2008 (real prices are not available but organizers provided us with the distributions of private values after the end of the tournament). As can be seen, the mean absolute percentage error for our estimation is less than 2% in all cases, validating the effectiveness of our method. It is interesting to mention that our specialist managed to successfully estimate this competitive equilibrium since the fifth, eighth and sixth trading day for the three final games respectively.

Figures 1(a) and 1(b) show the estimated probability density functions (pdf) of the prices of the lowest accepted asks and the highest accepted bids respectively that were obtained by Mertacor's log files for the first final game of TAC Market Design 2008. The form of the pdf is similar for the remaining games. The estimation of the pdf is based on the Parzen window method, also called kernel density estimation (Bowman and Azzalini 1997). A closer look at these figures reveals a relative symmetry of the distributions around the expected global competitive equilibrium price (equal to 100). This is further clarified in Figure 1(c), illustrating the distribution of the prices of all the shouts recorded. As can be observed, the resulting price distribution is very close to the real uniform distribution of the private values, $U(50, 150)$. We believe that the form of this distribution is caused by the traders' strategy mix selected, which was identical for buyer and seller populations, and the fact that private values were drawn from the same distribution for buyers and sellers, leading to an almost equal mean profit margin for both trading sides.

Table 1: Expected and estimated global competitive equilibrium price for the final games of CAT 2008.

Game	**Private Values Distribution**	**Exp. Gl. CE Price**	**Est. Gl. CE Price**	**Abs. Percentage Error**
1	$U(50, 150)$	100	100.199	0.199%
2	$U(100, 200)$	150	152.411	1.607%
3	$U(70, 170)$	120	118.931	0.891%



(a) Asks (b) Bids (c) Shouts

Figure 1: Kernel density estimation of the shout prices recorded by Mertacor for the first final game of CAT 2008.

Agent Mertacor

In this section we shortly describe the strategies followed by our entrant, which was placed 5^{th} in the finals of 2008. Mertacor's policies are primarily based on the theory of microeconomics, combined with heuristic techniques that improve the agent's performance for the specific setting of the tournament. Our agent's main design objective is the acquisition of a satisfactory level of profit-share score, compromising its target for the market-share. Figure 2 illustrates Mertacor's architecture. As shown, there are five different policies along with the *auctioneer* and the *market client* parts.

The latter is the communication component of the agent, converting incoming messages to a comprehensible form for the specialist and, conversely, transforming Mertacor's decisions according to CATP, the message protocol of JCAT, and then transmitting the resulting information to the CAT server.

The auctioneer acts as a coordinator among the rest of the components, assembling and transferring all the information required by each of them. In addition, it undertakes the responsibility to compute the global competitive equilibrium. The successful estimation of this point initiates the steady-state behavior of the agent discussed here. Before that, Mertacor acts like a modified CDA market that has proved to be effective in a variety of settings.

The *quote-accepting policy* determines the shouts that will be accepted for potential transactions in the market. Its filtering behavior is crucial, as this component is mainly responsible for the TSR score of the specialist. Mertacor implements a *global equilibrium beating accepting policy*, allowing only globally intra-marginal trades to take place during the first rounds of the trading day, and subsequently switches to a policy that implements the NYSE rule, according to which received shouts must beat the quote (current best offer placed in the market).

After the acceptance of the qualified shouts, a specialist must select the pairs of bids and asks that will lead to transactions. This is the task of the *matching policy*, which implements the 4-heap algorithm (Wurman, Walsh, and Wellman 1998) in the case of Mertacor.

The price of each transaction is specified by the *pricing policy*. Our contestant uses a *uniform global equilibrium pricing policy* which sets the price of all the transactions at

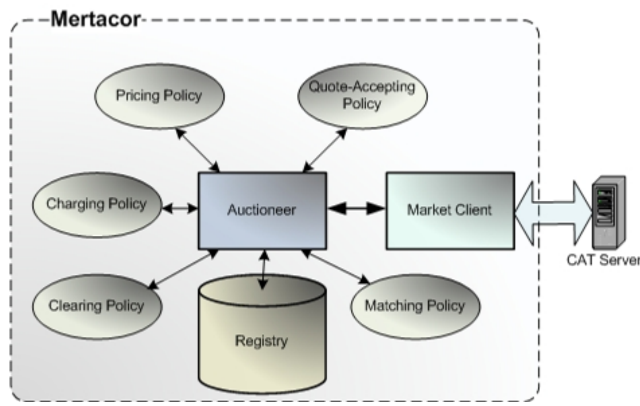
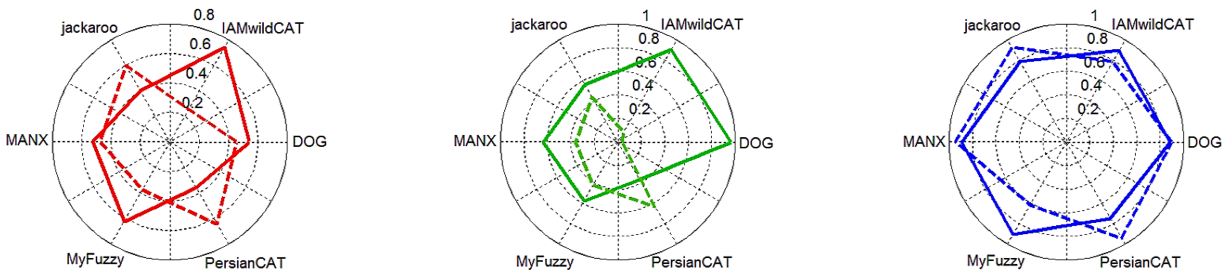


Figure 2: Mertacor's architecture.

the global competitive equilibrium price. This policy offers each individual the same profit that would be obtained in a global efficient allocation. For the last rounds of each day a modified version of the *side-based pricing policy*, originally introduced by IAMwildCAT (Vytelingum et al. 2008), is utilized, providing a higher amount of profit to the desired globally intra-marginal traders.

The time of the transactions is determined by the *clearing policy*. Mertacor uses a *round clearing policy* for the first rounds and then switches to a *continuous clearing* rule to increase the volume of its transactions.

Finally, the *charging policy* selects the type and the amount of the fees that registered traders should pay to obtain market services. There are four different kinds of fees in the CAT game: (a) the *registration fee*, charged for the registration of the traders in the market, (b) the *information fee*, for their access in an opponent market's accepted shouts and transactions executed, (c) the *shout fee*, for every shout placed, (d) the *transaction fee*, for every transaction carried out, and (e) the *profit fee*, which is a percentage of the profit obtained by each trader from a transaction. Our specialist charges a small profit fee, since we decided that only profitable traders ought to pay for their trades. Mertacor uses a



(a) Market-share (b) Profit-share (c) TSR

Figure 3: Tournament's evaluation metrics for the one-to-one experiments. The polar coordinates of each vertex of the dashed-line polygon represent opponent's score whereas the coordinates of the vertices of the solid-line polygon represent Mertacor's respective score.

Table 2: Results of one-to-one experiments for the games of CAT 2008. Each opponent is compared against Mertacor. The second value in each column refers to the respective mean score of Mertacor. Each experiment was repeated 9 times.

Opponent	Score	Efficiency	Conv. Coef.
DOG	0.465 - **0.797**	0.906 - 0.929	6.461 - 6.427
IAMwildCAT	0.382 - **0.848**	0.885 - 0.936	9.156 - 5.029
jackaroo	0.655 - 0.584	0.937 - 0.877	6.047 - 7.475
MANX	0.594 - **0.684**	0.913 - 0.935	7.414 - 6.943
MyFuzzy	0.473 - **0.701**	0.890 - 0.943	7.741 - 5.904
PersianCAT	0.738 - 0.493	0.952 - 0.853	4.396 - 7.905

limited score-based charging policy, keeping the fee in the interval $[0.1, 0.3]$, and setting its amount based on its market statistics and opponent scores. According to this policy, our entrant makes an attempt to beat better opponents in time intervals proportional to their score differences, also taking into account its rivals with lower cumulative score but higher daily score that might threaten its position in the game.

Mertacor's Evaluation

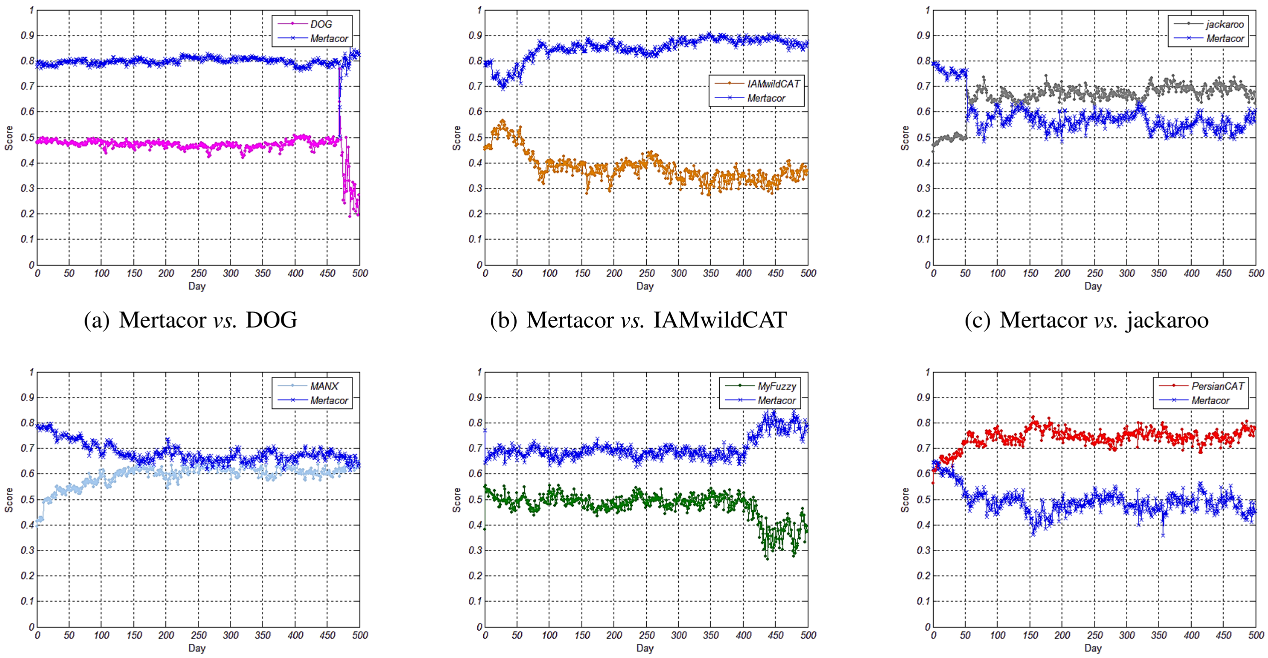
We conducted a number of experiments to evaluate our specialist's performance against its opponents. Our results are not statistically significant, considering the length of each game (four hours approximately), which presents an important limitation, as already mentioned in the analysis of (Niu et al. 2008a). We adopt a similar methodology to (Niu et al. 2007), originally introduced by (Tesauro and Das 2001) for trading strategies, comparing our specialist in homogeneous and heterogeneous market settings. We have used our post-tournament version of Mertacor, found as version 2 in the TAC agent repository. All specialists were obtained from the same repository, although we did not manage to include BazarganZebel, CrocodileAgent and PSUCAT because of their unstable operation.

Heterogeneous Markets

We ran a total of 18 experiments for the case of the heterogeneous markets. At first, we carried out one-to-one experiments where one market of each opponent competes with one Mertacor specialist. The remainder of the experiments concern one-to-many comparisons where five specialists of each entrant compete with one Mertacor and, conversely, five Mertacor markets operate against one opponent. Traders' strategy mix was identical to that of the games of 2008. More specifically, GD, ZIP, RE and ZI-C strategies were followed by 20%, 30%, 30% and 20% of the trading agents for the first game, 20%, 25%, 30% and 25% for the second game, and 15%, 30%, 35% and 20% of the total trader population for the third game respectively. All traders followed an ϵ-greedy market selection strategy (ϵ = 0.1, α = 1). Each experiment was repeated 9 times (3 iterations for each game). The duration of the games was 500 trading days and each day comprised 10 rounds. The performance criteria include the tournament's evaluation metrics as well as specialist's allocative efficiency and coefficient of convergence. The latter is proportional to the standard deviation of transaction prices around the market's competitive equilibrium price and constitutes a measure of their volatility.

One-to-One Experiments. The score of each competitor depends not only on its policies but also on the trader population and its opponents in a game. This kind of experiments isolates the influence of the specialists' mix on a market's performance, making a direct comparison between two competitors. Table 2 illustrates the mean values of score, allocative efficiency and coefficient of convergence for our comparisons with Mertacor.

These findings generally agree with the results of the tournament, except for the case of MANX, which was placed second in the finals, although this may be due to the minor modification of Mertacor's charging policy for our experimental setting. We found out that traders' strategy mix does not affect the quality of our results, as Mertacor is beaten by PersianCAT and jackaroo and continually wins over the other specialists. Figures 3(a)-3(c) show the market-share, profit-share and TSR in a polar system where the score of



(a) Mertacor *vs.* DOG

(b) Mertacor *vs.* IAMwildCAT

(c) Mertacor *vs.* jackaroo

(d) Mertacor *vs.* MANX

(e) Mertacor *vs.* MyFuzzy

(f) Mertacor *vs.* PersianCAT

Figure 4: Daily score results of one-to-one experiments.

each opponent is represented by the coordinates of its vertex in a dashed-line polygon and the respective score of Mertacor against each opponent is represented by the coordinates of the vertices of a similar solid-line polygon. The profit-share score shows that Mertacor succeeds in its objective for the profit, although PersianCAT obtains a higher profit-share because of its greater market-share and similar charging policy. As can be seen, TSR constitutes the weakest metric for Mertacor, being lower than the respective score for the majority of its rivals. Finally, there is an apparent impact of the opponent's selection on both allocative efficiency and coefficient of convergence, whose values are worse when Mertacor faces a stronger entrant and better for a weaker one.

It is also useful to see how specialists' total scores change on a daily basis, providing insights into the way that the duration of the game affects our findings. In addition, these graphs may help us segregate the start effect of the game, revealing specialists' steady state behavior. The results are illustrated in Figures 4(a)-4(f). Score differences are constant from the beginning of the game for the majority of the results but there seems to be a small convergence for Mertacor and jackaroo and an obvious one for the case of MANX. Figure 4(d) clearly illustrates the difference in the scores of Mertacor and MANX for the first days, mainly due to their different charging policies implemented. Mertacor exploits the fact that most of its opponents operate free of charge during the initial days, setting its fees from the start of the game. This has also helped PersianCAT to obtain a higher score difference since the second game of 2008, when it switched to an akin charging policy.

One-to-Many Experiments. In this experimental setting one specialist of each contestant is compared with many specialists of the same opponent. This kind of experiments demonstrates how a specialist might exploit the competition of the opponents in majority, revealing the benefits of its deviation from a homogeneous market markup.

Table 3 shows the results of the one-opponent-to-many-Mertacors experiments, evaluating opponents' performance against Mertacor. The mean values for the Mertacor refer to the best performing market of our specialist in terms of its total score. We have also included the mean ranking for the single specialist, providing more details on their ability to manipulate their Mertacor opponents. It becomes obvious from these results that PersianCAT and jackaroo dominate Mertacor in this kind of games, since they beat its markets in all of the experiments conducted. Mertacor was the winner for all of the remaining games except for the experiments with MANX. The results for the latter are ambiguous, given that it managed to win in one of the games and had a mean ranking of three. Finally, we observe a better performance of DOG than IAMwildCAT, even though the latter was placed above the former in the finals of 2008, thus uncovering a relative strength when it confronts our specialist.

The results of the experiments with one Mertacor versus many opponent markets are illustrated on Table 4. The values for the opponents in majority also refer to their score-maximizing specialist, whereas the mean ranking corresponds to the Mertacor competitor. The results are similar to the above except for the case of MANX and jackaroo. Mertacor takes advantage of the competition among MANX

Table 3: Results of one-to-many experiments for the games of CAT 2008. One opponent is compared with many Mertacor specialists. The second value in each column refers to the best-performing Mertacor market. The mean ranking corresponds to the single opponent specialist. Each experiment was repeated 9 times.

Opponent	Op. Rank	Score	Market-share	Profit-share	TSR	Efficiency	Conv. Coef.
DOG	6	0.351 - **0.424**	0.153 - 0.182	0.011 - 0.203	0.888 - 0.886	0.889 - 0.928	6.665 - 6.724
IAMwildCAT	6	0.316 - **0.429**	0.095 - 0.193	0.017 - 0.205	0.835 - 0.887	0.908 - 0.929	9.899 - 6.598
jackaroo	1	0.438 - 0.404	0.245 - 0.158	0.126 - 0.197	0.943 - 0.857	0.941 - 0.916	6.170 - 7.607
MANX	3	0.408 - **0.419**	0.184 - 0.182	0.102 - 0.189	0.939 - 0.886	0.926 - 0.934	7.233 - 7.234
MyFuzzy	6	0.279 - **0.419**	0.131 - 0.190	0.118 - 0.177	0.588 - 0.890	0.884 - 0.933	7.956 - 7.065
PersianCAT	1	0.482 - 0.388	0.274 - 0.152	0.232 - 0.175	0.938 - 0.837	0.952 - 0.903	4.630 - 7.410

Table 4: Results of one-to-many experiments for the games of CAT 2008. Many specialists of each opponent are compared with one Mertacor. The second value in each column refers to the best-performing opponent market. The mean ranking corresponds to the single Mertacor specialist. Each experiment was repeated 9 times.

Opponent	Mert. Rank	Score	Market-share	Profit-share	TSR	Efficiency	Conv. Coef.
DOG	1	0.364 - **0.686**	0.161 - 0.230	0.018 - 0.926	0.913 - 0.903	0.911 - 0.937	6.848 - 6.730
IAMwildCAT	1	0.375 - **0.683**	0.174 - 0.409	0.072 - 0.728	0.879 - 0.913	0.930 - 0.942	7.684 - 4.898
jackaroo	2.333	0.422 - 0.416	0.186 - 0.131	0.161 - 0.313	0.919 - 0.804	0.923 - 0.882	6.884 - 7.790
MANX	1.222	0.420 - **0.446**	0.173 - 0.174	0.158 - 0.301	0.930 - 0.862	0.924 - 0.931	8.378 - 8.072
MyFuzzy	1	0.329 - **0.502**	0.140 - 0.340	0.171 - 0.230	0.677 - 0.935	0.905 - 0.954	8.323 - 6.056
PersianCAT	5.889	0.430 - 0.339	0.187 - 0.124	0.194 - 0.137	0.910 - 0.755	0.926 - 0.852	5.919 - 8.035

specialists and wins in almost all of the games. This is true for the jackaroo opponents as well, where Mertacor was the winner of five games, obtaining a mean ranking of 2.333.

A closer look at the market-share and profit-share metrics of Tables 3 and 4 reveals an increased profit-share score per trader for Mertacor when it faces multiple homogeneous opponents for all of the results except for the games against PersianCAT. This mirrors the influence of our charging policy on Mertacor's score and is the main reason for its relative success over jackaroo for some of the last experiments, although the values of allocative efficiency and coefficient of convergence are better for the latter in these games.

Homogeneous Markets

In this kind of experiments all specialists implement the same policies, revealing their ability to cooperate and produce desirable global outcomes for the trading agents. This is very useful in cases where the designer owns all of the specialists. The metrics used for our evaluation in this setting include the *global allocative efficiency* and the *global coefficient of convergence*. The latter is proportional to the standard deviation of the prices from the global competitive equilibrium price divided by that price. The global allocative efficiency is defined as the ratio of the traders' actual profit to their theoretical maximum profit (obtained, according to microeconomic theory, when the price of all the transactions is set at the global competitive equilibrium price) had all the traders been in a single global market.

We ran 7 different experiments, one for each contestant. Each experiment was repeated 9 times and comprised 6 specialists and 240 ZI-C traders (120 buyers and 120 sellers). The selection of this strategy lies in the fact that these trading agents exhibit zero rationality, submitting offers at random, so we expect to obtain a lower bound for the allocative efficiency and an upper limit for the coefficient of convergence. Moreover, all trading agents follow an ϵ-greedy market selection strategy ($\epsilon = 0.1$, $\alpha = 1$). We have chosen not to use a random market selection strategy, as this would annihilate the influence of the specialists' charging policy on traders' movement among the markets and, consequently, on their global performance obtained.

Table 5 illustrates the results for these experiments. As shown, MANX, jackaroo and PersianCAT are the most globally efficient specialists, although the differences observed among markets are diminutive. Most important, the global allocative efficiency might be lower than its respective value

Table 5: Results for homogeneous markets populated by ZI-C traders following an ϵ-greedy market selection strategy. Each experiment was repeated 9 times.

Specialist	Global Efficiency(%)	Global Conv. Coef.
MANX	**93.516**	8.598
jackaroo	93.134	6.813
PersianCAT	93.081	**4.213**
Mertacor	92.542	8.094
IAMwildCAT	92.247	4.799
DOG	91.063	5.697
MyFuzzy	90.940	5.231

for a single market, as expected, but its mean value is above 90% in all cases, validating once again the effectiveness of the DA mechanism regardless of the traders' strategy mix utilized. PersianCAT presents a notably small value for the global coefficient of convergence, followed by the markets of IAMwildCAT and MyFuzzy, thus revealing a quick convergence of the transaction prices to the global competitive equilibrium. On the other hand, jackaroo and, particularly, MANX and Mertacor produced much higher values for this metric, despite being more efficient than the two last specialists above.

Conclusions and Future Work

In this paper we have shortly described CAT tournament as well as our agent's policies for the games of 2008.

Moreover, we have introduced a successful way of computing the global competitive equilibrium, constituting the most valuable component of our specialist's strategy. The importance of this point is twofold. From a market designer's perspective, the successful approximation of this point may help him meet his design objectives. The main challenge for a CAT specialist is to promote not only the quantity but also the quality of its traders' population, identifying and attracting the globally intra-marginal clients, thus increasing both its allocative efficiency and potential profit from trades. This computation along with the classification of the bidding strategies might be the key to the success. From the trading agent's view, this estimation might provide the opportunity to obtain novel bidding strategies for multiple markets. Moreover, the global competitive equilibrium could be utilized from an *arbitrageur* (trader that exploits the price difference of the same good exchanged in multiple markets, buying it low and then selling it high) to identify the most profitable stock exchanges for its trades irrespective of the markets' pricing policies implemented. However, we must examine how modifying private values' distribution, trading strategies and market rules might affect our estimation results.

In addition, we have provided a thorough analysis of the competition of CAT entrants against Mertacor. We have compared our agent in one-to-one and one-to-many, as well as in homogeneous market settings. We have concluded that the results of the tournament are in accordance to our findings, although there seems to be a slight divergence in the case of MANX, which deserves further investigation. The major problem with Mertacor is its low score of transaction success rate, which is possibly due to its quote-beating accepting policy for the last rounds of each day, allowing the submission of extra-marginal shouts. On the other side, Mertacor manages to obtain a respectable level of profit in the majority of the cases, being the second most profitable specialist in our experiments, thus accomplishing its main design objective.

The absence of a dominant strategy for the TAC Market Design setting provides each contestant with the incentives to attempt to improve his specialist's behavior. Hence, we intend to conduct more experiments to detect the main sources of Mertacor's inefficiency and eliminate them for our participation in the CAT tournament 2009.

Acknowledgements

We would like to thank the CAT organizing team for their support during the competition as well as for providing us with the tournament's parameter values.

References

Bowman, A. W., and Azzalini, A. 1997. *Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations*, volume 18 of *Oxford statistical science series*. Walton Street, Oxford OX2 6DP, UK: Oxford University Press.

Chatterjee, K., and Samuelson, W. 1983. Bargaining under incomplete information. *Operations Research* 31(5):835–851.

Cliff, D., and Bruten, J. 1997. Minimal-intelligence agents for bargaining behaviors in market-based environments. Technical Report HPL-97-91, Hewlett-Packard Labs.

Dash, R. K.; Vytelingum, P.; Rogers, A.; David, E.; and Jennings, N. R. 2007. Market-based task allocation mechanisms for limited capacity suppliers. *IEEE Transactions on Systems, Man, and Cybernetics - Part A* 37(3):391–405.

Gerding, E.; McBurney, P.; Niu, J.; Parsons, S.; and Phelps, S. 2007. Overview of cat: A market design competition. Technical Report ULCS-07-006, Dept. of Computer Science, Univ. of Liverpool.

Gjerstad, S., and Dickhaut, J. 1998. Price formation in double auctions. *Games and Economic Behavior* 22(1):1–29.

Gode, D. K., and Sunder, S. 1993. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy* 101(1):119–137.

Gomoluch, J., and Schroeder, M. 2003. Market-based resource allocation for grid computing: A model and simulation. In *Proceedings of the First International Workshop on Middleware for Grid Computing*, 211–218.

Kagel, J. H., and Vogt, W. 1993. Buyer's bid double auctions: Preliminary experimental results. In Friedman, D., and Rust, J., eds., *The Double Auction Market: Institutions, Theories and Evidence*. Cambridge, MA, USA: Perseus Publishing. chapter 10, 285–305.

Nisan, N.; London, S.; Regev, O.; and Camiel, N. 1998. Globally distributed computation over the internet - the popcorn project. In *Proceedings of the 18th International Conference on Distributed Computing Systems*, 592. Washington, DC, USA: IEEE Computer Society.

Niu, J.; Cai, K.; Parsons, S.; and Sklar, E. 2007. Some preliminary results on competition between markets for automated traders. In *Proceedings of the Workshop on Trading Agent Desing and Analysis (TADA-07)*.

Niu, J.; Cai, K.; McBurney, P.; and Parsons, S. 2008a. An analysis of entries in the first tac market design competition. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, volume 2, 431–437. Sydney, Australia: IEEE Computer Society.

Niu, J.; Cai, K.; Parsons, S.; Gerding, E.; McBurney, P.; Moyaux, T.; Phelps, S.; and Shield, D. 2008b. Jcat: a platform for the tac market design competition. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems*, 1649–1650. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Roth, A. E., and Erev, I. 1995. Learning in extensive form games: Experimental data and simple dynamic model in the intermediate term. *Games and Economic Behavior* 8:164–212.

Rust, J.; Miller, J. H.; and Palmer, R. 1993. Behavior of trading automata in a computerized double auction market. In Friedman, D., and Rust, J., eds., *The Double Auction Market: Institutions, Theories and Evidence*. Cambridge, MA, USA: Perseus Publishing. chapter 6, 155–198.

Satterthwaite, M. A., and Williams, S. R. 1993. The bayesian theory of the k-double auction. In Friedman, D., and Rust, J., eds., *The Double Auction Market: Institutions, Theories and Evidence*. Cambridge, MA, USA: Perseus Publishing. chapter 4, 99–123.

Smith, V. L. 1962. An experimental study of competitive market behaviour. *Journal of Political Economy* 70(2):111–137.

Sutton, R., and Barto, A. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press.

Tesauro, G., and Das, R. 2001. High-performance bidding agents for the continuous double auction. In *Proceedings of the Third ACM conference on Electronic Commerce*, 206–209. New York, NY, USA: ACM.

Tesfatsion, L. 2002. Agent-based computational economics: Growing economies from the bottom up. *Artificial Life* 8(1):55–82.

Vytelingum, P.; Vetsikas, I.; Shi, B.; and Jennings, N. 2008. Iamwildcat: The winning strategy for the tac market design competition. In *Proceedings of the 18th European Conference on AI*, 428–432.

Vytelingum, P. 2006. *The Structure and Behaviour of the Continuous Double Auction*. Ph.D. Dissertation, School of ECS, Univ.of Southampton, Southampton, UK.

Wurman, P. R.; Walsh, W. E.; and Wellman, M. P. 1998. Flexible double auctions for electronic commerce: theory and implementation. *Decision Support Systems* 24(1):17–27.