may fluctuate, but also the *range* of future offers may be different at different times. Establishing this range is not easy, because the strategy used by the opponent is of course unknown to us. The offers of any time dependent opponent with incomplete information can again be modeled by a stochastic distribution, but this time the distribution will change over time. In terms of optimal stopping, this means that the bid distribution $X_j$ can be different for every $j$.

## 4.1 Uniformly Unpredictable Offers

If we assume that the opponent's offers are uniformly distributed, we only need to know the *interval* of utilities we can expect in every round. If this is the case, then we are able to compute the optimal time to accept, as is stated in the following proposition.

PROPOSITION 4.1. *Against a time-dependent opponent who, with j rounds still to be observed, makes bids uniformly distributed in $X_j = [a_j, b_j]$, the optimal stopping cut-off is $v_j$, where $v_j$ satisfies the following equation:*

$$v_j = \begin{cases} 0, & \text{if } j = 0 \\ v_{j-1}, & \text{if } v_{j-1} \geq b_{j-1} \\ \frac{b_{j-1}+a_{j-1}}{2}, & \text{if } v_{j-1} \leq a_{j-1} \\ v_{j-1} + \frac{1}{2} \cdot \frac{(b_{j-1}-v_{j-1})^2}{b_{j-1}-a_{j-1}}, & \text{if } a_{j-1} < v_{j-1} < b_{j-1}. \end{cases}$$

PROOF. From equation (2), we have

$$v_j = E(\max(X_{j-1}, v_{j-1})),$$

so immediately, if $v_{j-1} \geq b_{j-1}$, then $v_{j-1} \geq X_{j-1}$, and thus $v_j = v_{j-1}$. On the other hand, if $v_{j-1} \leq a_{j-1}$, then $v_j = E(X) = \frac{b_{j-1}+a_{j-1}}{2}$. So therefore, the only case left is $a_{j-1} < v_{j-1} < b_{j-1}$, in which case we derive the following:

$$\begin{aligned} v_j &= P(X_{j-1} \leq v_{j-1}) \cdot v_{j-1} + \\ &\quad P(X_{j-1} > v_{j-1}) \cdot P(X_{j-1} \mid X_{j-1} > v_{j-1}) \\ &= \frac{v_{j-1} - a_{j-1}}{b_{j-1} - a_{j-1}} \cdot v_{j-1} + \frac{1}{2} \cdot \frac{b_{j-1} - v_{j-1}}{b_{j-1} - a_{j-1}} \cdot (v_{j-1} + b_{j-1}) \\ &= v_{j-1} + \frac{1}{2} \cdot \frac{(b_{j-1} - v_{j-1})^2}{b_{j-1} - a_{j-1}}. \quad \square \end{aligned}$$

Note how this proposition is an extension of Proposition 3.1: if we set $X_j = [0, 1]$ for every $j$, the equation simplifies to $v_j = \frac{1}{2} + \frac{1}{2} v_{j-1}^2$ again.

Also, we observe that in the special case of perfect information, the distributions would be singletons of the form $X_j = \{x_j\}$, with probability 1 for the outcome $x_j$. The equation of Proposition 4.1 then simplifies to

$$\begin{aligned} v_j &= \begin{cases} 0, & \text{if } j = 0 \\ \max(x_{j-1}, v_{j-1}), & \text{otherwise.} \end{cases} \\ &= \max_{0 \leq k < j} x_k. \end{aligned}$$

This means that the optimal stopping procedure has the desirable property that when it gets perfect estimates as input, it will also produce perfect output.

## 4.2 Arbitrarily Unpredictable Offers

Proposition 4.1 is useful to gain insight into the optimal acceptance policy, but in practice, the distributions $X_j$ are neither known, nor uniformly distributed, and therefore an estimation method is required against arbitrary opponents. Of course, the success of the optimal stopping rules will greatly depend on the fidelity of the estimating technique used to predict the opponent's behavior. Therefore, we first examine the case of a perfect estimator, to see how our method performs in the ideal case. After that, we will move our focus to an estimator that can be used in practice.

### 4.2.1 *Opponent Prediction Using Perfect Estimates*

The perfect estimation method that we employ divides the number of rounds $N$ into a number of time slots $S$. Then, by momentarily using perfect information, it gets the minimum and maximum utility that will be offered by the opponent during that time slot. This allows us to control exactly the precision of the estimate, where using more slots emulates having more information about the opponent's behavior. If we set the number of slots equal to the total number of rounds, we are in a full information state and the performance should be theoretically optimal. If we use only one slot, we have less information, knowing only the opponent's utility range over the entire $N$ rounds.

### 4.2.2 *Opponent Prediction Using Gaussian Process Regression*

Finally, we consider an estimation method that uses as input only the information that can be observed during the negotiation, namely the utility of the offers made by the opponent. For this, we opted for a Gaussian process regression (GPR) technique as described in [24]. We selected the GPR technique because it can be computed in real time during the negotiation, and it is specifically designed to be robust with respect to significantly varying observations. It works as follows: for each offer made by the opponent, the round at which the offer was made is recorded, along with the offered utility. From this, the future concessions of the opponent are estimated using regression with a Gaussian process. To reduce the effect of noise, the offers received are aggregated in a number of time windows, and only the maximum value that is received in each time window is used as input for the Gaussian process.

The output of the Gaussian process regression is a normal distribution for every upcoming round $k$, with mean $\mu_k$ and standard deviation $\sigma_k$. The mean $\mu_k$ gives a prediction of the most likely offered utility value in round $k$, whilst the standard deviation $\sigma_k$ gives an indication of how accurate the prediction is. When using GPR, the opponent bid distribution is estimated in real-time by a *normal* distribution, truncated to fit in the range $[0, 1]$.

**Algorithm 2:** Determining $v_j$

```
Input: The number of remaining rounds j, and all
       negotiation outcomes Ω.
Output: v_j.
begin
    if j = 0 then
        return 0
    // Use either perfect estimation, or GPR
    Y_{j-1} ⟵ estimated utility distribution at j − 1;
    // Use either uniform, or Gaussian
       distribution for X_{j-1}
    X_{j-1} ⟵ utility distribution of Y_{j-1} over Ω;
    // Recursively determine v_{j-1}
    return E(max(X_{j-1}, v_{j-1}))
```

### 4.3 Experiments

To analyze the performance of optimal stopping (OS) against time dependent negotiation strategies, we adopted the same experimental setup as before, this time testing it with both perfect estimates and the GPR technique. We set the number of Gaussian process regressions to 10, and we set the number of samples equal to the number of rounds (for details, see [24]).

We used two versions of perfect estimation: the full information state by setting $S = N$ (called "perfect estimation with full slots"), and a version with $S = 1$ (called "perfect estimation with one slot"). We also tested two variants of the GPR technique: in one, we simply set $X_j$ equal to the truncated Gaussian distribution with mean $\mu_j$ and standard deviation $\sigma_j$ as predicted by the GPR technique (called "Gaussian GPR prediction"). These predictions turned out to be overly optimistic in most cases, since the GPR technique uses as input the maximum utility received in each time slot. Therefore, we opted to include a simplified second version, which produces a uniform distribution between zero and the estimated maximum offered utility, which we set to $\mu + 2\sigma$ (called "uniform GPR prediction"). See also Algorithm 2.

As the specifics of the negotiation scenario influences the behavior of the opponent, we picked a total of six negotiation scenarios from ANAC, aiming for a large spread of negotiation characteristics (see Table 1).

| Scenario | Size | Opposition |
|---|---|---|
| *ADG* [1] | 15625 | Low |
| *Amsterdam* [1] | 3024 | Medium |
| *England-Zimbabwe* [19] | 527 | Medium |
| *Nice Or Die* [3] | 3 | High |
| *Itex–Cypress* [14] | 180 | High |
| *Travel* [2] | 188160 | Medium |

**Table 1: Characteristics of the negotiation scenarios**

For the opponents, we selected various TDT's from [8]. Our optimal stopping policy works against *any* type of time dependent negotiation strategy, but we selected TDT's because they are typical, well-known examples of strategies that change their range of bids over time. Additionally, as in the case of *Random Walker*, TDT's are non-adaptive and hence it is not important what counter-offers are sent out to them.

We selected the same TDT's used earlier, namely: *Hardliner*, *Boulware*, *Conceder Linear*, and *Conceder*. We generated many variants of each opponent by choosing the values 0.7, 0.8, and 0.9 for the *min* parameter (which controls the utility threshold up to where the agent will concede [8]). Note that this creates quite a competitive opponent pool, as the opponents will never fully concede. This leads to only small utility differences between the different acceptance strategies, but this should be regarded an artifact of our competitive setup. The average score of all agents is shown in Figure 6.

Optimal stopping with perfect estimation with full slots should be considered the theoretical upper bound here; and indeed, it outperforms all other methods. Among the agents that act with *incomplete* information, the *Boulware* agent obtains a surprisingly good score. Its strategy turns out to be particularly successful against TDT's since it waits for a long time to let the opponent concede as much as possible, until it quickly concedes in the end to obtain an agreement.

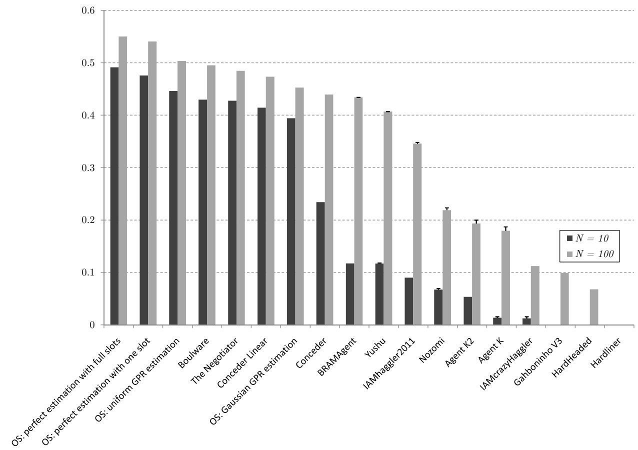


**Figure 6: The utilities obtained by ANAC agents and optimal stopping conditions with different estimation methods against time dependent opponents for 10 and 100 rounds.**

However, because it waits for so long, it misses out on good offers that are offered earlier.

Gaussian GPR prediction is not as successful, mainly because it was found to overestimate the opponent's willingness to concede, and hence it aimed for too much during the negotiation. It is optimal stopping with uniform GPR that performs significantly best (*one-tailed t-test*, $p < 0.01$), which shows that the optimal stopping policy is indeed a robust mechanism that can still perform well in an incomplete information setting.

## 5. RELATED WORK

As far as the authors are aware, this is the first work that deals with the optimal decision on the acceptance of an offer in a negotiation setting of incomplete information. In many settings of *complete* information ([21] is a typical example) the deal is usually formed right away and as such, sequential decisions whether to accept do not come into play.

In [26], a sequential decision making framework is also employed, using similar arguments for using it as we do. Furthermore, they also choose actions that maximize the expected payoff using a recursive formula; however, their approach uses Bayesian learning techniques and does not provide solutions specifically aimed at acceptance strategies.

The work by Fatima et al. [10] also treats optimal strategies in an incomplete information setting, but it primarily focuses on bidding strategies in the context of unknown deadlines and reservation values, and does not deal with acceptance strategies.

A work that comes closest to ours is [22], where optimal stopping is employed to decide when a party should reach an agreement in the context of conflict resolution. In contrast to our work, the scope of the paper is limited to simple bargaining games, and deals with one-sided incomplete information only.

## 6. DISCUSSION AND FUTURE WORK

This paper deals with the question of when to accept in a bilateral negotiation with incomplete information. Our approach has been to model the opponent's bids as a stochastic process, and to regard the decision of when to accept as a sequential decision problem. We first determined the optimal acceptance policies for particular opponent classes of which we were able to predict the behavior well. Of course, in the general case of unknown opponents, the solutions are only as good as the estimation of the opponent's behavior. We have shown however, that our techniques are robust, in the sense that they also perform well in practice. This demonstrates that our optimal stopping mechanism is a valuable element of a negotiating agent's strategy, whether in a complete or incomplete information setting.

Our work also opens up various lines of possible future work. First, there is an important aspect of negotiation that is not included in our model: negotiation is a dynamic two-party process. The opponent's behavior is influenced by whether or not we accept, and by what counter offers we make. In this paper we focused on the decision of when to accept, rather than on what counter offer should be generated in case the offer is unacceptable. Clearly, in practice, the bidding strategy is also very important in conducting a successful negotiation, and its effects on the acceptance strategy are yet to be determined.

Second, our model already incorporates the concept of negotiation costs, but in this paper we assumed them to be zero; however, it would be interesting to see the effects of costs on optimal acceptance behavior. Similarly, we also plan to study negotiation scenarios that have discounted payoffs. Both extensions will incentivize agents to employ more permissive acceptance conditions. On the other hand, adding reservation values to the agent's preferences would make an agent *less* inclined to accept. Combining this with discounted scenarios could induce agents to fall back on their reservation value by *ending* the negotiation prematurely. This 'outside option' gives rise to a new variety of optimal acceptance strategies that have to make the optimal choice between continuing, accepting, or walking away.

# 7. REFERENCES


[1] Tim Baarslag, Katsuhide Fujita, Enrico H. Gerding, Koen Hindriks, Takayuki Ito, Nicholas R. Jennings, Catholijn Jonker, Sarit Kraus, Raz Lin, Valentin Robu, and Colin R. Williams. Evaluating practical negotiating agents: Results and analysis of the 2011 international competition. *Artificial Intelligence Journal*, 2012.

[2] Tim Baarslag, Koen Hindriks, Catholijn M. Jonker, Sarit Kraus, and Raz Lin. The first automated negotiating agents competition (ANAC 2010). In *New Trends in Agent-based Complex Automated Negotiations*, pages 113–135, Berlin, Heidelberg, 2012. Springer-Verlag.

[3] Mai Ben Adar, Nadav Sofy, and Avshalom Elimelech. Gahboninho: Strategy for balancing pressure and compromise in automated negotiation. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, pages 205–208. Springer Berlin Heidelberg, 2013.

[4] Tibor Bosse and Catholijn M. Jonker. Human vs. computer behaviour in multi-issue negotiation. In *Proceedings of the Rational, Robust, and Secure Negotiation Mechanisms in Multi-Agent Systems*, RRS '05, pages 11–, Washington, DC, USA, 2005. IEEE Computer Society.

[5] Robert M. Coehoorn and Nicholas R. Jennings. Learning on opponent's preferences to make effective multi-issue negotiation trade-offs. In *Proceedings of the 6th international conference on Electronic commerce*, ICEC '04, pages 59–68, New York, USA, 2004. ACM.

[6] Morris H. DeGroot. *Optimal statistical decisions*. McGraw-Hill, New York, 1970.

[7] A.S.Y. Dirkzwager, M.J.C. Hendrikx, and J.R. Ruiter. The negotiator: A dynamic strategy for bilateral negotiations with time-based discounts. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, pages 217–221. Springer Berlin Heidelberg, 2013.

[8] P. Faratin, C. Sierra, and N. R. Jennings. Negotiation decision functions for autonomous agents. *Int. Journal of Robotics and Autonomous Systems*, 24(3-4):159–182, 1998.

[9] S. S. Fatima, M. Wooldridge, and N. R. Jennings. An agenda-based framework for multi-issue negotiation. *Artificial Intelligence*, 152(1):1–45, 2004.

[10] S. Shaheen Fatima, Michael Wooldridge, and Nicholas R. Jennings. Optimal negotiation strategies for agents with incomplete information. In *Revised Papers from the 8th International Workshop on Intelligent Agents VIII*, ATAL '01, pages 377–392, London, UK, UK, 2002. Springer-Verlag.

[11] Asaf Frieder and Gal Miller. Value model agent: A novel preference profiler for negotiation with agents. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, pages 199–203. Springer Berlin Heidelberg, 2013.

[12] Dhananjay K. Gode and Shyam Sunder. Allocative efficiency in markets with zero intelligence (zi) traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119–137, 1993.

[13] Shogo Kawaguchi, Katsuhide Fujita, and Takayuki Ito. Agentk2: Compromising strategy based on estimated maximum utility for automated negotiating agents. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, pages 235–241. Springer Berlin Heidelberg, 2013.

[14] Gregory E. Kersten and Grant Zhang. Mining inspire data for the determinants of successful internet negotiations. *InterNeg Research Papers INR 04/01 Central European Journal of Operational Research*, 2003.

[15] Sarit Kraus, Jonathan Wilkenfeld, and Gilad Zlotkin. Multiagent negotiation under time constraints. *Artificial Intelligence*, 75(2):297 – 345, 1995.

[16] Thijs Krimpen, Daphne Looije, and Siamak Hajizadeh. Hardheaded. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, pages 223–227. Springer Berlin Heidelberg, 2013.

[17] Bjorn Leonardz. *To stop or not to stop. Some elementary optimal stopping problems with economic interpretations.* Almqvist & Wiksell, Stockholm,, 1973.

[18] 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*, 2012.

[19] Raz Lin, Sarit Kraus, Jonathan Wilkenfeld, and James Barry. Negotiating with bounded rational agents in environments with incomplete information using an automated agent. *Artificial Intelligence*, 172(6-7):823 – 851, 2008.

[20] Chhaya Mudgal and Julita Vassileva. Bilateral negotiation with incomplete and uncertain information: A decision-theoretic approach using a model of the opponent. In *Proceedings of the 4th International Workshop on Cooperative Information Agents IV, The Future of Information Agents in Cyberspace*, CIA '00, pages 107–118, London, UK, UK, 2000. Springer-Verlag.

[21] Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–109, 1982.

[22] Santiago Sánchez-Pagés. The use of conflict as a bargaining tool against unsophisticated opponents. ESE discussion papers, Edinburgh School of Economics, University of Edinburgh, 2004.

[23] Ingolf Stahl. *Bargaining theory*. Economic Research Institute, Stockholm, 1972.

[24] 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 *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*. AAAI Press, January 2011.

[25] Colin R. Williams, Valentin Robu, Enrico H. Gerding, and Nicholas R. Jennings. Iamhaggler: A negotiation agent for complex environments. In *New Trends in Agent-based Complex Automated Negotiations*, pages 151–158, Berlin, Heidelberg, 2012. Springer-Verlag.

[26] D. Zeng and K. Sycara. Bayesian learning in negotiation. *International Journal of Human Computer Systems*, 48:125–141, 1998.