

# MCTS-based Automated Negotiation Agent

Cédric L R Buron  
Thales Research & Technology  
Palaiseau, France  
cedric.buron@thalesgroup.com

Zahia Guessoum  
LIP6, Sorbonne Université  
Paris, France  
zahia.guessoum@lip6.fr

Sylvain Ductor  
Universidade Estadual do Ceará  
Fortaleza, Brazil  
sylvain.ductor@uece.com.br

## ABSTRACT

This paper introduces a new Negotiating Agent for automated negotiation on continuous domains and without considering a specified deadline. The agent bidding strategy relies on Monte Carlo Tree Search, which is a trendy method since it has been used with success on games with high branching factor such as Go. It uses two opponent modeling techniques for its bidding strategy and its utility: Gaussian process regression and Bayesian learning. Evaluation is done by confronting the existing agents that are able to negotiate in such context: Random Walker, Tit-for-tat and Nice Tit-for-Tat. None of those agents succeeds in beating our agent; moreover the modular and adaptive nature of our approach is a huge advantage when it comes to optimize it in specific applicative contexts.

## KEYWORDS

Bargaining and negotiation, Learning agent-to-agent interactions (negotiation, trust, coordination)

### ACM Reference Format:

Cédric L R Buron, Zahia Guessoum, and Sylvain Ductor. 2019. MCTS-based Automated Negotiation Agent. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019*, IFAAMAS, 3 pages.

## 1 INTRODUCTION

Negotiation is a form of interaction in which a group of agents with conflicting interests and a desire to cooperate try to reach a mutually acceptable agreement on an object of negotiation [2]. The agents explore solutions according to a predetermined protocol in order to find an acceptable agreement. Being widely used in economic domains and with the rise of e-commerce applications, the question of automating negotiation has gained a lot of interest in the field of artificial intelligence and multi-agent systems.

Many negotiation frameworks have been proposed [12] and encompass different aspects: the set of participants, the agent preferences and the number of issues. One of the major issues in automated negotiation is considering the time pressure to be well defined over the negotiation, and materialized through a deadline. However, some potential applications of automated negotiation, typically in industrial context, may require an varying time pressure. Factoring is a good example: when a company sells goods or services to another company, it produces an invoice which may be paid after several weeks. This delay of payment may have a negative impact on the company activity, as it may not have sufficient

liquidity to fulfill other contracts. Factoring is an interesting answer to this issue. A funding company – called a *factor* – accepts to fund the invoices of the supplier, by paying them immediately less than their nominal amount and assuming the delay of payment of the principal. From the factor perspective, it can be seen as a short term investment. This setting may include issues of various kinds: continuous (discount rate), numeric (nominal amount of invoices), and categorical (principal). In this application, time pressure is not constant over the negotiation. For the factor, the time pressure depends both on the money it has to invest and on the investment opportunities. For the supplier, the negotiation may be suddenly interrupted at some point if the payment of an invoice makes it useless for the supplier.

In this paper, we introduce a loosely constrained adaptive strategy for automated negotiation that can negotiate with nonlinear preferences, over discrete and continuous issues and without predefined time pressure. We represent bargaining as an extensive game with each proposal considered as a move in the game [17, 19]. We state that our opponent is adaptive, which implies that negotiation history is important. We therefore propose to combine Monte Carlo Tree Search (MCTS) and opponent modeling. MCTS has proved to be a very adaptable game heuristics, in General Game Playing [9]. It has also proved efficient for large branching factor games [5, 20].

## 2 RELATED WORKS

In this section, we introduce the domains related to our agent: automated negotiation and Monte Carlo methods applied to games.

*Automated negotiation.* To review automated negotiation strategies, we rely on the “BOA” (Bidding strategy, Opponent modeling, Acceptance strategy) paradigm [1]:

Bidding strategies may depend on the history, *i.e.* the concessions made by the opponent, a negotiation deadline, the utility function of the agent, and the opponent model. In particular, Tit-for-tat [8] only relies on the opponent proposals. Nice Tit-For-Tat agent [4] uses learning techniques in order to improve it. The other methods rely on time pressure and cannot be applied in our context.

Acceptance strategies can be divided into two main categories [3]. The first category is called “myopic strategies” as they only consider the last bid of the opponent, the agent’s own last proposal or its bidding strategy. The second category consists of “optimal strategies” [3] and rely on the deadline.

Most of the opponent modeling techniques related to automated negotiation have been reviewed by [2]. In our context, they can be used to model: the opponent bidding strategy, its utility and the acceptance strategy. There are two main methods to model adaptive **bidding strategies** which do not rely on the deadline: neural networks and time series-based techniques. Among time series techniques, the Gaussian process regression is a stochastic

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada. © 2019 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

technique which has been used with success by [21]. The opponent **utility** is generally modeled through either frequency based techniques or Bayesian learning. Frequency methods are relevant in the cases where the negotiation domain only consists of discrete issues. Bayesian Learning [14] is well suited for the continuous case and can easily be extended to categorical domains. The opponent **acceptance strategy** can be learned in two ways, either by assuming that the opponent has a myopic strategy or by using neural networks [7]. The latter is computationally expensive though.

*Monte Carlo Tree Search.* Monte Carlo methods are often used as heuristics for games. Kocsis and Szepesvári [15] propose a method to combine the construction of a game tree with Monte Carlo techniques. This method is called Monte Carlo Tree Search (MCTS). It consists of 4 steps: **selection** explores the already built part of the tree, **expansion** generates a new node, **simulation** plays a game until a final state is reached and the utility of the agents is computed and **backpropagation** propagated it over all the selected nodes.

### 3 OUR MCTS-BASED AGENT

Our agent follows the BOA paradigm. It consists of a bidding strategy that implements MCTS and an opponent modeling module divided into two submodules: one for the opponent utility, the other for its bidding strategy. The last module is the acceptance strategy, it accepts the opponent proposal if it is better than the bid generated by the bidding strategy.

*Opponent modeling.* To model the opponent bidding strategy, we use Gaussian Process Regression [18, 21]. One of the capital aspects of this method is the choice of its kernel. In order to choose the best one, we compared them using an automated negotiation setting similar to the one of our experiments. The table 1 shows the results of GPR for each of the most common ones over 50 negotiation sessions. This method can also be used on the categorical issues as explained in chapter 3 of [18].

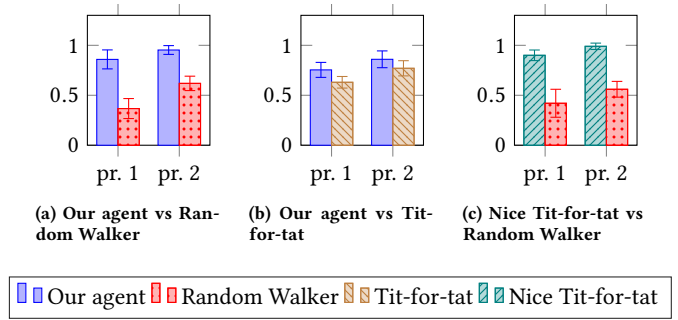
Kernel	RBF	RQF	Matérn	ESS
avg. dist.	43.288	<b>17.766</b>	43.228	22.292

**Table 1: Avg dist. between proposals and GPR predictions**

Bayesian learning [14] considers that an agent makes concessions at roughly constant rate. It relies on triangular functions. It first generates a predetermined number of hypotheses on the utility functions and the estimate of their probability based on received proposals. The estimated utility of the opponent is the sum of these hypotheses weighted by their probability. This method can be naturally extended to the categorical issues by using traditional Bayesian inference.

A simulated agent accepts the proposal from its opponent if its utility is better than the utility generated by its bidding strategy.

*MCTS-based bidding strategy.* As explained in introduction, negotiation can be considered as a 2 players extensive game [17, 19]. However, we must adapt the heuristics traditionally used for games to its peculiarities: **Selection** is based on progressive widening [6]:



**Figure 1: Average utility of negotiating agents**

a new node is expanded if  $n_p^\alpha \geq n_c$ , with  $n_p$  the number of simulations of the parent,  $n_c$  its number of children and  $\alpha$  a parameter of the model. If there is no expansion, the selected node  $i$  maximizes  $W_i = \frac{s_i}{n_i+1} + C \times n^\alpha \sqrt{\frac{\ln(n)}{n_i+1}}$  with  $n$  the number of simulations of the tree,  $s_i$  the score of the node  $i$  and  $C$  a parameter of the model. **Expansion** is chosen randomly among possible bids. **Simulation** is made according to the opponent models. **Backpropagation** is made both on the agent score and the opponent modeled score.

We use the agent knowledge on the game to prune the less promising branches of the tree: *i.e.* all the branches of the tree with lower utility than the best proposal of the opponent (from our agent’s point of view).

### 4 RESULTS

Our agent is evaluated using the GENIUS [16] framework against Tit-for-tat [8], Nice Tit-for-Tat and RandomWalker [4] in the ANAC 2014 setting [10] except that there is no deadline. Figure 1 displays the utility of the agents when negotiating with each other. The results are averaged over 20 negotiation sessions with each profile, with error bars representing the standard deviation from the average.

Our agent is able to beat the Random Walker in every situation and get a significantly better result whatever the profile. Our agent gets a lower utility with Tit-for-Tat but is still able to beat it significantly. The negotiations with Nice Tit-for-Tat never end: the agents keep negotiating forever. We propose instead an indirect evaluation by confronting Nice Tit-for-Tat with Random Walker, in the same setting. The performances of both agents are equal, considering the standard deviation of the series.

### 5 CONCLUSION

In this paper, we presented a negotiating agent able to negotiate in a context where agents do not have predetermined deadline, with both continuous and categorical issues. The experimental results are promising: against all the agents that can negotiate in this domain, our agent outperformed Random Walker and Tit-for-Tat and draws with Nice Tit-for-Tat. Among the perspectives of this work, we propose to adapt it to the multilateral context and try improvements as AMAF and RAVE [11, 13]

## REFERENCES

- [1] Tim Baarslag. *Exploring the Strategy Space of Negotiating Agents: A Framework for Bidding, Learning and Accepting in Automated Negotiation*. PhD thesis, Delft University of Technology, 2016.
- [2] Tim Baarslag, Mark J C Hendriks, Koen V Hindriks, and Catholijn M Jonker. Learning about the opponent in automated bilateral negotiation: a comprehensive survey of opponent modeling techniques. *Autonomous Agents and Multi-Agent Systems*, 20(1):1–50, 2015.
- [3] Tim Baarslag and Koen V. Hindriks. Accepting optimally in automated negotiation with incomplete information. In *AAMAS '13*, pages 715–722, Richland, SC, 2013. International Foundation for Autonomous Agents and Multiagent Systems.
- [4] Tim Baarslag, Koen V Hindriks, and Catholijn Jonker. A tit for tat negotiation strategy for real-time bilateral negotiation. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, volume 435, pages 229–233. Springer Berlin Heidelberg, 2013.
- [5] Cameron C Browne, Edward Powley, Daniel Whitehouse, Simon M. Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Taverner, Diego Perez, Spyridon Samothrakis, and Simon Colton. A survey of Monte Carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games*, 4(1):1–43, 2012.
- [6] Adrien Couëtoux. *Monte Carlo Tree Search for Continuous and Stochastic Sequential Decision Making Problems*. PhD thesis, Université Paris XI, 2013.
- [7] Fang Fang, Ye Xin, Yun Xia, and Xu Haitao. An opponent's negotiation behavior model to facilitate buyer-seller negotiations in supply chain management. In *2008 International Symposium on Electronic Commerce and Security*, 2008.
- [8] Peyman Faratin, Nicholas R Jennings, and Carles Sierra. Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems*, 24(3-4):159–182, 1998.
- [9] Hilmar Finnsson. Generalized Monte Carlo tree search extensions for general game playing. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, AAAI 12, pages 1550–1556. AAAI Press, 2012.
- [10] Naoki Fukuta, Takayuki Ito, Minjie Zhang, Katsuhide Fujita, and Valentin Robu, editors. *Recent Advances in Agent-based Complex Automated Negotiation*, volume 638 of *Studies in Computational Intelligence*. Springer International Publishing, 2016.
- [11] Sylvain Gelly and David Silver. Monte-carlo tree search and rapid action value estimation in computer go. *Artificial Intelligence*, 175(11):1856–1875, 2011.
- [12] Robert H Guttman, Alexandros G Moukas, and Pattie Maes. Agent-mediated electronic commerce: a survey. *The Knowledge Engineering Review*, 13(02):147–159, Jul 1998.
- [13] David P Helmbold and Aleatha Parker-Wood. All-moves-as-first heuristics in monte-carlo go. In *IC-AI*, pages 605–610, 2009.
- [14] Koen Hindriks and Dmytro Tykhonov. Opponent modelling in automated multi-issue negotiation using bayesian learning. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems*, volume 1, pages 331–338, 2008.
- [15] Levente Kocsis and Csaba Szepesvári. Bandit based Monte Carlo planning. In Johannes Fürnkranz, Tobias Scheffer, and Myra Spiliopoulou, editors, *Machine Learning: ECML 2006: 17th European Conference on Machine Learning Berlin, Germany, September 18-22, 2006 Proceedings*, pages 282–293. Springer Berlin Heidelberg, 2006.
- [16] Raz Lin, Sarit Kraus, Tim Baarslag, Dmytro Tykhonov, Koen Hindriks, and Catholijn M Jonker. Genius: an integrated environment for supporting the design of generic automated negotiators. *Computational Intelligence*, 30(1):48–70, 2014.
- [17] John F Nash Jr. The bargaining problem. *Econometrica: Journal of the Econometric Society*, pages 155–162, 1950.
- [18] Carl E Rasmussen and Christopher K I Williams. *Gaussian processes for machine learning*. MIT Press, 2006.
- [19] Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, pages 97–109, 1982.
- [20] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, and al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.
- [21] Colin R Williams, Valentin Robu, Enrico H Gerding, and Nicholas R Jennings. Using gaussian processes to optimise concession in complex negotiations against unknown opponents. In *IJCAI'11*, pages 432–438, 2011.