

# IAMhaggler2011: A Gaussian Process Regression Based Negotiation Agent

Colin R. Williams, Valentin Robu, Enrico H. Gerding, and Nicholas R. Jennings

**Abstract.** We describe the strategy used by our agent, IAMhaggler2011, which finished in third place in the 2011 Automated Negotiating Agent Competition. A key feature of this agent is the way in which it models the likely negotiation behaviour of its opponent. Specifically, it first uses a Gaussian process regression technique to estimate the future concession of its negotiation opponent. Its concession is then set as a best response to this prediction.

## 1 Introduction

In this work, we give a brief overview of the core parts of the strategy used by our agent, *IAMhaggler2011*, which finished in third place in the second international Automated Negotiating Agent Competition (ANAC2011). The competition setup is described in [1]. The overall framework used by the agent is based on that of our previous (2010) competition entry, *IAMhaggler*[3]. A more detailed description of *IAMhaggler2011* and an analysis of its performance can be found in [4].

The core of the strategy used by *IAMhaggler2011* consists of three parts, which we describe in turn. The first predicts the concession of the opponent (Section 2). The second sets the concession rate such that it optimises the expected utility given that prediction (Section 3). The final part generates a multi-issue offer according to the concession rate (Section 4).

## 2 Predicting the Opponent's Concession

In order to set its behaviour as a best response to that of the opponent, our agent first needs to predict how the opponent will concede throughout the negotiation, using

---

Colin R. Williams · Valentin Robu · Enrico H. Gerding · Nicholas R. Jennings  
School of Electronics and Computer Science, University of Southampton,  
University Road, Southampton, SO17 1BJ  
e-mail: {crw104, vr2, eg, nrj}@ecs.soton.ac.uk

only the information which it can observe: the offers made by the opponent, and the utility of these offers according to our agent's utility function. This is done using a Gaussian process regression technique, in order to provide both a prediction of the opponent's future concession and a measure of the level of confidence in that prediction (see [2] for details).

The output of the Gaussian process is a Gaussian distribution, for each time  $t$ , denoted by  $f(u; \mu_t, \sigma_t)$ , where the mean,  $\mu_t$ , gives an indication of the most likely value for  $u$  at time  $t$ , whilst the standard deviation,  $\sigma_t$  is an indication of how accurate the prediction of  $\mu_t$  is likely to be.

As input to the Gaussian process, *IAMhaggler2011* uses the maximum value offered by the opponent in a particular time window of a fixed duration, and the time of that window. This windowed approach reduces the effect of noise on the Gaussian process and reduces the amount of input data. The maximum value in each time window is used, rather than the average, as the maximum represents the best offer that has been observed, and therefore that our agent can expect to reach agreement at.

### 3 Setting the Concession Rate

We now show how the prediction of the opponent's future concession is used by our strategy to set its concession rate by optimising the expected utility given that prediction. Specifically, we discuss the way in which both the mean,  $\mu$ , and standard deviation,  $\sigma$ , output by the Gaussian process are used in setting an optimal concession rate. Optimal in this context means that our strategy maximises the agent's expected utility given its prediction of the opponent. The aim is to calculate the best time,  $t^*$  and utility value,  $u^*$  at which to reach agreement. We consider the best time,  $t^*$ , to be the point in future time ( $t \in [t_c, t_{\max}]$ ) at which the expected utility,  $E_{\text{rec}}(t)$  of the opponent's offer is likely to be maximised. The expected utility of reaching an agreement at time  $t$  is given by  $E_{\text{rec}}(t) = \int_0^1 D(R(u)f(u; \mu_t, \sigma_t), t) du$  where  $D(\cdot, \cdot)$  is the effect of the discounting factor,  $R(\cdot)$  is the risk function, and  $f(\cdot)$  is the probability distribution over the values of  $u$ , as given by the Gaussian process.<sup>1</sup>

Our strategy considers the risk associated with reaching agreement at a particular utility. In a tournament setting, such as the one used in ANAC2011, the primary aim of the strategy is to 'win' the negotiation by achieving a higher utility than that of its opponent. In such a setting, reaching an agreement with a high utility is not always good enough, since it is possible that the opponent may have achieved an even higher utility, thereby winning the negotiation. Consequently, the strategy may need to take a more aggressive approach than it would if it were simply maximising its own utility.

---

<sup>1</sup> Due to the constraints on the utility functions used in the competition, we assume that the utility of the opponent's offers must lie in the range  $[0, 1]$ . Therefore we use a truncated normal distribution, with the utility constrained to fit in that range.

In order to deal with this trade-off, we adjust the behaviour of our agent by including the concept of risk attitude in the design of our concession strategy. Formally, the risk function used is  $R(u) = u^r$  where  $r$  is the risk parameter. If  $r = 1$ , the strategy is risk-neutral, for  $r > 1$ , the strategy is risk-seeking and for  $r < 1$ , it is risk-averse. A risk-seeking strategy would result in more aggressive behaviour, since such an agent will concede more slowly, as it will regard lower utilities to be of even lower value than their true value. This has the effect of making the agent more likely to win in a tournament setting, and is therefore more appropriate in the negotiation competition.

Thus, the effect of the standard deviation,  $\sigma_t$  is as follows. If, at two points in time,  $t_1$  and  $t_2$ , the mean values are the same ( $\mu_{t_1} = \mu_{t_2}$ ), but the standard deviation differs such that  $\sigma_{t_1} < \sigma_{t_2}$ , then a risk-seeking agent will consider the expected utility at time  $t_2$  to be greater than at time  $t_1$ . That is, the risk-seeking agent is prepared to wait for the less certain offer at time  $t_2$ , as there is a higher chance that the utility may exceed the value of  $\mu_{t_2}$ , than it would for  $\mu_{t_1}$ . The converse applies for risk-averse agents, with risk-neutral agents being indifferent between the two solutions. In the ACAN 2011 tournament, we used a risk-seeking approach, with  $r = 3$ .

Having selected the time,  $t^*$ , at which the expected utility of the opponent's offers is maximised, our agent needs to choose a utility,  $u^*$ , to offer at that time. The approach that our strategy takes here is to maximise the expected utility, of making an offer of utility  $u$  at time  $t^*$ , as given by  $u^* = \arg\max_{u \in [0,1]} E_{\text{offer}}(u, t^*)$ .

The expected utility is calculated based on the probability that an offer of a given utility will be accepted. We assume that an offer of utility  $u$  will be accepted at time  $t^*$  if  $u \leq u_{t^*}$ . Since we have a probability distribution over  $u_{t^*}$ , we can calculate the probability that  $u \leq u_{t^*}$  using the cumulative distribution  $F(u; \mu_t, \sigma_t)$ . Therefore, our agent's expectation for the adjusted utility (taking into consideration the risk attitude and time discounting) of an offer with utility  $u$  is given by  $E_{\text{offer}}(u, t^*) = D(R(u)F(u; \mu_{t^*}, \sigma_{t^*}), t^*)$  where  $D(\cdot, \cdot)$  and  $R(\cdot)$  are as before, and  $F(\cdot)$  is the c.d.f. for  $f(\cdot)$ . Due to the effect of the standard deviation, note that the risk-seeking agent will find a higher optimal utility than a risk-neutral or risk-averse agent.

Finally, having determined  $u^*$  as the utility to offer at time  $t^*$ , our agent needs to choose a target utility,  $u_t$  to offer at the current time,  $t_c$ . The agent should not concede immediately to offer  $u^*$  at the current time, nor should it wait until  $t^*$ . Either of those approaches is likely to result in concession behaviour which is too extreme, especially since they are based on predictions which may be inaccurate. Therefore, and to avoid any additional parameters, our approach is simply to concede linearly between  $[t_{lr}, u_{lr}]$  and  $[t^*, u^*]$ , where  $t_{lr}$  is the time at which the regression was last performed and  $u_{lr}$  is the target utility at that time. The overall concession will not generally be linear, as the predictions of  $t^*$  and  $u^*$  are revised at the end of each time window, throughout the negotiation.

## 4 Generating an Offer

Having selected a target utility,  $u_\tau$ , our strategy needs to generate an offer which has a utility close to that target. In a multi-issue negotiation, there may be many different packages with a similar utility. Under the real-time constraints of the competition, the goal is to reach an agreement within a shorter time period, but not necessarily to limit the number of offers made.

Consequently, our agent uses a fast strategy, which simply chooses a random offer with a utility in the range  $[u_\tau - 0.025, u_\tau + 0.025]$ . A range is used, since, in a domain which consists entirely of discrete issues, it is highly likely that there are no possible outcomes with a utility of exactly  $u_\tau$ . If an offer cannot be found even within this range, the range is expanded, until a solution is found. In addition, if the target drops below the highest value of the offers made by the opponent, we instead propose the package with that utility that was offered by the opponent. This is since we assume that, for a set of possible offers with utility greater than  $u_\tau$ , the one which is most likely to be accepted is the one which has previously been offered by the opponent.

## 5 Conclusions and Future Work

In this work, we have developed a negotiation strategy that uses a principled approach to concession, by setting its behaviour as a best response, according to its estimate of the opponent's future concession. In future work, we plan to improve our offer generation algorithm by modelling the opponent's preferences, and then choose a package at a given utility level which maximises the likelihood that it is accepted by the opponent.

## References

1. Baarslag, T., Hindriks, K., Jonker, C.M., Kraus, S., Lin, R.: The second automated negotiating agents competition (ANAC2011). In: *Complex Automated Negotiations*. SCI, vol. 435, pp. 183–197. Springer, Heidelberg (2012)
2. Rasmussen, C.E., Williams, C.K.I.: *Gaussian Processes for Machine Learning*. The MIT Press (2006)
3. Williams, C.R., Robu, V., Gerding, E.H., Jennings, N.R.: IAMhaggler: A Negotiation Agent for Complex Environments. In: Ito, T., Zhang, M., Robu, V., Fatima, S., Matsuo, T. (eds.) *New Trends in Agent-Based Complex Automated Negotiations*. SCI, vol. 383, pp. 151–158. Springer, Heidelberg (2012)
4. Williams, C.R., Robu, V., Gerding, E.H., Jennings, N.R.: Using gaussian processes to optimise concession in complex negotiations against unknown opponents. In: *Proc. of the Twenty-Second International Joint Conference on Artificial Intelligence*, vol. 1, pp. 432–438 (2011)