



# MCTS-based Automated Negotiation Agent

Cédric Buron, Zahia Guessoum, Sylvain Ductor

## ► To cite this version:

Cédric Buron, Zahia Guessoum, Sylvain Ductor. MCTS-based Automated Negotiation Agent. International Conference on Principles and Practice of Multi-Agent Systems (PRIMA), Oct 2019, Torino, Italy. pp.186-201, 10.1007/978-3-030-33792-6\_12 . hal-02282222

**HAL Id: hal-02282222**

**<https://hal.science/hal-02282222>**

Submitted on 12 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MCTS-based Automated Negotiation Agent

Cédric L.R. Buron\*, Zahia Guessoum†, Sylvain Ductor§

September 12, 2019

## Abstract

This paper introduces a new negotiating agent model for automated negotiation. We focus on applications without time pressure with multidimensional negotiation on both continuous and discrete domains. 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 also exploits opponent modeling techniques thanks to 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. Also, the modular and adaptive nature of our approach is a huge advantage when it comes to optimize it in specific applicative contexts.

**keywords:** Automated negotiation, MCTS, Supply Chain

## 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 [13]. They may be characterized along different aspects, whether concerning the set of participants (e.g. bilateral or multilateral), agent preferences (e.g. linear or not), issues of negotiated objects (e.g. discrete or continuous), or even the characteristics of the interaction protocol (e.g. globally bounded in time or number of rounds). They run negotiating agents that use strategies to evaluate the received information and make proposals. Several strategies have been proposed. Either fixed or adaptive, most of them rely on a known deadline (either in time or in rounds). However in several applications, the deadline of an agent may change over the negotiation. The negotiation horizon may vary depending on external elements such as other opportunities. To the best of our knowledge, these elements have not been taken into account so far.

In this paper, we propose to handle this issue by designing a loosely constrained adaptive strategy for automated negotiation. This strategy considers that: 1) the agent preferences are nonlinear, 2) the issues of negotiated objects can be both continuous and discrete and 3) the time pressure is undefined, and therefore the deadline of the negotiation. To cope with this objective, our agent is based on General Game Playing [11] and Machine Learning [2]. Its strategy relies on both Monte Carlo Tree Search (MCTS), a heuristic technique that has been used with success for many kinds of games (see for instance [5, 23]), and opponent modeling techniques in order to be more efficient.

The paper is organized as follows. Section 2 describes our targeted industrial application. Section 3 provides some background on automated negotiation and AI strategies for games. Section 4 gives the theoretical and formal setting for bar-

---

\*Thales Research & Technology, Palaiseau France  
cedric.buron@thalesgroup.com

†LIP6, Sorbonne Université, Paris, France

‡CReSTIC, Université de Reims Champagne Ardennes, France

§Universidade Estadual do Ceará

Cite as Cédric L. R. Buron, Zahia Guessoum, and Sylvain Ductor. “MCTS-based Automated Negotiation Agent”. In *The 22nd International Conference on Principles and Practice of Multi-Agent Systems (PRIMA2019)*, Turin, 2019

gaining in order to motivate the use of AI for games. Section 5 introduces our strategy. Section 6 gives some details on the agent implementation and shows its performance against a Random Walker agent, Tit-for-Tat agent and Nice Tit-for-Tat agent. The last section gives concluding remarks and perspectives.

## 2 Target application

This work is part of an industrial project that addresses an economic application, the *factoring*. This application requires a solution that complies with the specific scope we consider here and which is neglected by the literature.

When a company sells goods or services to another company, it produces an invoice. The selling company is called *supplier* and the customer is called *debtor*. Each country may define a legal payment term of generally several weeks. Moreover, the principal may not pay within this payment term. In the supply chain, the debtor is often much larger than the supplier. It can therefore impose its own conditions at the expense of the supplier. The consequences are quite harmful for the latter: during those payment delays, its *working capital* is reduced and hence its capacity to produce, fulfill future orders or pay its own suppliers.

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 – generally a bank or an investment fund – the principal can be seen as a short-term investment, where the risk of defaulting on payment depends on the reputation of the principal. Since we consider the case where the principal is much larger than the supplier, this risk is lower and the rate is more affordable for the supplier.

As this kind of funding may be recurrent, there is a strong interest in automating the negotiation between the supplier and the factor. Moreover, some recent works have shown an increase in the number of factoring marketplaces all around the world [8]. However, there are several specificities to this setting. The first specificity is the negotiation domain. Several elements are negotiated at the same time: the nominal amount to be funded and the discount rate are the primary elements

of the negotiation. Also, for identical nominal amounts, a factor may ask the supplier to sell invoices of certain principals it trusts. Finally, when several invoices are available, the expected number of financing days may also be negotiated. Therefore, we consider complex issues that combine at the same time elements of various kinds: continuous (the discount rate), numeric (the numerical amount and the financing days), and categorical (the principal).

The second specificity is related to uncertainty and resource availability. Automated negotiation often considers a deadline, which defines the time allocated to the negotiation. Most of the negotiation strategies rely on this time pressure to compute a concession rate. In our 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. If the factor has a lot of money and few opportunities, the time pressure increases: the factor sees not invested money as a loss. On the contrary, when resources are limited and opportunities are common, the factor tries to get a better discount rate and the time pressure decreases. For the supplier, the situation is even more unpredictable. Time pressure depends on its opportunities to get new credit lines (including bank loans) and even on the time the principal takes to pay it: the negotiation may be brutally interrupted at some point if the payment of an invoice makes it useless for the supplier.

Automated negotiation components relying either on a deadline commonly known by the agents, or on a deadline private to each of them is therefore not applicable to our target applications. In our target application the negotiation domain consists of numerical, continuous and categorical issues, the agent preferences are nonlinear, and the time pressure is dynamic.

## 3 Related work

Our agent is at the meeting point of automated negotiation and Monte Carlo methods applied to games. In this section, we introduce both domains.

### 3.1 Automated Negotiation

Various authors have explored the negotiation strategies with different perspectives. Most of those works have identified three components that make up the “BOA” architecture [1]: a **Bidding strategy** defines the offers the agent sends to its opponent, an **Acceptance strategy** defines whether the agent accepts the offer it just received or if it makes a counterproposal, and an **Opponent modeling** models some features of the opponents, as its bidding strategy, its preference domain and its acceptance strategy. The latter aims to improve the efficiency of the bidding strategy and/or the acceptance strategy of the agent. The following subsections present the related work for each component.

#### 3.1.1 Bidding Strategy

Bidding strategies may depend on several elements: the history, including the concessions made by the opponent and/or a negotiation deadline, the utility function of the agent, and the opponent model. Faratin *et al.* proposed the so-called tactics that mainly rely on the criticality of the resources, the remaining time before the deadline is reached or the concessions made by the opponent. All of them except the last rely on a known deadline. The latter has been extended to create more complex strategies, as the Nice tit-For-Tat agent [4], which uses learning techniques in order to improve it. Genetic algorithms [15] use generated proposals as individuals. They rely on the time pressure for the variation of their proposals and make some that are acceptable for their opponent.

This paper advances the state of the art in the bidding strategies by introducing and evaluating a strategy that considers the negotiation as a game and uses the very efficient Monte Carlo Methods.

#### 3.1.2 Acceptance Strategy

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. An agent may accept an offer when (1) it is better than the new one produced by its own bidding strategy, (2) it is better than the last one made by the agent, (3) it is above a predetermined threshold, or (4) it embodies any

combination of the previous ones. The second category consists of “optimal strategies” [3] that rely on an opponent bidding strategy in order to optimize the expected utility. They are based on the concessions made by the opponent and a prediction that the expected utility of the agent should increase while the deadline is getting closer. So, the first category is not suitable for our context. We therefore propose to use the second category in our agent model.

#### 3.1.3 Opponent modeling

[2] presents an exhaustive review of the opponent modeling techniques related to automated negotiation. They are generally used to model (1) opponent bidding strategy, (2) its utility and (3) its acceptance strategy as well as private deadline and a reserve price, depending on which of these elements are relevant in each context.

There are two main methods to model adaptive **bidding strategy** which does not rely on the deadline: neural networks and time series-based techniques. Neural networks use a fixed number of previous offers as input, and the expected value for the next proposal as output. Time series methods can be generalized very easily. Among them, the Gaussian process regression is a stochastic technique which has been used with success by [24]. Due to its nature, it can generate various proposals at each negotiation turn. Those proposals are proportional to a likelihood provided by the regression. We select this technique to model the opponent bidding strategy since it is particularly adapted to the Monte Carlo Tree Search.

The opponent **utility** is generally considered as the weighted average of partial utility functions for each issue. Two families of methods are used to model them. The first one is based on the frequency of each value among the previous opponent bids. The methods of this family make the hypothesis that the most frequent values are the ones the opponent prefers, and that the most stable issues are the most important for it. They are relevant in the cases where the negotiation domain only consists of discrete issues; their extension to the continuous case is not suitable to complex domains, as it requires the definition of a distance function which depends on the negotiation domain. The second family of methods is based on Bayesian Learning [14]. It is well suited for the continuous case and can be easily extended to categorical

domains. We use it as is for the numerical issues and make an extension for the categorical ones. The latter is presented in section 5.1.2.

There are two ways to learn the opponent **acceptance strategy**: either by assuming that the opponent has a myopic strategy or by using neural networks [9]. The latter is quite expensive in terms of computation time. The weights of the network must be updated each time the opponent makes a new proposal.

### 3.2 Monte Carlo Tree Search

Monte Carlo methods are regularly used as heuristics for games. Rémi Coulom [6] proposes a method to combine the construction of a game tree – a traditional method for games that has proved to be very effective – with Monte Carlo techniques. This method is called Monte Carlo Tree Search (MCTS) and it has been improved using various extensions [5], including pruning the less promising branches of the tree. It has met great success in games, particularly games with high branching factor [23].

MCTS consists of 4 steps. A **selection** is dedicated to the exploration of the already built part of the tree, based on a predefined strategy. While exploring a node, the algorithm chooses whether to explore a lower-level branch or expand a new branch. In the latter case, there is an **expansion** of a new node; it is created just below the last explored one. Once a new node has been expanded, a **simulation** of the game is performed until a final state is reached. Finally, outcomes of the final state are computed and a **backpropagation** of them is made over all the nodes that have been explored.

[16] presents a recent attempt to exploit MCTS for General Game Playing, with Automated Negotiation as a potential application. The negotiation domain considered is limited to a single-issue, discrete domain with complete information (each agent knows what is the optimal deal for its opponent). Our work is specifically made for recent evolution of Automated Negotiation, focusing thus on multi-issues, combining continuous, numerical and categorical domains with incomplete information (the agent has no information on the opponent utility profile). These differences have consequences on the technical aspects of these works. The referenced works use Upper Confidence Trees, inapplicable in our case since it imposes the num-

ber of possible moves at each step to be finite. Also, [16] does not require opponent modeling, since the opponent utility profile is known.

## 4 Negotiation as a game

Monte Carlo Tree Search has been applied with success to extensive games. In this section, we show how it is possible to represent negotiation using this model. We first associate each aspect of negotiation to a game element. We then describe specificities of negotiation that prevent us from using most common MCTS selection, expansion and simulation strategies.

An extensive game [20] consists of a set of players, the set of all possible game histories, a function mapping each non-terminal history to the player who must play then and a preference profile. By using this definition, it is possible to define a bargaining  $\mathcal{B}$  as an extensive game:

**Definition 1** (Bargaining). A bargaining can be represented as triple  $\mathcal{B} = (H, A, (u_i)_{i \in \llbracket 1, 2 \rrbracket})$  where:

1.  $A$  is the set of two **players**: the *buyer* (player 1) and the *seller* (player 2),
2.  $H$  is the set of **possible histories** of the negotiation. Each history consists of the sequence of messages the agents sent to each other: *proposals*, *acceptance* and *rejection* messages. Terminal histories are the histories ending by acceptance or rejection, and infinite histories. Each message is a pair  $(\alpha, c)$  where  $\alpha$  is the speech act (performative) of the message and  $c$  is the content of the message, *i.e.* a list of couples  $(k, v)$  where  $k$  is the key of an issue of the negotiation domain and  $v$  is the corresponding value. The history can be divided in two parts, each part corresponding to the messages sent by one of the agents:  $h_i = (\alpha, c)_i$ ,
3. the **player** function is based on the parity of the size of the history (we suppose that the buyer (player 1) always plays first). Therefore  $\forall h \in H, \text{player}(h) = 2 - (|h| \bmod 2)$  where  $|h|$  is the size of  $h$ ,
4. the **preference profile**,  $u_i, i \in \{1, 2\}$ , is an evaluation of terminal histories with regards to each player. If the history ends with an

acceptance of the agent,  $u_i$  returns the utility associated by the agent, if not it returns a specific value that may depend on the engaged resources.

The representation of bargaining as a game has already been investigated in [19, 22] for single issue bargaining. More complex domains has been initially dealt by making the assumption that the agents proposals are independent of the opponent's ones [10]. However in recent advances in automated negotiation, the agent bidding strategies are generally adaptive *i.e.* the proposals made by an agent depend on its opponent's ones.

However, negotiation is not a classical combinatorial game as Chess or Go. Its resolution is a challenge to the MCTS approach for three reasons. First, it is a non-zero sum game: agents try to find a mutually beneficial agreement which is often much better for both agents than their reserve utility *i.e.* the situation where the agents do not find an agreement. Second, it is an incomplete information game: the agent preference profiles are unknown to their opponents and generally modeled by them. These two specificities make it impossible to use the most common implementation of MCTS, the Upper Confidence Tree [17]. Last, we consider a large and complex domain that encompasses numeric, continuous and categorical issues, with nonlinear utility functions and possibly infinite game trees. This has several consequences on the way the tree is explored, in particular on the criterion followed to expand a new node.

## 5 MCTS-based agent

As we explained in the previous section, negotiation is a particular game. It is therefore required to adapt the heuristics traditionally used for games to its specificities. In this section, we present our automated negotiation agent relying on MCTS. The agent architecture is composed of three modules presented in Figure 1. The bidding strategy module implements MCTS and uses the opponent modeling module. The latter consists of two submodules: one models the opponent utility, the other models its bidding strategy. The last module is the acceptance strategy, which makes a comparison between the last proposal from the opponent and the bid generated by the bidding

strategy. Each of the agent submodules and their interactions are described in this section.

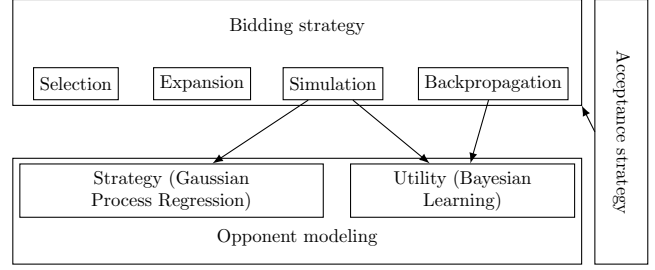


Figure 1: Interaction between the modules of our agent

### 5.1 Opponent modeling

In order to improve the efficiency of MCTS, we model both the bidding strategy and the utility of the opponent.

#### 5.1.1 Bidding Strategy Modeling

The goal of this model is to predict what proposal the opponent will make at turn  $x_*$ . To do so, we use Gaussian Process Regression [21]. This method produces a Gaussian with the mean corresponding to the value predicted by the algorithm and a standard deviation corresponding to the uncertainty induced by the model.

The first step is to compute the covariance matrix  $K$  which represents the proximity between the turns  $(x_i)_{i \in \llbracket 1, n \rrbracket}$  of the sequence, based on a covariance function, also called kernel  $k$ . Let

$$K = \begin{pmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) \\ \vdots & & \vdots \\ k(x_n, x_1) & \dots & k(x_n, x_n) \end{pmatrix} \quad (1)$$

we then compute the distance between the turn of the predicted proposal  $x_*$  and the previous turns in the vector  $K_*$ :

$$K_* = (k(x_*, x_1), \dots, k(x_*, x_n)) \quad (2)$$

The Gaussian process regression relies on the supposition that all these values are the dimensions of a multivariate Gaussian. Using results on multivariate Gaussian, we can compute

$$\bar{y}_* = K_* K^{-1} \mathbf{y} \quad (3)$$

$$\sigma_*^2 = \text{Var}(y_*) = K_{**} - K_* K^{-1} K_*^\top \quad (4)$$

where  $K_{**} = k(x_*, x_*)$ . The result corresponds to a Gaussian random variable with mean  $\bar{y}_*$  and standard deviation  $\sigma_*$ .

One of the capital aspects of Gaussian process regression is the choice of the kernel. The most common ones are radial basis functions (RBF), rational quadratic functions (RQF), Matérn kernel and exponential sine squared (ESS). These kernels are used to define distance between the turns of the bargaining. We tested these four kernels on various negotiations among finalists of ANAC 2014 (bilateral general-purpose negotiation with nonlinear utilities). The table 1 shows the results of GPR for each of the aforementioned kernels. We generated randomly 25 negotiation sessions and modeled two agents using each kernel. We got a total of 50 models by kernel. Each bid of each sequence is predicted using previous proposals and is used to predict following ones. The table shows average Euclidean distance between the actual proposals and the predicted sequences. The lower the value, the closer is the prediction to the actual proposal. The kernel that got the best result is the Rational Quadratic Function. Our agent therefore uses this one.

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

Table 1: Average Euclidean distance between actual proposals of a bargaining and values predicted by GPR, depending on the used kernel.

This method also allows to make predictions on the categorical issues. The method presented in chapter 3 of [21] is also possible and relies on Monte Carlo method for some integration estimation.

### 5.1.2 Preference Profile Modeling

Bayesian learning described in [14] makes the only supposition that an agent makes concessions at roughly constant rate. Though this constraint may seem tough, it is relatively low in comparison with other methods.

The opponent utility is approximated by a weighted sum of triangular functions. A function  $t$  of  $[a, b] \subset \mathbb{R}$  in  $[0, 1]$  is called triangular if and only if:

- $t$  is linear and either  $t(a) = 0$  and  $t(b) = 1$  or  $t(a) = 1$  et  $t(b) = 0$ , or:

- there is some  $c$  in  $[a, b]$  such that  $t$  is linear on  $[a, c]$  and  $[c, b]$ ,  $t(a) = 0$ ,  $t(b) = 0$  and  $t(c) = 1$ .

The method can be divided into two steps. First, the agent generates a predetermined number of hypotheses on the utility function. These hypotheses are composed of weighted sums of triangular functions (one per issue). Each issue is therefore associated with a weight and a triangular function.

The estimated utility of the opponent is the weighted sum of these hypotheses where each weight is the probability computed using Bayesian learning. This method does not make any supposition on the opponent strategy, which makes it more general than the frequency based techniques.

This method can be naturally extended to the categorical issues. Given a categorical issue  $C = \{C_1, \dots, C_n\}$ , the partial utility function is chosen among the set of functions of  $[0, 1]^C$ .

### 5.1.3 Acceptance Strategy Model

Acceptance strategy is modeled in a very simple way: a simulated agent accepts the proposal from its opponent if and only if its utility is better than the utility generated by the bidding strategy model. This method presented by Baarslag et al. in [25] is not computationally expensive.

## 5.2 MCTS-Based Bidding Strategy

Monte Carlo methods are very adaptive, and achieve promising results in various games, including games with high branching factor. In this section, we describe the way we have adapted these methods to the negotiation context.

### 5.2.1 Raw MCTS

As we explained before, MCTS is a general algorithm and relies on several strategies. Each time an agent needs to take a decision, it generates a new tree and explores it using MCTS. The most common implementation of MCTS is Upper Confidence Tree (UCT). This method has proved to be very efficient, particularly for Go. Nevertheless, it expands a new node whenever it explores a node whose children have not all been explored, which is not applicable when issues are continuous. Beyond that, in the case where the branching factor is not small in comparison with the number of simulation, UCT keeps expanding a new node without exploring deeper nodes, which loses the

interest of MCTS compared to flat Monte Carlo. In our context, it is therefore necessary to define a different implementation of MCTS:

**Selection** For this step, we use progressive widening, as described by [7]. The expansion criterion of the progressive widening states that a new node is expanded if and only if:

$$n_p^\alpha \geq n_c \quad (5)$$

where  $n_p$  is the number of times the parent has been simulated,  $n_c$  is its number of children and  $\alpha$  is a parameter of the model. If the result is that a new node is not expanded, the selected node is the node  $i$  maximizing:

$$W_i = \frac{s_i}{n_i + 1} + C \times n^\alpha \sqrt{\frac{\ln(n)}{n_i + 1}} \quad (6)$$

where  $n$  is the total number of simulations of the tree,  $s_i$  is the score of the node  $i$  and  $C$  is also a parameter of the model.

**Expansion** The content of the expanded node is chosen randomly among all the possible bids of the domain with an even distribution

**Simulation** During the simulation, the model of the opponent bidding strategy is used in order to make the simulation more representative. The model of the opponent utility and its acceptance strategy are used to decide when it accepts a proposal.

**Backpropagation** The backpropagation also uses the opponent utility model. The utility of both agents is computed and the scores of both agents is updated for each visited node. The utility of the agent itself is computed using its real preference profile.


### 5.2.2 Pruning

In order to explore only the interesting nodes for the agent, it is possible to use the knowledge of the agent on the game to prune the less promising branches of the tree. Though we do not have much information in our context, we decide that the opponent should find any offer it made acceptable. We therefore decide to prune all the branches of the tree where our agent makes a proposal less interesting than the best proposal received from the opponent: the goal is to use the best proposal of the opponent as a lower bound and to try to improve on this basic value.

## 6 Experiments

In order to evaluate our agent, we use the GENIUS [18] framework. We confront our agent with the only agents that do not require a deadline for their strategy to the best of our knowledge: two variants of the Tit-for-Tat agent and a RandomWalker agent. In this section, we describe the implementations of our agent and those three. We then describe in detail the experimental protocol and finally we present the achieved results.

### 6.1 Implementation

Our agent is developed in Java and consists of a set of independent modules connected to GENIUS platform thanks to interfaces. Figure 2 presents this architecture (the  represents the messages sent and received by the agent). It consists of the components of Section 5: a module for the MCTS-based bidding strategy (with classes for the Monte Carlo Tree, its nodes, etc.), a module for opponent utility modeling and another one for the opponent bidding strategy modeling. We use the same acceptance strategy for both our agent and the model it uses for its opponent. As they are very simple, there is no module dedicated to them.

As MCTS are computationally expensive and take quite a long time, we parallelized them. The stochastic aspect of the opponent bidding strategy model ensures that two simulations going from the same branch of the tree are not similar, while still prioritizing the most probable values. The Gaussian process regression, which relies on matrix computation, has been developed using the Jama library<sup>1</sup>. The parameters of the kernel are optimized using the Apache Commons Math library utilities<sup>2</sup>.

The time taken by our agent for each round has been set empirically. When our agent makes a choice among 200 bids, there is generally one that results into a high utility both for it and its opponent. In order to explore the tree in depth, we let it enough time to generate about 50'000 simulations. Then, to make a proposal, our agent takes about 3 minutes. This duration meets the expectation of the real-world applications of our industrial partners. From these values, we get  $\alpha = 0.489$ .

<sup>1</sup><http://math.nist.gov/javanumerics/>

<sup>2</sup><http://commons.apache.org/proper/commons-math/>



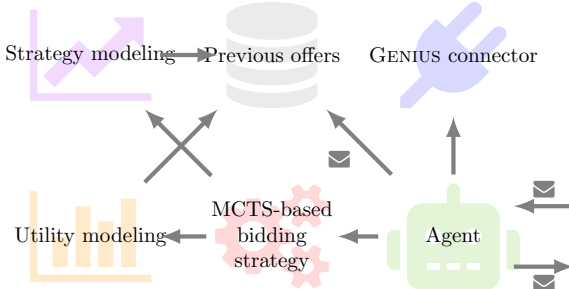


Figure 2: Software architecture of the modules of our agent

## 6.2 Opponent’s Description

In this section, we provide a description of the confronted agents. It is impossible to compare our agents with state-of-the-art ones, as they all rely on the supposition that there is a publicly known deadline. The only three agents able to negotiate without a given deadline are the RandomWalker, the Tit-for-Tat agent and its more evolved variant, the Nice Tit-for-Tat.

### 6.2.1 RandomWalker

The RandomWalker is described by Baarslag in [3] and makes random proposals.

### 6.2.2 Tit-for-Tat

Tit-for-tat agent was first described by Faratin *et al.* in [10]. This agent makes a concession whenever its opponent makes concessions itself. Several possible implementations are given. Here, when it has received fewer than 2 proposals the agent makes the most interesting proposal from its own perspective (it generates 10’000 random proposals and chooses the best one for its utility function). For the other proposals, the agent looks at the last two proposals of its opponent and computes the made concession. This concession may be positive or negative. It adds this concession to the utility of its own last proposal. It then searches for a proposal with the closest utility to this target value. In our implementation, 10’000 proposals are generated to this end.

### 6.2.3 Nice Tit-for-Tat

Nice Tit-for-Tat is somehow, a more evolved version of the Tit-for-Tat. It has been described by

Baarslag *et al.* in [4]. The goal of this work is to comply with domains where mutual agreement is possible. The agent then uses the same opponent utility modeling as our agent and uses it to estimate the Nash point of the setting, i.e. the agreement maximizing the product of the utilities of the agents. The concession rate is computed between the first and the last bids of the opponent as the percentile of the distance between its first bid and the Nash point. The corresponding concession is made from the agent point of view. Among the equivalent offers, the utility model is also used in order to choose the best bid for the opponent among equivalent ones for the agent. The only difference with the version used in our experiments and the agent proposed in [4] is the acceptance strategy. Indeed, the strategy proposed in the original version which is presented in [3] depends on the deadline of the negotiation in order to take the time pressure into account. Here, we provide the Nice Tit-for-Tat with a simplified version of its acceptance strategy, which corresponds to the same acceptance strategy as our agent.

## 6.3 Experimental Protocol

GENIUS makes it possible to negotiate on numerical or categorical issues, but not yet on continuous ones. In order to evaluate our agent, we want to target a negotiation domain that is at the same time neutral enough to show the generic aspects of our work, but complex enough to motivate its use. ANAC is an international competition used to determine the effective negotiation strategies. The negotiation domain used in ANAC 2014 [12] fits well to this objective. Subsequent competitions focused on multilateral negotiation and specific application domains. In ANAC 2014 domain, issues are numerical, varying from 1 to 10. Several domains have been proposed, varying from 10 to 50 issues. In order to reduce computational complexity which is not the concern of this work, we use the 10-issue version. The utility functions are non-linear, and the reserve utility is set to 0, which is the minimal outcome value for the agents. As the time pressure is supposed to vary over time, we do not use a discount rate. While it does not exactly correspond to the targeted application, its complexity ( $10^{10}$  possible proposals) makes a suitable test bench for our negotiation strategy. In order to simulate the fact that the time pressure is unknown to the agent, we put a very large dead-

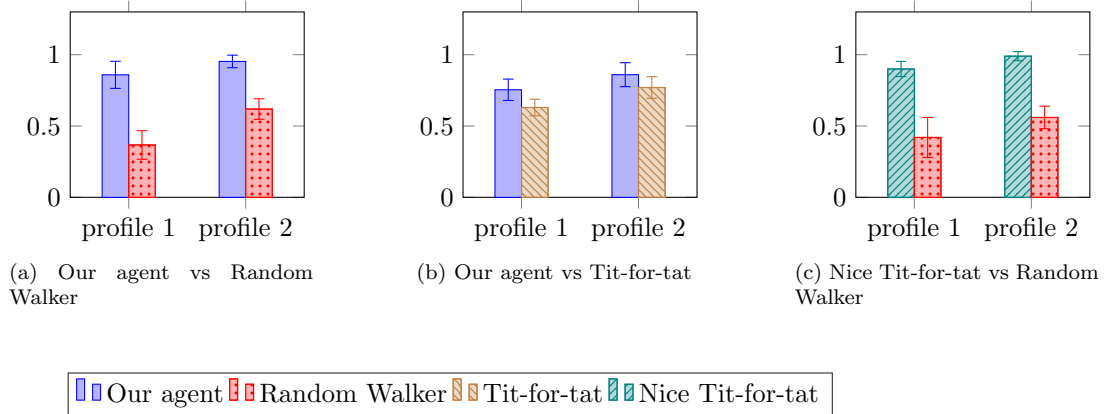


Figure 3: Average utility of negotiating agents

line, so that it is never reached.

## 6.4 Results

Figure 3 displays the utility of the agents when negotiating with each other using a histogram. The results are averaged over 20 negotiation sessions with each profile, with error bars representing the standard deviation from the average.

Note that the two preference profiles are very different from each other, and not symmetrical at all. This specificity explains the fact that for all the agents in all configurations, utility is always higher with Profile 2 than it is with Profile 1.

As shown on Figure 3a, our agent is able to beat the Random Walker in every situation, even when its preference profile is Profile 1 and Random Walker’s is Profile 2. It is interesting to note that the negotiations with Random Walker are very short with only 3.1 proposals in average: 2.5 when it gets Profile 1 and 3.7 proposals when our agent gets Profile 2. This difference can be explained by the fact that it is easy to find agreements with very high results for Profile 1 (0.9 or more) and high results for Profile 2 (0.6). In most of the negotiations, the first proposal of our agent is of this kind. In that case, Random Walker is more likely to generate a proposal with utility lower than the one proposed by our agent and accepts it, generating a utility of 0.6 for Random Walker and a negotiation session consisting of a single proposal.

Negotiating with Tit-for-Tat is harder, as we can see on Figure 3b. Our agent gets a lower utility than the Random Walker but is able to beat Tit-

for-Tat. The expectation level of Tit-for-Tat also generates much longer negotiations: 34.2 proposals on average with 31.55 proposals when it gets Profile 1 and 36.85 proposals when our agent gets Profile 2. This result can be explained the same way as the results of the negotiation with Random Walker.

The negotiations with Nice Tit-for-Tat never ends: the agents keep negotiating forever. Our MCTS-based method refuses to make a concession significant enough to have a chance to be accepted by Nice Tit-for-Tat, considering the high expectation it has by using the Nash point. Reciprocally, Nice Tit-for-Tat, without time pressure, and dynamic adaptation of its acceptance strategy, does not accept the proposals of our agent. By looking at the internal state of the Nice Tit-for-Tat, we also see that its estimation of the Nash point is incorrect: it expects a utility above the real one.

We propose instead an indirect evaluation by confronting Nice Tit-for-Tat with Random Walker, in the same setting. The results are represented on Figure 3c. The performances of both agents are comparable, considering the standard deviation of the series.

## 7 Conclusion

In this paper, we presented an automated negotiation agent able to negotiate in a context where agents do not have predetermined deadline, neither in time nor in rounds, and where the negotiation domain can be composed of numerical, continuous and categorical issues. We described this

setting as an extensive game and described a negotiation strategy based on a specific implementation of MCTS relying on two opponent models and a pruning strategy. One of them is the Gaussian process regression, which relies on a covariance function. We tested several covariance functions and chose the one that provides better results in context similar to ours.

Experiments were run in the context of a large negotiation domain, with nonlinear utility functions using different preference profiles. The experimental results are promising: against all the agents that can negotiate in its negotiation domain, our agent outperformed Random Walker and Tit-for-Tat and draws with Nice Tit-for-Tat. This work therefore indicates that techniques from games such as MCTS can be used with success in automated negotiation. However, the modularity of the architecture and the variety of strategies proposed on General Game Playing and Machine Learning areas are a huge advantage when it comes to optimizing the agent for a specific application domain.

Among the perspectives of this work, we would like to create a customized version of our agent and adapt it to the context where the deadline is known, in order to make it available for these applications. Our agent can already be used in this context, but the fact that it does not exploit this information may make it less efficient than its opponents. Another possible direction would be to adapt it to the multilateral context. In fact, there would be little modification to make it available for a context of stacked alternate protocol or a many-to-many bargaining scenario, since MCTS has already been used in n-player games. The use of our agent in its industrial context, in particular with corresponding negotiation domains would yield very interesting results. Last, we would like to improve our agent by using MCTS variations. It would be interesting for instance to test other kinds of pruning. The use of traditional MCTS techniques such as Rapid Action Value Estimation or All Moves As First to reduce the number of simulations while keeping their intrinsic qualities would be interesting.

## References

- [1] Baarslag, T.: Exploring the Strategy Space of Negotiating Agents: A Framework for Bidding, Learning and Accepting in Automated Negotiation. Ph.D. thesis, Delft University of Technology (2016)
- [2] Baarslag, T., Hendriks, M.J.C., Hindriks, K.V., Jonker, C.M.: 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). <https://doi.org/10.1007/s10458-015-9309-1>
- [3] Baarslag, T., Hindriks, K.V.: Accepting optimally in automated negotiation with incomplete information. In: *AAMAS '13*. pp. 715–722. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2013), <http://dl.acm.org/citation.cfm?id=2484920.2485033>
- [4] Baarslag, T., Hindriks, K.V., Jonker, C.: A tit for tat negotiation strategy for real-time bilateral negotiation. In: *Complex Automated Negotiations: Theories, Models, and Software Competitions*, vol. 435, pp. 229–233. Springer Berlin Heidelberg (2013). [https://doi.org/10.1007/978-3-642-30737-9\\_18](https://doi.org/10.1007/978-3-642-30737-9_18)
- [5] Browne, C.C., Powley, E., Whitehouse, D., Lucas, S.M., Cowling, P.I., Rohlfshagen, P., Taverner, S., Perez, D., Samothrakis, S., Colton, S.: A survey of Monte Carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games* **4**(1), 1–43 (2012). <https://doi.org/10.1109/TCIAIG.2012.2186810>
- [6] Coulom, R.: Efficient selectivity and backup operators in monte-carlo tree search. In: *International conference on computers and games*. pp. 72–83. Springer (2006)
- [7] Couëtoux, A.: Monte Carlo Tree Search for Continuous and Stochastic Sequential Decision Making Problems. Ph.D. thesis, Université Paris XI (2013)
- [8] Dziuba, D.T.: Crowdfunding platforms in invoice trading as alternative financial markets. *Roczniki Kolegium Analiz Ekonomicznych/Szkoła Główna Handlowa* **49**, 455–464 (2018)

- [9] Fang, F., Xin, Y., Xia, Y., Haitao, X.: 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). <https://doi.org/10.1109/ISECS.2008.93>
- [10] Faratin, P., Jennings, N.R., Sierra, C.: Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems* **24**(3-4), 159–182 (1998). [https://doi.org/10.1016/S0921-8890\(98\)00029-3](https://doi.org/10.1016/S0921-8890(98)00029-3)
- [11] Finnsson, H.: Generalized Monte Carlo tree search extensions for general game playing. In: Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence. pp. 1550–1556. AAAI'12, AAAI Press (2012), <http://dl.acm.org/citation.cfm?id=2900929.2900948>
- [12] Fukuta, N., Ito, T., Zhang, M., Fujita, K., Robu, V. (eds.): Recent Advances in Agent-based Complex Automated Negotiation, Studies in Computational Intelligence, vol. 638. Springer International Publishing (2016)
- [13] Guttman, R.H., Moukas, A.G., Maes, P.: Agent-mediated electronic commerce: a survey. *The Knowledge Engineering Review* **13**(02), 147–159 (Jul 1998). [https://doi.org/10.1007/3-540-58266-5\\_1](https://doi.org/10.1007/3-540-58266-5_1), <http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=71015&fileId=S0269888998002082>
- [14] Hindriks, K., Tykhonov, D.: Opponent modelling in automated multi-issue negotiation using bayesian learning. In: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems. vol. 1, pp. 331–338 (2008), <http://dl.acm.org/citation.cfm?id=1402383.1402433>
- [15] de Jonge, D., Sierra, C.: Recent Advances in Agent-based Complex Automated Negotiation, Studies in Computational Intelligence, vol. 638, chap. GANGSTER: An Automated Negotiator Applying Genetic Algorithms, pp. 225–234. Springer International Publishing (2016). [https://doi.org/10.1007/978-3-319-30307-9\\_14](https://doi.org/10.1007/978-3-319-30307-9_14), [https://doi.org/10.1007/978-3-319-30307-9\\_14](https://doi.org/10.1007/978-3-319-30307-9_14)
- [16] de Jonge, D., Zhang, D.: Automated negotiations for general game playing. In: AAMAS '17. pp. 371–379. Richland, SC (2017), <http://dl.acm.org/citation.cfm?id=3091125.3091183>
- [17] Kocsis, L., Szepesvári, C.: Bandit based Monte Carlo planning. In: Fürnkranz, J., Scheffer, T., Spiliopoulou, M. (eds.) Machine Learning: ECML 2006: 17th European Conference on Machine Learning Berlin, Germany, September 18–22, 2006 Proceedings. pp. 282–293. Springer Berlin Heidelberg (2006). [https://doi.org/10.1007/11871842\\_29](https://doi.org/10.1007/11871842_29)
- [18] Lin, R., Kraus, S., Baarslag, T., Tykhonov, D., Hindriks, K., Jonker, C.M.: Genius: an integrated environment for supporting the design of generic automated negotiators. *Computational Intelligence* **30**(1), 48–70 (2014). <https://doi.org/10.1111/j.1467-8640.2012.00463.x>
- [19] Nash Jr, J.F.: The bargaining problem. *Econometrica: Journal of the Econometric Society* pp. 155–162 (1950)
- [20] Osborne, M.J., Rubinstein, A.: A course in game theory. MIT press, 12 edn. (1994)
- [21] Rasmussen, C.E., Williams, C.K.I.: Gaussian processes for machine learning. MIT Press (2006)
- [22] Rubinstein, A.: Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society* pp. 97–109 (1982)
- [23] Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., al.: Mastering the game of go with deep neural networks and tree search. *Nature* **529**(7587), 484–489 (2016)
- [24] Williams, C.R., Robu, V., Gerding, E.H., Jennings, N.R.: Using gaussian processes to optimise concession in complex negotiations against unknown opponents. In: IJCAI'11. pp. 432–438 (2011). <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-080>