

# Competitive Benchmarking: Lessons Learned from the Trading Agent Competition

*Wolfgang Ketter, Andreas L. Symeonidis*

■ *Over the years, competitions have been important catalysts for progress in artificial intelligence. We describe the goal of the overall Trading Agent Competition (TAC) and highlight particular competitions. We discuss its significance in the context of today's global market economy as well as AI research, the ways in which it breaks away from limiting assumptions made in prior work, and some of the advances it has engendered over the past 10 years. Since its introduction in 2000, TAC has attracted more than 350 entries and brought together researchers from AI and beyond.*

Major advancements in artificial intelligence have been stimulated by well-designed competitions that tackle intriguing and often very complex problems. Chess, poker, stock trading, real-time strategy games, robot soccer, robot rescue or planning, and autonomous vehicles are among the most well known. Adaptability, proactiveness, and interoperability are essential characteristics of these games. Many of these competitions focus on development of rational behavior that would allow dominance over competitors. Apart from game play, though, one of the main reasons for these competitions is to bridge the gap between AI and real-life domains.

In many real-life domains, such as trading environments, self-interested entities need to operate subject to limited time and information. Additionally, the web has mediated an ever broader range of transactions, urging participants to concurrently trade across multiple markets. All these have generated the need for technologies that empower prompt investigation of large volumes of data and rapid evaluation of numerous alternative strategies in the face of constantly changing market conditions (Bichler, Gupta, and Ketter 2010). AI and machine-learning techniques, including neural networks and genetic algorithms, are continuously gaining ground in the support of such trading scenarios. User modeling, price forecasting, market equilibrium prediction, and strategy optimization are typical cases where AI typically provides reliable solutions. Yet, the adoption and deployment of AI practices in real trading environments remains limited, since the proprietary nature of markets precludes open benchmarking, which is critical for further scientific progress. To

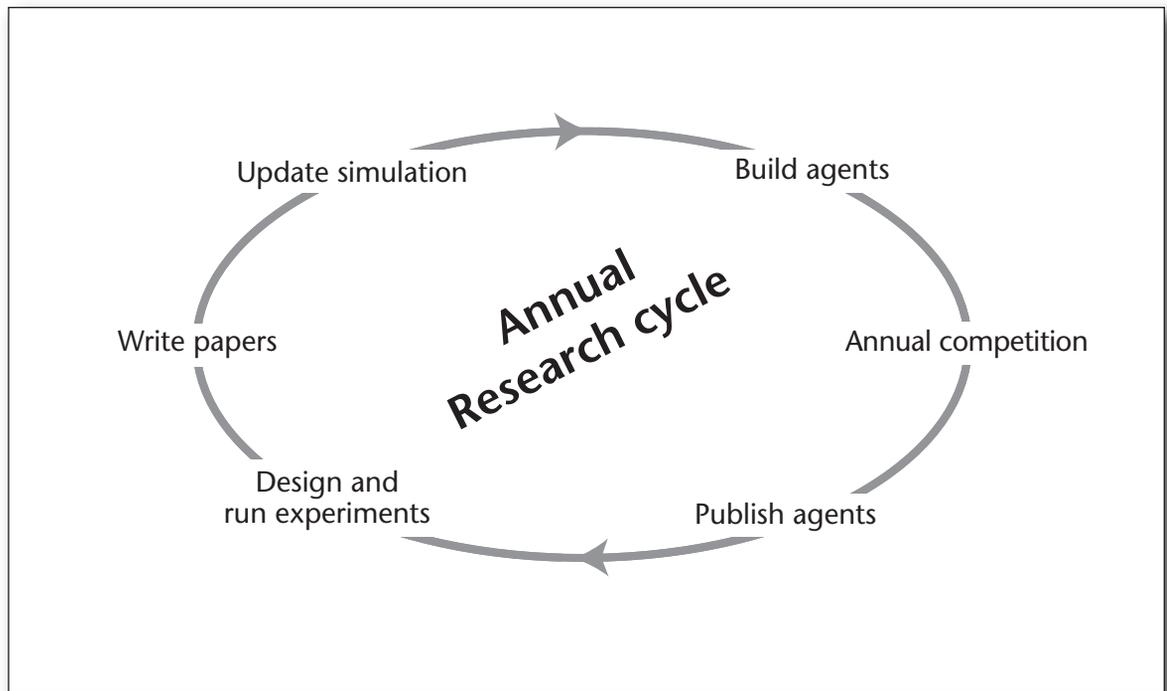


Figure 1. Annual Research Cycle for Competitive Simulation.

this end, the Association for Trading Agent Research has introduced the TAC framework.

### The TAC Framework

The Association for Trading Agent Research is a nonprofit organization aiming to promote research and education related to agents participating in trading environments. Issues related to trading agent architectures, decision making algorithms, theoretical analysis, empirical evaluations of agent strategies in negotiation scenarios, game-theoretic analyses, trading mechanisms, and market architectures are all of interest to the association (Wellman 2011).

The association fosters collaboration and innovation through the dissemination of competitive benchmarks through the Trading Agent and Design Analysis (TADA) workshop, typically colocated with a high-impact AI conference and the design of competitions similar to real-life economic environments (known as TAC), in order to provide robust benchmarks for researchers to test their trading strategies against other top performing research groups.

The Trading Agent Competition (TAC) was initially conceived to provide a platform for the study of agent behavior in competitive economic environments. Research teams from around the world develop agents for these environments. TAC is designed to run as an annual competition, a mod-

el that has been very effective in stimulating research. The basic annual research cycle is shown in figure 1. Each year, research groups build or update their agents and enter the competition. The competition is typically held in conjunction with a relevant major conference where participants can present their work, discuss what they have learned, and begin planning for the next competition cycle.<sup>1</sup> After the competition, teams are encouraged to release their agent code, so that all teams can design and run their own experiments using a range of state-of-the-art strategies and market design details. The results are published, and teams incorporate new insights into their agent designs for the following year.

TAC games aspire to address all facets of real-life trading, such as (1) challenges in the design and implementation of attractive and affordable product basket compilation, given specific customer preferences—TAC travel (Wellman, Greenwald, and Stone 2007), and (2) efficient strategies for increasing profit in a typical supply chain management scenario—TAC SCM (Collins, Ketter, and Sadeh 2010), (3) a stock exchange attracting as many traders as possible—TAC market design (CAT) (Cai et al. 2009), (4) adjustable broker policies with respect to expertise and capacity in sponsored keyword searches on the web—TAC ad auctions (TAC AA) (Jordan et al. 2009) and, (4) all of these, in a fully liberalized power market scenario—Power TAC (Ketter et al. 2011). Each year, the simulation may be updated to add new challenges

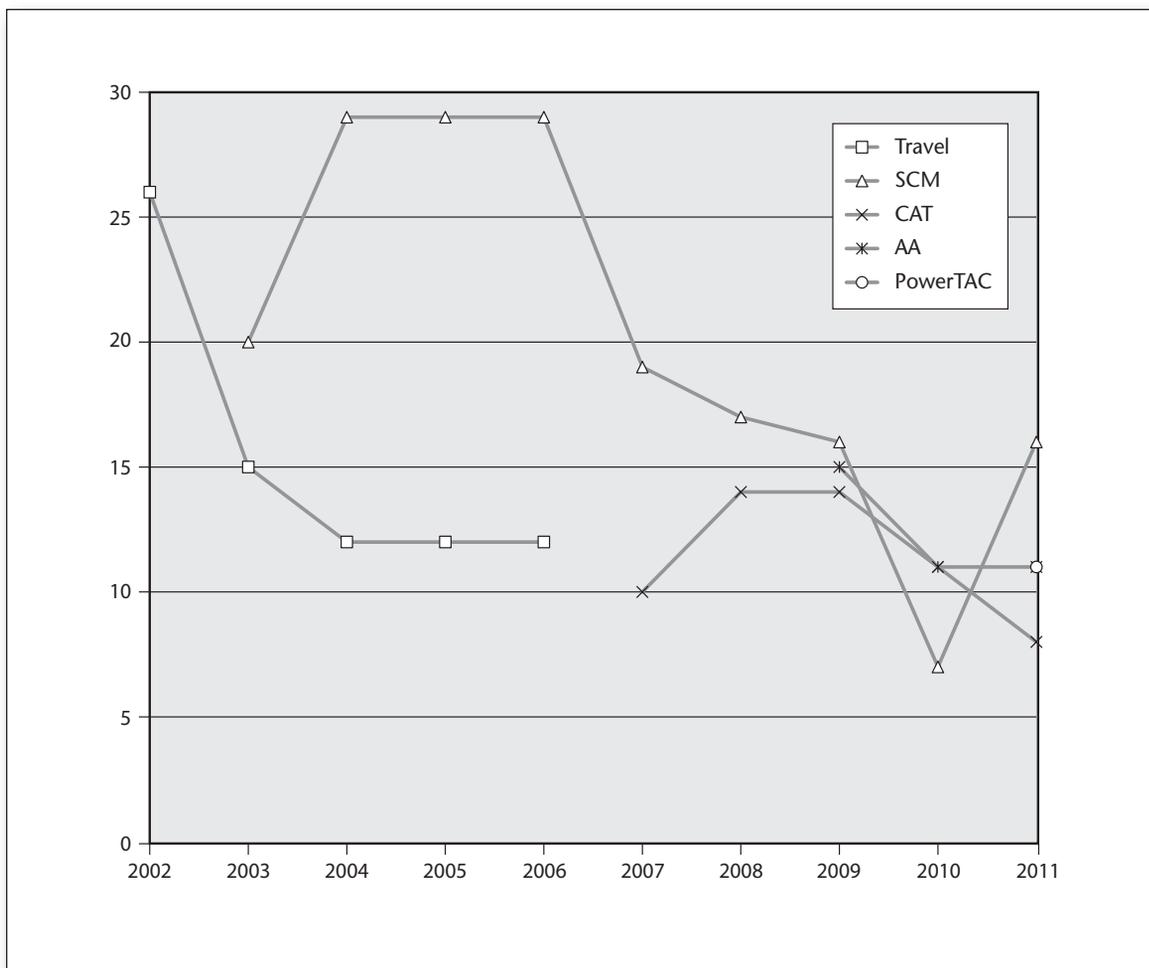


Figure 2. Number of Participating Teams in all TAC Games.

and, if necessary, to tune the market designs and level of realism to enhance the relevance of the shared enterprise for both research value and policy guidance. Figure 2 illustrates the participation in each one of the TAC games from 2002 on.

### The Game and Agent Design Challenges

Designing a competitive simulation that serves a research agenda and provides an interesting and accessible challenge to researchers is a difficult undertaking. And creating a successful agent that competes within that environment is not less challenging. TAC games have been designed in such a way, so that research challenges are introduced progressively, giving time for research groups to adapt and develop the ideas and designs to meet the next research question.

Each of the TAC games is designed to promote research in a specific field of economics and trad-

ing. The challenges met are Travel, SCM, CAT, AA, and Power TAC.

In the Travel challenge, agents participate in continuous one-sided auctions for the procurement of flight tickets, Standard English multiunit auctions for hotel rooms, and continuous double auctions for entertainment activities. They attempt to attract as many clients as possible, through maximizing traveler satisfaction (one party satisfaction) and outperforming competitors.

In the SCM challenge, the game simulates the entire supply chain management process. Agents have to procure computer parts, schedule production and assemble PC units, manage inventory costs, forecast inventory needs, and sell products. Bidding through request for quotes (RFQs) is performed both at the supplier as well as the client side. The development of an efficient solution for the SCM challenge requires multiobjective optimization.

The CAT challenge aspires to solve the inverse SCM problem. Agents have to design an efficient

Game	Years Played	Complexity	Modularity	Adaptability	Extensibility	Domain Applicability
Travel	2002–2006	Low	Low	Low	Low	Basic
SCM	2003 on	High	High	Low	Low	Partial
CAT	2007 on	High	High	Low	Low	Partial
AA	2009 on	Medium	Medium	Medium	High	Partial
Power TAC	2011 on	High	High	High	High	Full

Table 1. TAC Games Features.

stock exchange mechanism for continuous double auctions. To do so, they have to design the market rules for effectively matching buyers and sellers given a dynamic set of traders, compete against other matchmakers by attracting traders to their own market, and maximize their own profit and market share.

In the AA challenge, agents have to develop an efficient strategy for a sponsored ad search mechanism. They must strategically decide which keywords to bid on, and what prices to offer for their ad to be displayed on the search engine page. To decide which keywords to bid on, they have to estimate the game parameters and forecast future game states. Additionally, agents participate in continuous one-sided auctions to define the offer price.

In the Power TAC challenge, agents function as brokers in a power market, trying to attract a profitable set of customers through carefully designed tariffs, but also procuring energy from the wholesale market through day-ahead periodic double auctions (PDA). Power TAC is the most complex TAC problem to date, ensuring versatility in the behavior of all involved entities. Agents have to perform multi-objective optimization in two subdomains and successfully integrate the decisions in an optimal strategy, while ensuring prompt adaptability, due to constant changes in the game environment.

Table 1 summarizes game complexity (in terms of the problem to be solved), modularity (the number of entities involved and subproblems to be solved), adaptability (the ability to modify the behavior of the involved entities), extensibility (the ability to add new entities/concepts to the research agenda of the game), and domain applicability (how easily can research findings be applied to real-life domains).

### Lessons Learned

Combining statistics from participants (omitted due to limited space) with the data from table 1, and feedback provided by the teams yields several interesting insights. TAC counts more than 350 participants in all games from 2002 on. More than 30 research teams have been systematically benchmarking their research strategies under the TAC

framework, indicating their commitment to the trading research effort. What is evident, though, is that up until now only teams illustrating high programming skills could be involved, thus discouraging scientists from other disciplines (such as economics). This is a problem that Power TAC aspires to solve.

With respect to applied strategies, we find two dominant approaches. Teams either perform empirical game evaluations, trying to model the game facets (TacTex) (Pardoe and Stone 2005) or the opponent strategies (Deep Maize) (Jordan, Kiekintveld, and Wellman 2007). Or they attempt to model the bigger picture problem, for example by predicting market dynamics (MinneTAC) (Ketter et al. 2009), or playing a strategy robust to steep market fluctuations (Chatzidimitriou and Symeonidis 2009).<sup>3</sup> The combination of and the competition between these approaches leads to highly interesting game dynamics.

Finally, it is interesting to explore how TAC games have evolved. TAC Travel, the first simulation benchmark designed, dealt with a relatively easy problem. It provided the grounds for establishing a community that could test their own strategies, gain creditability through competitive benchmarking, and investigate competitors' approaches through the reproduction of the research results in offline competitions. After a 5 year course, the travel problem was well-understood, and the community moved on to a more complex problem, CAT. CAT is a much more complex problem, in the sense that end-user preferences play a pivotal role in the way a market is formulated.

TAC SCM was built from the beginning as a high complexity competition, trying to mimic a real supply chain (procurement, inventory, manufacturing, sales, and shipping) as much as possible. And though an information gain study showed that procurement performance dominated in winning solutions (Andrews et al. 2007) and a design flaw on the procurement side was identified (Ketter et al. 2004) (and subsequently solved), the SCM environment still holds numerous interesting research challenges. Even today, the game attracts participants with continuously improving strategies. In 2009, AA was introduced, providing a

much more modular and extensible framework and addressing a much more challenging problem. Many teams with experience in SCM have shifted their interest towards this game, making it very competitive from the beginning.

Lessons learned from the established competitions have led to the design and development of Power TAC, a highly modular and adaptable benchmark that aspires to shed light on four fundamental research tasks related to agent technology and trading environments, namely (1) developing a robust agent architecture for the market trading domain (no silver bullet exists yet), (2) drawing the optimal line between decision support versus decision automation, (3) promoting research for handling steep market fluctuations and, ultimately, (4) increasing trust in agent-based architectures for real-world applications. Power TAC will be officially launched at TAC 2012,<sup>3</sup> where numerous research organizations, as well as business corporations are expected to participate.

### Acknowledgements

We would like to thank all research teams who have participated in TAC over the years; they make the various TAC competitions stimulating and highly research-driven. A big thanks goes to the TAC community as a whole, who provides with its annual Trading Agent Design and Analysis (TADA) workshop a platform for vivid exchange among the various researchers.

### Notes

1. For further information regarding the different competitions and tournament schedules, please visit the Association for Trading Agents Research website ([www.tradingagents.org](http://www.tradingagents.org)).
2. See [issel.ee.auth.gr/doku.php/software/mertacor](http://issel.ee.auth.gr/doku.php/software/mertacor)
3. In 2012, TAC will be held together with the AAMAS and ACM EC conferences in Valencia, Spain, from June 4–8.

### References

- Andrews, J.; Benisch, M.; Sardinha, A.; and Sadeh, N. 2007. What Differentiates a Winning Agent: An Information Gain Based Analysis of TAC-SCM. In *Trading Agent Design and Analysis: Papers from the 2007 AAAI Workshop*. Technical Report WS-07-13. Menlo Park, CA: AAAI Press.
- Bichler, M.; Gupta, A.; and Ketter, W. 2010. Designing Smart Markets. *Information Systems Research* 21(4): 688–699.
- Cai, K.; Gerding, E.; McBurney, P.; Niu, J.; Parsons, S.; and Phelps, S. 2009. Overview of CAT: A Market Design Competition (Version 2.0). Technical Report ULCS-09-005, Computer Science Department, University of Liverpool, Liverpool, UK.
- Chatzidimitriou, K. C. and Symeonidis, A. L. 2009. Agents in Dynamic Supply Chain Management Environments: Data Mining-Driven Design Choices. *Intelligent Systems* 24(3): 54–63.

Collins, J.; Ketter, W.; and Sadeh, N. 2010. Pushing the Limits of Rational Agents: The Trading Agent Competition for Supply Chain Management. *AI Magazine* 31(2): 63–80.

Jordan, P. R.; Cassell, B.; Callender, L. F.; and Wellman, M. P. 2009. The Ad Auctions Game for the 2009 Trading Agent Competition. Technical Report, University of Michigan, Department of Computer Science and Engineering.

Jordan, P. R.; Kiekintveld, C.; and Wellman, M. P. 2007. Empirical Game-Theoretic Analysis of the TAC Supply Chain Game. In *Proceedings of the Seventh International Conference on Autonomous Agents and Multiagent Systems*, 1188–1195. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems ([ssrn.com/abstract-1975237](http://ssrn.com/abstract-1975237)).

Ketter, W.; Collins, J.; Gini, M.; Gupta, A.; and Schrater, P. 2009. Detecting and Forecasting Economic Regimes in Multiagent Automated Exchanges. *Decision Support Systems* 47(4): 307–318.

Ketter, W.; Collins, J.; Reddy, P.; Flath, C.; and Weerd, M. D. 2011. The Power Trading Agent Competition. ERIM Report ERS-2011-027-LIS, RSM Erasmus University, Rotterdam, The Netherlands.

Ketter, W.; Kryzhnyaya, E.; Damer, S.; McMillen, C.; Agovic, A.; Collins, J.; and Gini, M. 2004. Analysis and Design of Supply-Driven Strategies in TAC-SCM. Paper presented at the Trading Agent Design and Analysis Workshop at the Third International Conference on Autonomous Agents and Multiagent Systems, New York, New York, 19–20 July.

Pardoe, D., and Stone, P. 2005. Bidding for Customer Orders in TAC SCM. In *Agent Mediated Electronic Commerce VI: Theories for and Engineering of Distributed Mechanisms and Systems (AMEC 2004)*, ed. P. Faratin and J. A. Rodriguez-Aguilar, volume 3435 of Lecture Notes in Artificial Intelligence, 143–157, Berlin: Springer Verlag.

Wellman, M. P. 2011. *Trading Agents*. Number 12 in Syntheses Lectures on Artificial Intelligence and Machine Learning. San Rafael, CA: Morgan & Claypool.

Wellman, M. P.; Greenwald, A.; and Stone, P. 2007. *Autonomous Bidding Agents*. Cambridge, MA: The MIT Press.

**Wolfgang Ketter** is an associate professor at the Department of Decision and Information Sciences at the Rotterdam School of Management of the Erasmus University and Chair of the Association for Trading Agent Research. He founded and runs the Learning Agents Research Group at Erasmus (LARGE) and the Erasmus Center for Future Energy Business.

**Andreas Symeonidis** is a lecturer with the Department of Electrical and Computer Engineering at the Aristotle University of Thessaloniki and a faculty affiliate with the Center for Research and Technology Hellas. He is TAC 2012 general chair and TADA 2011 workshop chair.