

An Analysis of Entries in the First TAC Market Design Competition

Jinzhong Niu, Kai Cai
Computer Science
Graduate Center, CUNY
{jniu, kcai}@gc.cuny.edu

Peter McBurney
Computer Science
University of Liverpool
mcburney@liverpool.ac.uk

Simon Parsons
Computer and Information Science
Brooklyn College, CUNY
parsons@sci.brooklyn.cuny.edu

Abstract

This paper presents an analysis of entries in the first TAC Market Design Competition final that compares the entries across several scenarios. The analysis complements previous work analyzing the 2007 competition, demonstrating some vulnerabilities of entries that placed highly in the competition. The paper also suggests a simple strategy that would have performed well.

1. Introduction

Auctions are widely used in solving real-world resource allocation problems, and in structuring stock and futures exchanges. As a result, the field of auction mechanism design has drawn much attention in recently years from economists, mathematicians, and computer scientists [2, 10]. In traditional auction theory, auctions are viewed as games of incomplete information, and traditional analytic methods from game theory have been successfully applied to some *single-sided auctions*, where a single seller has goods for sale and multiple buyers bid for the goods, and some simple forms of *double-sided auctions* (DAs), where there are multiple sellers and multiple buyers and both sides may make *offers* or *shouts*.

However, as, for example Friedman [6], has pointed out, DAs, and particularly *continuous double auctions* (CDAs),¹ are too complex to analyze in this way since at every moment, a trader must compute expected utility-maximizing shouts based on the history of shouts and transactions, and the time remaining in the auction. This difficulty led researchers to seek experimental approaches. Smith [18] pioneered this field and showed, through a series of experiments with human subjects, that even CDAs with just a handful of traders can give high efficiency and quick convergence to the theoretical equilibrium. Software agents

¹A CDA is a continuous DA in which any trader can accept an offer and make a deal at any time during the auction period.

armed with various learning algorithms and optimization techniques have been shown to produce outcomes similar to those obtained by human subjects [4, 9] and are capable of generating higher individual profits [5].

In parallel with the automation of traders, computer scientists have started to take the approach of automated auction mechanism design. Thus, Cliff [3] used a genetic algorithm to discover a completely new kind of auction with good convergence properties. Phelps *et al.* [15] showed that genetic programming can be used to find an optimal point in a space of pricing policies. Niu *et al.* [13] identified a mechanism that minimizes variation in transaction price, confirming the mechanism through an evolutionary exploration. Pardoe and Stone [14] suggested a self-adapting auction mechanism that adjusts auction parameters in response to past auction results.

All this work has one common theme — it all studies single markets or compares different market mechanisms indirectly. In contrast, not only do traders in an auction compete against each other, real market institutions face competition [17]. In addition, previous studies usually present comparisons of auction mechanisms in different proprietary settings which differ in available information, computational resources and so on. As a result, mechanisms are difficult to compare, and it is desirable to have some platform that allows multiple markets to compete against each other, and evaluates market mechanisms in a uniform way.

The Trading Agent Competition (TAC) Market Design Tournament [7], also known as the CAT game, addresses these very issues². Prior TAC competitions have competing trading agents that aim to maximize their payoffs by interacting in a single market. CAT games do just the opposite. Each entrant in the competition provides a *specialist* that regulates a market with a set of auction rules, and these specialists compete against each other to attract traders and make profits. Traders in CAT games are provided by the competition platform and each of them learns to choose the best market to trade in.

²The first competition was held in July 2007 and a second competition was held in July 2008.

In a previous paper, we analyzed those entries from the first CAT competition (CAT 2007) that are available in the TAC agent repository³, and tried to identify effective auction rules [12]. In this paper, we extend this previous work by more closely examining the relative strength and weakness of the specialist agents.

2. The CAT competition

CAT games are designed to allow markets to compete against each other in a direct fashion. Each market serves a set of *traders*, each of which makes shouts indicating either what it is prepared to pay to buy some good (a *bid*), or what it expects to be paid to sell such a good (an *ask*). A CAT game lasts a certain number of *days*, each day consists of *rounds*, and each round lasts a certain number of *ticks*, or milliseconds.

Each trader is assigned private values for the goods it will trade. For buyers, the private value is the most it will pay for a good. For sellers, the private value is the least it will accept for a good. The private values and the number of goods to buy or sell make up the demand and supply of the markets. Private values remain constant during a day, but may change from day to day, depending upon the configuration of the game server. Each trading agent is endowed with a *trading strategy* and a *market selection strategy*. The first specifies how to make offers, the second specifies which market to choose to make shouts in.

Specialists facilitate trade by matching shouts and determining the trading price in an exchange market. Each specialist operates its own exchange market and may choose its own auction rules — the aim of the CAT competition is to create a specialist that optimizes a particular set of measures, including *market share*,⁴ *profit share*,⁵ and *transaction success rate*.⁶ The specialist with the highest cumulative score — the sum of these three metrics — is the winner of a game. Specialists may have adaptive strategies such that the policies change during the course of a game in response to market conditions.

We developed JCAT [11], the platform that is used in the CAT competition.

2.1. Parameterized market mechanisms

An auction mechanism can be parameterized into components that each regulates an aspect of the market. The following gives a framework extending that in [22]:

³<http://www.sics.se/tac/showagents.php>.

⁴The number of traders attracted to the market relative to the total number of traders.

⁵The amount of profit made by a specialist relative to the total amount of profit made by all the specialists in a game.

⁶The number of shouts matched by the market relative to the total number of shouts placed in the market.

- *Matching* policies define the set of matching offers in a market at a given time.
- *Quoting* policies determine the ask quote and bid quote — indicators of where traders need to place asks and bid in order to trade — from existing asks and bids.
- *Shout accepting* policies judge whether a request by a trader to place a shout in the market should be accepted or rejected.
- *Clearing* conditions define when to clear the market and execute transactions between matched shouts.
- *Pricing* policies are responsible for determining transaction prices for matched shouts.
- *Charging* policies determine the charges that a specialist imposes on traders using the market. In JCAT, a specialist can set fees for registration with a specialist, for market information, for making an offer, for completing a transaction, and impose a tax on profits.

For example, the classic CDA mechanism is a combination of the following auction rules (without considering the charging component): the market is cleared whenever a new shout is placed; the market matches the highest bid with the lowest ask that it exceeds; the pricing policy picks the midpoint of a matched ask-bid pair; the quoting policy uses the lowest unmatched ask as the ask quote and the highest unmatched bid as the bid quote; and the shout accepting policy only accepts shouts that beat the corresponding market quote.

2.2. A simple, but powerful market design

We developed `MetroCat`, a market mechanism that instantiates this parameterized framework, based on the following insights about the CAT game.

- It is crucial to maintain a high transaction success rate, since this rate is not immediately affected by the performance of other markets in contrast to market share and profit share. Thus a strong shout accepting policy, which only allows those shouts that are likely to match with other shouts, is desirable.
- Registration and information fees should be avoided, for these fees cause losses to extra-marginal traders⁷ and drive them away. Keeping extra-marginal traders in the market allows them to contribute through their impact on market share.

⁷Traders that theory says should not trade at market equilibrium and will not trade in efficient markets.

- Moderate charges on shouts, transactions, and trader profit only impacts intra-marginal traders, and because of this they stay with the market as long as they can make a considerable amount of profit through transactions after covering fees.

These insights led us to develop a CDA-based market mechanism, which uses a history-based shout accepting policy, denoted as AH. AH is based on the GD trading strategy [8]. GD selects a price that maximizes expected payoff, assuming that, for a given ask price a :

- if another ask price $a' < a$ was offered and was not accepted by a seller, a would not be accepted either;
- if another ask price $a' > a$ was offered and accepted by a seller, a would have been accepted as well; and
- if a bid price $b > a$ was offered in the market, a would have been accepted.

Based on these assumptions, the probability of a being matched is calculated as:

$$\Pr(a) = \frac{\sum_{d \geq a} MA(d) + \sum_{d \geq a} B(d)}{\sum_{d \geq a} MA(d) + \sum_{d \geq a} B(d) + \sum_{d \leq a} RA(d)}$$

where: $MA(d)$ is the number of asks with price d that have been matched; $RA(d)$ is the number of asks with price d that were not matched, and $B(d)$ is the number of bids with price d . It is not realistic to keep a full history of shouts and transactions, so GD maintains a sliding window and only considers those shouts and transactions in the window. Computed like this, $\Pr(a)$ is a monotonically decreasing function, since the higher a is, the lower $\Pr(a)$. It is also assumed that when $a = 0.0$, $\Pr(a) = 1$, and there is a certain value u_a , when $a > u_a$, $\Pr(a) = 0$. The probability $\Pr(b)$ of a given bid being accepted is computed analogously.

AH uses exactly $\Pr(a)$ and $\Pr(b)$ to estimate how likely a shout is to be matched, and only accepts those shouts with a probability higher than a specified threshold $\lambda \in [0, 1]$. When it is close to 1, the restriction may become too tight for intra-marginal traders to be able to place shouts in the market. When it is close to 0, the restriction may become so loose that extra-marginal traders are able to place shouts which do not stand much chance of being matched. The former would cause both the market and the traders to lose part of the expected profit and lead those traders to leave, and the latter would cause a low transaction success rate. `MetroCat` uses $\lambda = 0.5$, which we found to be optimal for a game configuration similar to CAT 2007.

In addition to AH, `MetroCat` uses a simple charging policy that imposes low, fixed fees on shouts, transactions, and trader profit, and no charges on registration and in-

formation⁸. Since we developed the competition platform, `MetroCat` was not an entrant in the competition,⁹ but we have used it as a benchmark in our post-competition experiments.

3. The analysis of CAT entries

Trading competitions have been an effective tool in fostering innovative approaches and advocating enthusiasm and exchange among researchers [20, 21]. However, the competitions themselves usually cannot provide a complete view of the relative strength and weakness of entries. In a competition, the performance of one player closely depends upon the composition of its opponents and the competition configuration, and the scenarios considered are usually limited. Thus we typically turn to post-competition analysis to tell us which entries are most interesting.

Ideally, such an analysis will cover all possible scenarios, but this usually presents too large a possible space. As a result, a common practice is to deliberately select a limited number of representative strategies and run games corresponding to a set of discrete points in the infinite space, assuming that the results are representative of what would happen in the whole space were one to explore it [19].

There are two common types of approaches to post-competition analysis; *white-box* approaches and *black-box* approaches. A white-box approach attempts to relate the internal logic and features of strategies and game outcomes. A black-box approach considers strategies as atomic entities. Niu *et al.* [12] performed a white-box analysis of CAT 2007, and examined how the dynamics in the CAT games are affected by the policies of each entry and their adaptation over time. This paper takes a complementary black-box approach and also examines the weakness and strength of CAT 2007 entries against `MetroCat`.

3.1. Multi-lateral simulation

Now, a full analysis of a set of strategies can only be achieved by considering many runs (to eliminate randomness) of every possible combination of strategies. This is not feasible for the CAT competition where each game runs for around 5 hours (irrespective of the hardware — the length of each trading days is hardcoded at a constant that permits each specialist to take time to perform possibly complex computations — any reduction in this time would potentially distort the results). Inspired by ecological analyses like [1, 16] — in which more copies of successful strategies, and less copies of unsuccessful strategies are run for

⁸The fees `MetroCat` imposes on shouts, transactions, and trader profit are respectively 0.1, 0.1, and 10% during the post-tournament experiments described in later sections.

⁹Instead it was included in the JCAT source code provided to entrants in the first CAT competition to support the development of their entries.

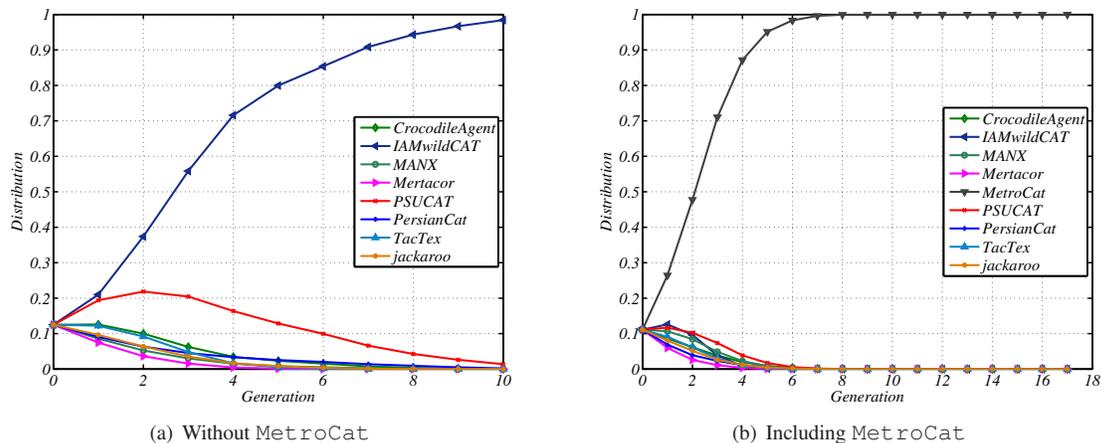


Figure 1: Simulation of CAT 2007 entries based on multi-lateral CAT games.

each successive game — but constrained by the number of specialists that we could have in a single game, we modified each strategy’s playing time in proportion to its score. That is, in a game that included all specialists, we decreased the number of trading days for less successful strategies, and increased the days for more successful strategies.

Figures 1(a) and 1(b) show the result of this simulation. The distribution on the y-axis shows the proportion of the total number of trading days for all markets that are allotted to each market, indicating how this evolves in populations without and with *MetroCat* respectively. Figure 1(a) shows that without *MetroCat* (i) the results of this analysis agrees with the results reported in [12], again confirming that *IAMwildCAT* was the strongest entrant in the 2007 competition; and (ii) the days allotted to *PersianCat* shrink more slowly than those of other losing specialists. This agrees with the results of bilateral games between *IAMwildCAT* and *PersianCat* (described below) and suggests that *PersianCat* was a strong entry. Figure 1(b) shows that with *MetroCat* (i) *MetroCat* quickly dominates the other entries, doing so faster than *IAMwildCAT* in Figure 1(a), and by generation 8 only *MetroCat* has any trading days; and (ii) the CAT 2007 champion, *IAMwildCAT*, loses trading days faster than other entrants after generation 1, indicating some weakness in its design when facing an opponent like *MetroCat*.

3.2. Bilateral simulation

One-on-one games closely examine the strength and weakness of a specialist when it faces different opponents. As a result, we ran 81 experiments in total between the 9 specialists including *MetroCat*. Nine of these are self-play games. Table 1 shows the resulting payoffs of specialists — their average daily scores — in these CAT games. Each

payoff is averaged over 10 iterations and entry (i, j) is the payoff of specialist i in the game against specialist j .

Figure 3 compares these payoffs pictorially using a polar coordinate system. Each plot shows the 9 specialists evenly distributed on the outer circle, the radial coordinates of the 9 vertices of the solid-line polygon represent a given specialist’s payoffs against all 9 specialists, and the radial coordinates of the 9 vertices of the dashed-line polygon represent its opponents’ payoffs in these games. The solid-line polygon and the dashed-line polygon overlap on the vertex that corresponds to the self-play game of the particular specialist. In Figure 3(a), the solid-line polygon completely encloses the dashed-line one, meaning *MetroCat* wins over all the other 8 specialists in bilateral competitions¹⁰. Figure 3(i) shows the opposite situation in that *Mertacor* loses all the games. The two polygons for any other specialist intersect somehow, showing their advantages in some games and disadvantages in others.

Both Figure 3(b) and 3(h) show that *IAMwildCAT*, the CAT 2007 champion, surprisingly loses (although barely) against *PersianCat* which placed sixth. This explains why in Figure 1(a) the days for *PersianCat* shrink more slowly than those for other specialists. *IAMwildCAT*’s loss, given the defeat of *PersianCat* by *PSUCAT* and *jackaroo*, suggests that *IAMwildCAT* has some particular weakness that is taken advantage of by *PersianCat*.

Other discrepancies, when compared to the results of the 2007 competition, include *jackaroo* (which placed 4th) winning over *PSUCAT* (2nd) and *Crocodile* (3rd). These may be significant, or may be caused by differences in the configurations for *PSUCAT* and *Crocodile* used in the sim-

¹⁰*MetroCat* maintains a better balance than those CAT 2007 entries between market share and profit share by keeping extra-marginal traders and preventing them from placing uncompetitive shouts.

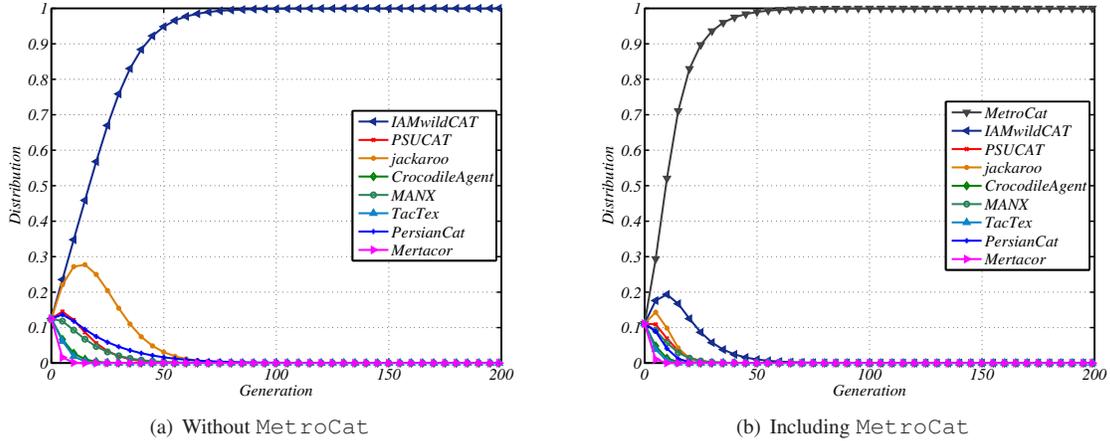


Figure 2: Simulation of CAT 2007 entries based on bilateral CAT games.

specialist	Metro	IAM	PSU	jack	Croc	MANX	TacTex	Persian	Mertacor
MetroCat	0.6451	0.7134	0.7461	0.7804	0.8217	0.7524	0.8592	0.7773	0.8885
IAMwildCAT	0.5895	0.6568	0.7207	0.6793	0.7681	0.7070	0.8008	0.6145	0.7632
PSUCAT	0.5366	0.5687	0.6152	0.5534	0.6950	0.6121	0.6420	0.7409	0.8307
jackaroo	0.4786	0.5926	0.6989	0.6279	0.7537	0.7088	0.7839	0.6902	0.8602
CrocodileAgent	0.4357	0.5245	0.5420	0.5145	0.4865	0.4614	0.6210	0.5879	0.7257
MANX	0.5383	0.5930	0.6067	0.5790	0.5101	0.6434	0.7150	0.6166	0.6944
TacTex	0.3362	0.4123	0.5743	0.4344	0.6271	0.5369	0.5546	0.6126	0.7238
PersianCat	0.4326	0.6200	0.5155	0.5925	0.7041	0.6686	0.6399	0.6446	0.7710
Mertacor	0.2677	0.3831	0.2947	0.3172	0.5068	0.4026	0.4479	0.4650	0.5503

Table 1: The payoff matrix of bilateral CAT games between CAT 2007 entries and MetroCat.

ulations and CAT 2007 games.

3.3 Ecological simulation

The payoff table for bilateral CAT games can be used to approximate ecological dynamics for populations involving more than 2 specialist types. The payoff of each specialist type for a certain population mixture is computed as the expected payoff for this specialist assuming that each specialist obtains the payoff it would have obtained had it competed one-on-one with each of the other specialists in the mix. Under this assumption, Figure 2(a) and 2(b) show how a population with an initial even distribution of specialists evolves over time when, as in [1], every specialist plays against every other specialist in every generation in bilateral games, and the number of specialists in any generation is proportional to the payoff achieved by that “breed” of specialist in the previous generation.

Comparing Figure 2(a) with Figure 1(a) and Figure 2(b)

with Figure 1(b), shows that while the winning strategies are the same, the ecological simulations based on multi-lateral games converge much faster than those based on bilateral games. This may be explained by the more epidemic effects of the strength of particular specialists in multi-lateral games compared with bilateral games. Another noticeable phenomenon is that PSUCAT performs much worse in the simulations with bilateral games than those with multi-lateral games, while jackaroo and IAMwildCAT do the opposite. These discrepancies indicate that, as one might expect, different game setups may lead to very different results. However, our results may be helpful to identify the weakness in strategies by looking at the particular scenarios in which a strategy performs poorly.

4. Conclusions and future work

This paper reports results of post-competition simulations of entries in the First TAC Market Design Competition

based on both bilateral games and multi-lateral games. The results basically agree with those in [12] and the result of the actual tournament, and also unveil weaknesses of specialists in particular scenarios, including the defeat of `IAMwildCAT` by `PersianCat` in bilateral games and the poorer relative performance of `jackaroo` in multi-lateral scenarios than in bilateral scenarios.

Some simulations also consider an additional specialist, `MetroCat`, which uses a history-based shout accepting policy that is derived from the GD trading strategy for double auctions. `MetroCat` claims victories in all the scenarios addressed in this paper, showing the importance of a shout accepting policy in a market mechanism.

The bilateral games and multi-lateral games can be viewed as the two ends of a spectrum of CAT games. The aim of running simulations based on both configurations is to explore whether the different competition configurations lead to results that differ very much. It is hoped that if they make no much difference — and our results suggest that they do not — the low cost of bilateral games can be used to approximate the games involving more different individual types and different population distributions.

Acknowledgments

The authors are grateful for financial support from the National Science Foundation under grant IIS-0329037 and from the EPSRC under grant GR/T10657/01.

References

- [1] R. Axelrod. *The Evolution of Cooperation*. Perseus Books Group, December 2006.
- [2] S. H. Clearwater, editor. *Market-Based Control: A Paradigm for Distributed Resource Allocation*. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 1996.
- [3] D. Cliff. Evolution of market mechanism through a continuous space of auction-types. Technical report, Hewlett-Packard Research Laboratories, Bristol, England, 2001.
- [4] D. Cliff and J. Bruten. Minimal-intelligence agents for bargaining behaviours in market-based environments. Technical report, Hewlett-Packard Research Laboratories, Bristol, England, 1997.
- [5] R. Das, J. E. Hanson, J. O. Kephart, and G. Tesauro. Agent-human interactions in the continuous double auction. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, Seattle, U.S.A., August 2001.
- [6] D. Friedman. The double auction institution: A survey. In D. Friedman and J. Rust, editors, *The Double Auction Market: Institutions, Theories and Evidence*, Santa Fe Institute Studies in the Sciences of Complexity, chapter 1, pages 3–25. Perseus Publishing, Cambridge, MA, 1993.
- [7] E. Gerding, P. McBurney, J. Niu, S. Parsons, and S. Phelps. Overview of CAT: A market design competition. Technical Report ULCS-07-006, Department of Computer Science, University of Liverpool, Liverpool, UK, 2007. Version 1.1.
- [8] S. Gjerstad and J. Dickhaut. Price formation in double auctions. *Games and Economic Behavior*, 22:1–29, 1998.
- [9] D. K. Gode and S. Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119–137, 1993.
- [10] P. Klemperer. How (not) to run auctions: The European 3G telecom auctions. *European Economic Review*, 46(4–5):829–845, 2002.
- [11] J. Niu, K. Cai, E. Gerding, P. McBurney, T. Moyaux, S. Phelps, D. Shield, and S. Parsons. JCAT: A platform for the TAC Market Design Competition. In Padgham, Parkes, Müller, and Parsons, editors, *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, Estoril, Portugal, May 2008. Demo Paper.
- [12] J. Niu, K. Cai, E. Gerding, P. McBurney, and S. Parsons. Characterizing effective auction mechanisms: Insights from the 2007 TAC Mechanism Design Competition. In Padgham, Parkes, Müller, and Parsons, editors, *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, pages 1079–1086, Estoril, Portugal, May 2008.
- [13] J. Niu, K. Cai, S. Parsons, and E. Sklar. Reducing price fluctuation in continuous double auctions through pricing policy and shout improvement rule. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 1143–1150, Hakodate, Japan, 2006.
- [14] D. Pardoe and P. Stone. Developing adaptive auction mechanisms. *ACM SIGecom Exchanges*, 5(3):1–10, April 2005.
- [15] S. Phelps, S. Parsons, E. Sklar, and P. McBurney. Using genetic programming to optimise pricing rules for a double auction market. In *Proceedings of the Workshop on Agents for Electronic Commerce*, Pittsburgh, PA, 2003.
- [16] J. Rust, J. H. Miller, and R. G. Palmer. Behaviour of trading automata in a computerized double auction market. In D. Friedman and J. Rust, editors, *The Double Auction Market: Institutions, Theories and Evidence*, Santa Fe Institute Studies in the Sciences of Complexity, chapter 6, pages 155–199. Perseus Publishing, Cambridge, MA, 1993.
- [17] A. Shah and S. Thomas. David and Goliath: Displacing a primary market. *Journal of Global Financial Markets*, 1(1):14–21, Spring 2000.
- [18] V. L. Smith. An experimental study of competitive market behaviour. *Journal of Political Economy*, 70(2):111–137, April 1962.
- [19] E. Sodomka, J. Collins, and M. L. Gini. Efficient statistical methods for evaluating trading agent performance. In *Proceedings of the 27th Conference on Artificial Intelligence*, pages 770–775, Vancouver, Canada, July 2007.
- [20] P. Stone and A. Greenwald. The first international trading agent competition: Autonomous bidding agents. *Electronic Commerce Research*, 5(2):229–265, 2005.
- [21] M. P. Wellman, S.-F. Cheng, D. M. Reeves, and K. M. Lochner. Trading agents competing: Performance, progress, and market effectiveness. *IEEE Intelligent Systems*, 18(6):48–53, 2003.
- [22] P. R. Wurman, M. P. Wellman, and W. E. Walsh. A parametrization of the auction design space. *Games and Economic Behavior*, 35:304–338, 2001.

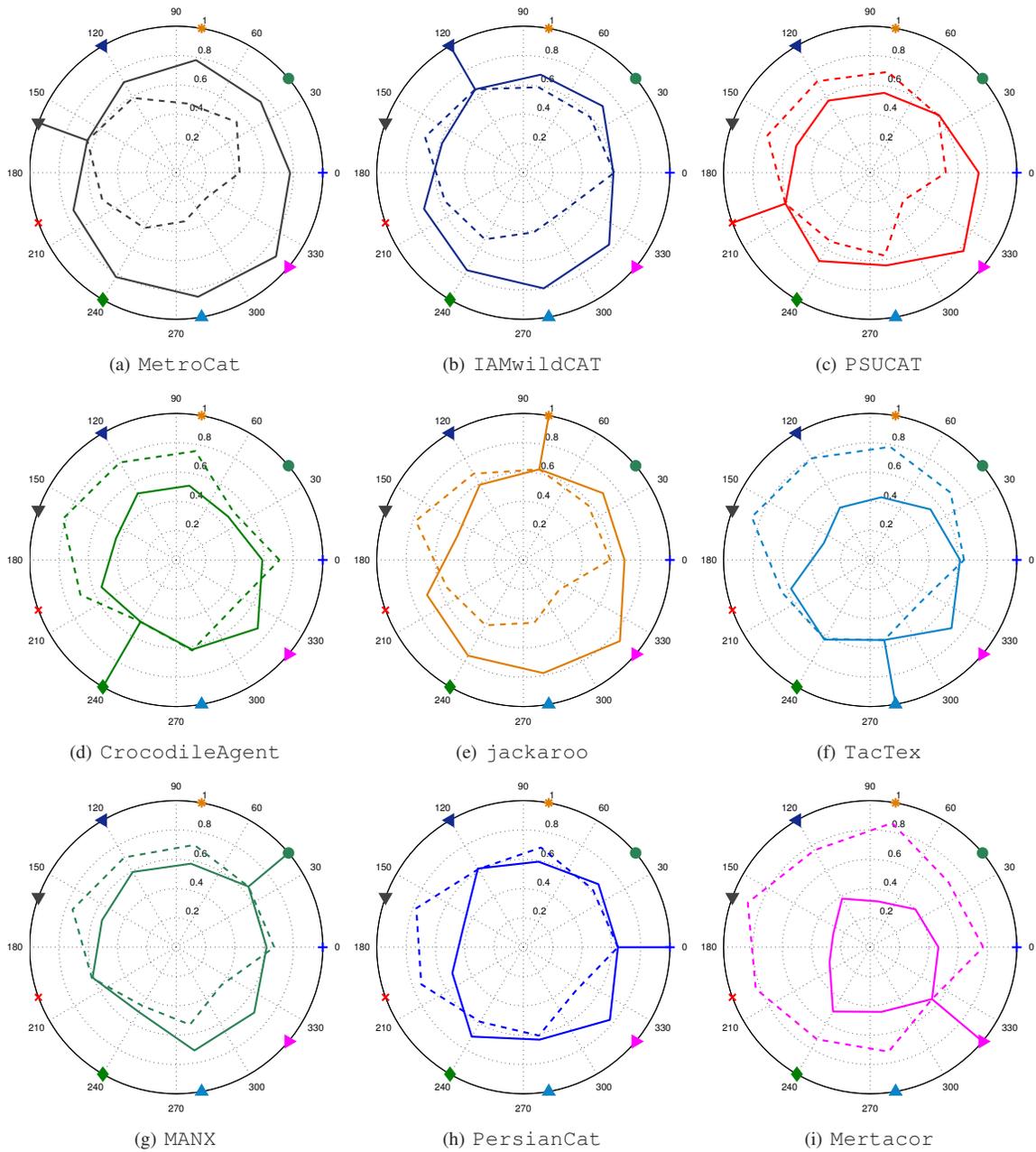


Figure 3: Payoffs of self and opponents in bilateral CAT games. On the outer circles starting from polar angle 0° lists the 9 specialists anti-clockwise : PersianCat (0°), MANX (40°), jackaroo (80°), IAMwildCAT (120°), MetroCat (160°), PSUCAT (200°), CrocodileAgent (240°), TacTex (280°), and Mertacor (320°). The radial coordinates of the 9 vertices of a solid-line polygon represent the corresponding specialist's payoffs against the 9 specialists respectively, and those of a dashed-line polygon represent payoffs of its opponents respectively. The overlapping vertex of the two polygons in each sub-figure is the self-play game for the particular specialist.