Evaluating practical negotiating agents: Results and analysis of the 2011 international competition ☆

Tim Baarslag [a], Katsuhide Fujita [c], Enrico H. Gerding [b], Koen Hindriks [a], Takayuki Ito [d], Nicholas R. Jennings [b], Catholijn Jonker [a], Sarit Kraus [e,f], Raz Lin [e], Valentin Robu [b], Colin R. Williams [b,∗]

[a] Man–Machine Interaction Group, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
[b] Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, United Kingdom
[c] Institute of Engineering Innovation, Graduate School of Engineering, The University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113-8656, Japan
[d] Department of Computer Science and Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
[e] Department of Computer Science, Bar-Ilan University, Ramat-Gan, 52900 Israel
[f] Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

ABSTRACT

This paper presents an in-depth analysis and the key insights gained from the Second International Automated Negotiating Agents Competition (ANAC 2011). ANAC is an international competition that challenges researchers to develop successful automated negotiation agents for scenarios where there is no information about the strategies and preferences of the opponents. The key objectives of this competition are to advance the state-of-the-art in the area of practical bilateral multi-issue negotiations, and to encourage the design of agents that are able to operate effectively across a variety of scenarios. Eighteen teams from seven different institutes competed. This paper describes these agents, the setup of the tournament, including the negotiation scenarios used, and the results of both the qualifying and final rounds of the tournament. We then go on to analyse the different strategies and techniques employed by the participants using two methods: (i) we classify the agents with respect to their concession behaviour against a set of standard benchmark strategies and (ii) we employ empirical game theory (EGT) to investigate the robustness of the strategies. Our analysis of the competition results allows us to highlight several interesting insights for the broader automated negotiation community. In particular, we show that the most adaptive negotiation strategies, while robust across different opponents, are not necessarily the ones that win the competition. Furthermore, our EGT analysis highlights the importance of considering metrics, in addition to utility maximisation (such as the size of the basin of attraction), in determining what makes a successful and robust negotiation agent for practical settings.

© 2012 Elsevier B.V. All rights reserved.

Glossary

Glossary of notation

d: discounting factor
P: set of players

p: individual player
p′: opponent of player *p*
S: set of strategies of all the players
S(p): specific strategy of player *p*
t: time
$U(s_1, s_2)$: utility achieved by player using strategy s_1 when negotiating against a player using strategy s_2
$U_D^t(s_1, s_2)$: discounted utility achieved at time *t* by player using strategy s_1 when negotiating against a player using strategy s_2
$U_1(\omega)$: utility to player 1 of outcome ω
$U_2(\omega)$: utility to player 2 of outcome ω
Ω: domain
ω: outcome
ω_{Nash}: Nash bargaining solution

1. Introduction

In May 2011 we held the Second International Automated Negotiating Agents Competition (ANAC 2011)[1] in conjunction with the Tenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011). This competition follows in the footsteps of a series of successful competitions that aim to advance the state-of-the-art in artificial intelligence (other examples include the Annual Computer Poker Competition[2] and the various Trading Agent Competitions (TAC) [45]). ANAC focuses specifically on the design of practical negotiation strategies. In particular, the overall aim of the competition is to advance the state-of-the-art in the area of bilateral, multi-issue negotiation, with an emphasis on the development of successful automated negotiators in realistic environments with incomplete information (where negotiators do not know their opponent's strategy, nor their preferences) and continuous time (where the negotiation speed and number of negotiation exchanges depends on the computational requirements of the strategy). More specifically still, the stated goals of the competition include: (i) encouraging the design of agents that can proficiently negotiate in a variety of circumstances, (ii) objectively evaluating different negotiation strategies, (iii) exploring different learning and adaptation strategies and opponent models, and (iv) collecting state-of-the-art negotiating agents and negotiation scenarios, and making them available and accessible as benchmarks for the negotiation research community.

Negotiation is an important topic of research in a number of disciplines. However, in recent years there has been an increasing focus on the design of *automated* negotiators, i.e., autonomous agents capable of negotiating with other agents in a specific environment [18,21]. Furthermore, many of these agents, e.g. [8,11,17,19,22,27], are designed to operate in specific and relatively simple scenarios, and are often based on simplified assumptions, that do not model well practical, real-life settings where negotiation may be applied. For example, such work often assumes that the opponent strategies and preferences are known or partially known. This is unrealistic especially in multi-issue negotiations, where agents can have different preferences for these issues, and they are unlikely to reveal them. Some work does consider settings with incomplete information, but still assumes agents to have partial information about the opponent's preferences [8,19], or probabilistic information about the opponents [11], which is again often not available in practice. Another assumption, both in the heuristic and game-theoretic literature, is that negotiation interactions occur at fixed time intervals and that negotiation ends after a fixed number of rounds [40]. In practice, however, negotiation often happens in real time, and the time required to reach an agreement depends on the deliberation time of the agents (i.e., the amount of computation required to evaluate an offer and produce a counter offer). This is particularly important when utility is *discounted*, i.e. when the value of an agreement decreases over time, or when there is a deadline. Furthermore, another problem is that many negotiation strategies are evaluated against relatively simple strategies, such as the time-dependent tactics introduced in [7], instead of the state-of-the-art. This is partly due to the fact that the results of the different implementations are difficult to compare, as various setups are used for experiments in ad hoc negotiation environments [16,26].

Against this background, the competition was established to address the above shortcomings, and to enable negotiating agents to be evaluated in realistic environments and with a wide variety of opponents and scenarios. Moreover, since the opponents, as well as the scenarios in which negotiation occurs are unknown in advance, competition participants are compelled to design *generic* negotiation agents that perform effectively in a variety of circumstances. These agents, together with a wide range of negotiation scenarios, provide a comprehensive repository against which negotiation agents can be benchmarked. This, in turn, allows the community to push forward the state-of-the-art in the development of automated negotiators and their evaluation and comparison to other automated negotiators.

To achieve this, GENIUS,[3] a **G**eneral **E**nvironment for **N**egotiation with **I**ntelligent multi-purpose **U**sage **S**imulation, was developed. GENIUS [26] is the underpinning platform of ANAC and it allows easy development and integration of existing negotiating agents. It can be used to simulate individual negotiation sessions, as well as tournaments between negotiating agents in various negotiation scenarios. It implements an open architecture that allows easy development and integration of existing negotiating agents using design patterns. Lastly, it enables the specification of negotiation domains and utility functions by means of a graphical user interface.

---

[1] http://www.itolab.nitech.ac.jp/ANAC2011/.
[2] http://www.computerpokercompetition.org.
[3] http://mmi.tudelft.nl/genius.

In more detail, eighteen teams (as compared to seven in the first competition) from seven different institutes (University of Alcalá, Bar-Ilan University, Ben-Gurion University, Politehnica University of Bucharest, Delft University of Technology, Nagoya Institute of Technology, and University of Southampton) and six different countries (Spain, Israel, Romania, the Netherlands, Japan, and the United Kingdom) submitted negotiating agents to the tournament. Eight of these teams continued to the finals after undergoing a qualifying round.

In this paper we present detailed results from the tournament, as well as an in-depth analysis of the strategies from the finalists and the techniques employed by the different agents. In particular, we evaluate the agents using two approaches. First, we consider the concession behaviour of each agent and how they perform when faced with different types of opponents. This provides an insight into what type of behaviour (e.g. being adaptive or hardheaded) is more likely to do well in different situations. Second, we use *empirical game theory* (EGT) to investigate the *robustness* of the strategies. That is, we study how well the strategies perform against different combinations of strategies and, from the strategies in the finals, which of these form either an *equilibrium* (where there is no incentive to deviate to another strategy) or, when no equilibrium exists, a *best reply cycle* (where there is a set of profiles without incentives to deviate a profile outside that set).

In terms of concessive behaviour, analysing the competition results yields some interesting insights into the properties exhibited by agents which successfully negotiate in realistic negotiation environments. We develop a measure of the concession rate and compare, for each strategy, the correlation between the total amount of concession made during the negotiation, and the utility it achieves. We find that there is a direct correlation between the performance of a strategy and its concession rate against a simple *Linear Conceder* strategy, but there is less correlation between performance and concession against a non-concessive, *Hardliner* strategy. This shows that the best strategies try to exploit concessive opponents, but against a non-concessive opponent, some aim to reach any agreement (even with a low utility) while others prefer not to concede and may therefore may receive the disagreement payoff (which is zero).

In terms of robustness, the EGT analysis shows that there are considerable differences in the performance of a negotiation strategy, depending on the pool of agents it is negotiating against and on the criteria used for ranking. For example, some strategies only do well in a tournament setting if they can exploit weaker opponents, but do not necessarily perform well when faced against the top players, or indeed in self-play. This highlights the importance of considering a variety of criteria, besides utility maximisation, when evaluating negotiation agents, and in determining what makes a successful and robust negotiation agent in practical settings. Such criteria may include the existence of either a pure Nash equilibria or a best reply cycle as stable sets of a set of deviations, as well as the size of the basins of attraction for each of these stable sets. We believe that these insights carry over to research on designing negotiating agents outside the setting of the competition.

The remainder of the paper is structured as follows. First, we introduce the competition environment in Section 2. The participating agents and the submitted scenarios are then described in Section 3, followed by the competition results in Section 4. The performance of the automated negotiators is then further evaluated in Section 5. We collect the insights gathered by this analysis in Section 6 and conclude in Section 7.

2. The competition setup

Given the goals outlined in the introduction, in this section we introduce the format of the 2011 competition. We begin by describing the negotiation model that is used during each negotiation encounter (Section 2.1), and then we describe how the competition tournament is formed as a series of such encounters (Section 2.2).

2.1. The negotiation model

In this competition, we consider *bilateral* negotiations, i.e. negotiation between two parties. The interaction between negotiating parties is regulated by a *negotiation protocol* that defines the rules of how and when proposals can be exchanged. Specifically, we use the alternating-offers protocol for bilateral negotiation as formalised in [37], in which the negotiating parties exchange offers in turns. We choose this protocol due to its simplicity, and moreover, it is a protocol which is widely studied and used in the literature, both in game-theoretic and heuristic settings (a non-exhaustive list includes [9,21,23,31, 32]).

Now, the parties negotiate over a set of *issues*, and every issue has an associated range of alternatives or *values*. A negotiation *outcome* consists of a mapping of every issue to a value, and the set, Ω, of all possible outcomes is called the negotiation *domain*. The domain is common knowledge to the negotiating parties and stays fixed during a single negotiation encounter. In ANAC 2011, we focused on settings with a finite set of discrete values per issue.

In addition to the domain, both parties also have privately-known preferences described by their *utility functions*. Each utility function, U, maps every possible outcome $\omega \in \Omega$ to a real-valued number in the range $[0, 1]$. In ANAC 2011, the utilities are *additive*. That is, the overall utility consists of a weighted sum of the utility for each individual issue. While the domain (i.e. the set of outcomes) is common knowledge, the utility function of each player is private information. This means that the players do not have access to the utility function of the opponent. In more detail, even the opponent's orderings of the issue values are unknown, and the agents are not provided with any prior distribution over the utility

functions. However, the player can attempt to learn during the negotiation encounter.[4] Moreover, we use the term *scenario* to refer to the domain and the pair of utility functions (for each agent) combined.

Finally, we supplement the scenario with a deadline and discount factors. The reasons for doing so are both pragmatic and to make the competition more interesting from a theoretical perspective. Without a deadline or discounting factor, the negotiation might go on forever. Also, with unlimited time an agent may simply try a large number of proposals to learn the opponent's preferences. In addition, as opposed to having a fixed number of rounds, both the discount factor and deadline are measured in *real time*. This, in turn, introduces another factor of uncertainty since it is now unclear how many negotiation rounds there will be, and how much time an opponent requires to compute a counter offer. Also, this computational time will typically change depending on the size of the outcome space. In ANAC 2011, the discount factors depend on the scenario, but the deadline is fixed and is set to three minutes in total.[5] The implementation of discount factors in ANAC 2011 is as follows. Let d in $[0, 1]$ be the discount factor. Let t in $[0, 1]$ be the current time, normalised between the beginning of the negotiation and the deadline. We compute the discounted utility U_D^t as follows:

$$U_D^t(s_1, s_2) = U(s_1, s_2) \cdot d^t \tag{1}$$

If $d = 1$, the utility is not affected by time, and such a scenario is considered to be undiscounted, while if d is very small there is high pressure on the agents to reach an agreement. Note that, in the setup used in ANAC 2011, discount factors are part of the scenario, are known to both agents and are always *symmetric* (i.e. d always has the same value for both agents).

2.2. Running the tournament

GENIUS incorporates several mechanisms that support the design of a general automated negotiator. The first mechanism is an analytical toolbox, which provides a variety of tools to analyse the performance of agents, the outcome of the negotiation and its dynamics. The second mechanism is a repository of scenarios. Lastly, it also comprises repositories of automated negotiators. In addition, GENIUS enables the evaluation of different strategies used by automated agents that were designed using the tool. This is an important contribution as it allows researchers to empirically and *objectively* compare their agents with others in different domains and settings.

The time-line of ANAC 2011 consists of three phases: the qualifying round, the updating period and the final round. The domains and utility functions used during the competition were not known in advance and were designed by the participants themselves (see Section 3). Therefore, in a given negotiation, an agent does not know the utility function of its opponent, apart from that the fact that it is additive. In more detail, the participants have no prior knowledge of the distribution over the function's parameters and, furthermore, they do not even know the opponent's preference ordering over the values for an individual issue.[6] An agent's success is measured according to the average utility achieved in all negotiations of the tournament for which it is scheduled.

First, a *qualifying round* was played in order to select the best 8 agents from the 18 agents that were submitted by the participating teams. Each participant also submitted a domain and pair of utility functions for that domain. All these scenarios were used in the qualifying rounds. For each of these scenarios, negotiations were carried out between all pairings of the 18 agents.

Since there were 18 agents, which each negotiate against 17 other agents, in 18 different domains, a single tournament in the qualifying round consists of $18 \times 17/2 \times 2 \times 18 = 5508$ negotiation sessions.[7] To reduce the effect of variation in the results, the tournament was repeated 3 times, leading to a total of 16,524 negotiation sessions, each with a time limit of three minutes. In order to complete such an extensive set of tournaments within a limited time frame, we used five high-spec computers, made available by Nagoya Institute of Technology. Specifically, each of these machines contained an *Intel Core i7* CPU and at least 4 GB of DDR3 memory.

The 8 agents that achieved the best average scores during qualifying were selected as participants for the final round. Between the rounds, we allowed a two week updating period, in which the 8 selected finalists were given the chance to improve their agents for the final round. The detailed results and all scenarios for the qualifying round were revealed to all finalists, and they could use this additional information to tune their agents.

The domains and utility functions in the final were the 8 scenarios submitted by all finalists for the final round. The entire set of pairwise matches were played among 8 agents, and the final ranking of ANAC 2011 was decided. In the final,

---

[4] Note however, that the player can only perform this learning *within* a negotiation session and that any learning cannot be used *between* different negotiation encounters. This is done so that agents need to be designed to deal with unknown opponents. In order to prevent the agents learning across instances, the competition is set up so that a new agent instance is created for each negotiation. The rules prohibit the agents storing data on disk, and they are prevented from communicating via the Internet.

[5] In contrast, in ANAC 2010, the agents had three minutes *each* to deliberate. This means the agents had to keep track of both their own time and the time the opponent had left, otherwise they run the risk of the opponent walking away unexpectedly.

[6] The pairs of utility functions which form a scenario are designed by the participants (who also develop the agents), but the rules prohibit designing an agent to detect a particular scenario (and therefore the opponent's utility function) based on knowledge of such a pair of functions. Furthermore, the scenarios were changed during the updating period so that the finalists would not benefit from tuning their strategies to the scenarios of the qualifying round.

[7] The combinations of 18 agents are $18 \times 17/2$, however, agents play each domain against each other twice (once for each profile).

Table 1
The scenarios used in ANAC 2011 final round, ordered in increasing size.

Scenario name	Number of issues	Size of outcome space	Opposition	Discount factor	Mean distance to Pareto frontier
Nice or Die	1	3	0.840	1.000	0.000
Laptop	3	27	0.160	0.424	0.295
Company Acquisition	5	384	0.117	0.688	0.121
Grocery	5	1600	0.191	0.806	0.492
Amsterdam Trip	6	3024	0.223	1.000	0.254
Camera	6	3600	0.219	0.891	0.450
Car	6	15,625	0.092	1.000	0.136
Energy	8	390,625	0.445	1.000	0.149

a single tournament consists of $8 \times 7/2 \times 2 \times 8 = 448$ negotiation sessions. Again, each tournament was repeated three times.

In order to improve the statistical significance of the results reported in this work, we subsequently carried out a total of 30 repetitions of the tournament, using the University of Southampton's IRIDIS high-performance computing cluster. The results that were obtained by this more extensive set of tests are consistent with those from the original final round. For consistency, throughout the remainder of this work, we present the results according to these 30 repetitions, and all evaluations are based on these results.

3. Competition scenarios and agents

Having introduced the setup of the competition, we now describe the negotiation scenarios (Section 3.1) and agents (Section 3.2) entered in the 2011 competition.

3.1. Scenario descriptions

The competition is targeted towards modelling multi-issue negotiations in uncertain, open environments, in which agents do not know the preferences of their opponent. The various characteristics of a negotiation scenario such as size, number of issues, opposition and discount factor can have a significant influence on the negotiation outcome [16]. Therefore, in order to ensure a good spread of negotiation characteristics, and to reduce any possible bias on the part of the organisers, we gathered the domains and profiles from the participants in the competition. Specifically, in addition to submitting their agents, each participant submitted a scenario, consisting of both a domain and a pair of utility functions. In the qualifying round, we used all 18 scenarios submitted by the participants. In the final round, the eight scenarios submitted by the eight finalists were used. The final scenarios vary in terms of the number of issues, the number of possible outcomes, the opposition of the utility functions and the mean distance of all of the points in the outcome space to the Pareto frontier (see Table 1). The opposition of the utility functions is determined by the minimum distance from all of the points in the outcome space to the point which represents complete satisfaction of both parties $(1, 1)$. Formally:

$$\text{opposition}(\Omega) = \min_{\omega \in \Omega} \text{dist}(\omega, \overline{\omega}) \tag{2}$$

where Ω is the set of all possible outcomes, $\overline{\omega}$ is the point in the outcome space at which both parties would receive their maximum utility (though it is not a possible outcome) and the 'dist' function is the distance between two points in the outcome space, as defined in (4). The mean distance to the Pareto frontier is defined formally as:

$$\text{meanParetoDistance}(\Omega) = \sum_{\omega \in \Omega} \frac{\min_{\omega_P \in \Omega_P} \text{dist}(\omega, \omega_P)}{|\Omega|} \tag{3}$$

where $\Omega_P \subset \Omega$ is the set of Pareto efficient possible outcomes. The 'dist' function gives the Euclidean distance between two points in the outcome space, defined formally as:

$$\text{dist}(\omega_1, \omega_2) = \sqrt{\left(U_1(\omega_1) - U_1(\omega_2)\right) * \left(U_2(\omega_1) - U_2(\omega_2)\right)} \tag{4}$$

where $U_1(\cdot)$ and $U_2(\cdot)$ give the utilities to players 1 and 2, respectively.

The properties of the scenarios can also be observed in the shape of the outcome space of each scenario, as presented graphically in Fig. 1. In more detail, very large scenarios, such as the *Energy* scenario, are displayed with a large number of points representing the many possible agreements, whereas smaller scenarios, such as *Nice Or Die*, have only very few points. Furthermore, scenarios which have a high mean distance to the Pareto frontier, such as the *Grocery* and *Camera* scenarios, appear very scattered, whereas those with a low mean distance, such as the *Company Acquisition* scenario, are much more tightly clustered. The other 10 scenarios (which were eliminated along with their agents in the qualifying round) contained broadly similar characteristics to those of the final 8 scenarios. Therefore, since the final 8 scenarios capture a good distribution of the characteristics we would like to examine, in the rest of this work, we consider only these scenarios, which are described in more detail as follows:
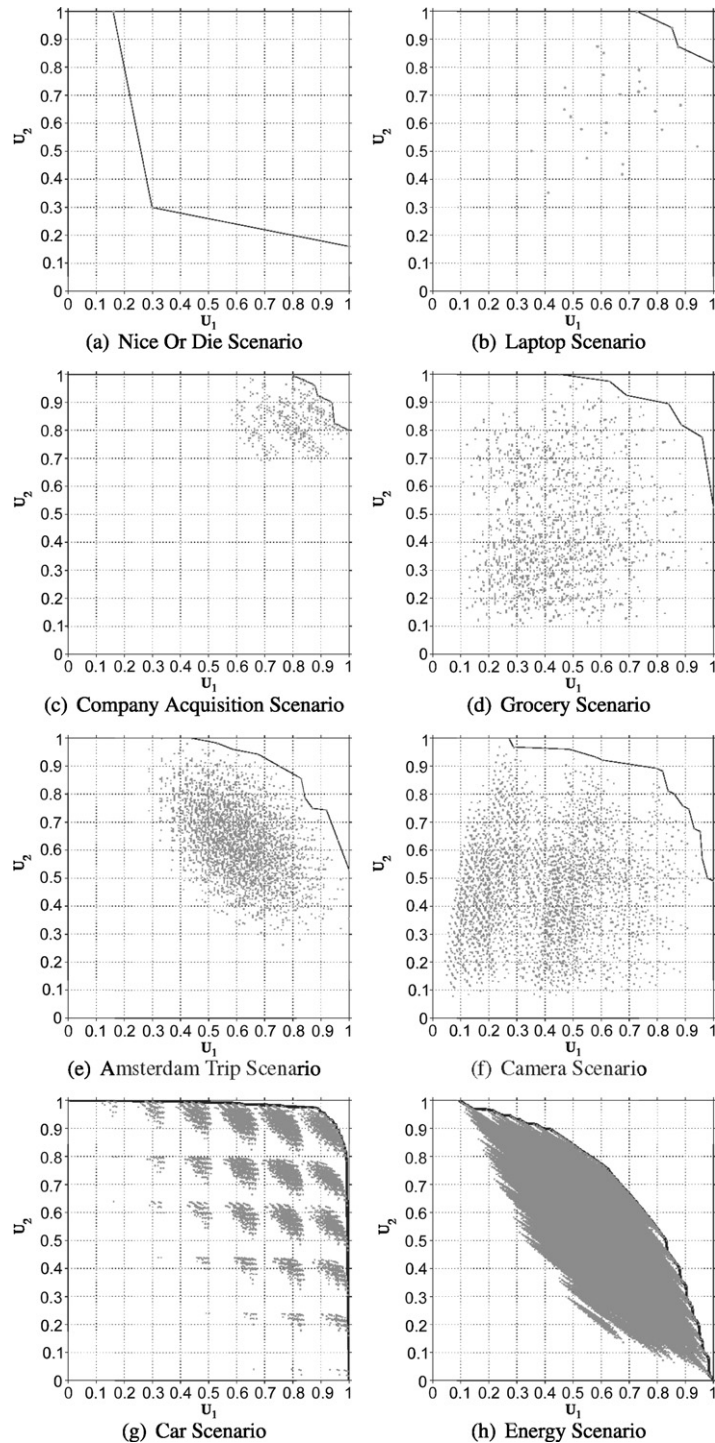
Fig. 1. Outcome spaces of all 8 scenarios for final round. The points represent all of the outcomes that are possible in each scenario. The solid line is the Pareto frontier, which connects all of the Pareto efficient outcomes.

Nice or Die. This scenario is the smallest used in the competition, with agents having to select between only 3 possible agreement points: a fair division point (nice), which is less efficient (in the sense that the sum of the agent's utilities is smaller) or one of two selfish points (die). The scenario is symmetric, in that neither player has an advantage over the other. The fair division point allows each player to achieve the same, relatively low score, while the other two selfish points allow one agent to get a high utility while its opponent achieves a very low one. As a result, the scenario has strong opposition between the participants. This means that if both agents try to get high utilities, it is hard for them to reach agreements.

However, if agents would like to make an agreement in this scenario, the social welfare is small (as the agents cannot learn from previous interactions with an opponent).

Laptop. In this scenario, a seller and a buyer, negotiate over the specifications of a laptop. There are three issues: the laptop brand, the size of the hard disk, and the size of the external monitor. Each issue has only three options, making 27 possible outcomes. For example, in a negotiation about buying a laptop the buyer may prefer to have a middle-sized screen but the seller may prefer to sell laptops with small screens because s/he has more of those in stock. They could, however, agree on the brand of laptop that they want to buy/sell. A successful outcome of a negotiation reconciles such differences and results in a purchase.

Company Acquisition. This scenario represents a negotiation between two companies, in which the management of Intelligent Solutions Inc. (IS) wants to acquire the BI-Tech company. The negotiation includes five issues: the price that IS pays for BI Tech, the transfer of intellectual property, the stocks given to the BI-Tech founders, the terms of the employees' contracts and the legal liability of Intelligent Solutions Inc. Each company wants to be the owner of the intellectual property. For IS, this issue is much more important. IS and BI-Tech have common interest that the BI-Tech co-founders would get jobs in IS. IS prefers to give BI-Tech only 2% of the stocks, while the BI-Tech co-founders want 5%. IS prefer private contracts, while firing workers is less desirable by them. BI-Tech prefers a 15% salary raise. For both sides this is not the most important issue in the negotiation. Each side prefers the least legal liability possible. In this case, the utility range is narrow and has high utility values such that all outcomes give both participants a utility of at least 0.5. The scenario is relatively small, with 384 possible outcomes.

Grocery. This scenario models a discussion in a local supermarket. The negotiation is between two people living together who have different tastes. The discussion is about five types of product: bread, fruit, snacks, spreads, and vegetables. Each category consists of four or five products, resulting in a medium sized scenario with 1600 possible outcomes. For their daily routine it is essential that a product of each type is present in their final selection, however only one product can be selected for each type. Besides their difference in taste, they also differ in what category of product they find more important. The profiles for agents Mary and Sam are modelled in such a way that a good outcome is achievable for both. Sam has a slight advantage, since he is easier to satisfy than Mary, and therefore is likely to have better outcomes. This scenario allows outcomes that are mutually beneficial, but the outcome space is scattered so agents must explore it considerably to find the jointly profitable ones.

Amsterdam Trip. This scenario concerns the planning of a tourist trip to Amsterdam and includes issues representing the day and time of travel, the duration of the trip, the type of venues to be visited, the means of transportation and the souvenirs to buy. This scenario is moderately large as the utility space has 3024 possible bid configurations. The utility functions specify a generous win-win scenario, since it would be unrealistic for two friends to make a trip to Amsterdam and to have it be a zero-sum game. The size of the scenario enables the agent to communicate their preferences (by means of generating bids), without having to concede far. The size also puts agents which use a random method of generating bids at a disadvantage, since the odds of randomly selecting a Pareto optimal bid in a large scenario are small. So this scenario will give an advantage to agents that make some attempt to learn the opponents' utility function, and those capable of rapidly choosing offers.

Camera. This scenario is another retail based one, which represents the negotiation between a buyer and a seller of a camera. It has six issues: maker, body, lens, tripod, bags, and accessories. The size of this scenario is 3600 outcomes. The seller gives priority to the maker, and the buyer gives priority to the lens. The opposition in this negotiation scenario is medium. The range of the contract space is wide, which means the agents need to explore it to find the jointly profitable outcomes. While jointly profitable outcomes are possible (since the Pareto frontier is concave) [34], no party has an undue advantage in this (since the Nash point is at an impartial position).

Car. This scenario represents a situation in which a car dealer negotiates with a potential buyer. There are 6 negotiation issues, which represent the features of the car (such as CD player, extra speakers and air conditioning) and each issue takes one of 5 values (good, fairly good, standard, meagre, none), creating 15,625 possible agreements. Although the best bids of the scenario are worth zero for the opponent, this scenario is far from a zero sum game. For example, agents can make agreements in which one of them can get close to the maximum possible utility, if it persuades its opponent to accept a utility only slightly below this. The scenario also allows agents to compromise to a fair division point in which both agents achieve a utility very close to the maximum possible. Consequently, the scenario has very weak opposition between the two participants.

Energy. This scenario considers the problem faced by many electricity companies to reduce electricity consumption during peak times, which requires costly resources to be available and puts a high pressure on local electricity grids. The application scenario is modelled as follows. One agent represents the electricity distribution company whilst the other represents a large consumer. The issues they are negotiating over represent how much the consumer is willing to reduce its consumption

Table 2
Team members in the final round.

Agent name	Affiliation	Team members	Scenario name
AgentK2	Nagoya Institute of Technology	Shogo Kawaguchi Katsuhide Fujita Takayuki Ito	Camera
BRAMAgent	Ben-Gurion University	Radmila Fishel Maya Bercovitch Ayelet Urieli Betty Sayag	Company Acquisition
Gahboninho	Bar Ilan University	Mai Ben Adar Nadav Sofy Avshalom Elimelech	Nice Or Die
HardHeaded	TU Delft	Thijs van Krimpen Daphne Looije Siamak Hajizadeh	Amsterdam Trip
IAMhaggler2011	University of Southampton	Colin R. Williams Valentin Robu	Energy
Nice Tit-For-Tat Agent	TU Delft	Tim Baarslag Koen Hindriks Catholijn Jonker	Laptop
TheNegotiator	TU Delft	Alex Dirkzwager Mark Hendrikx Julian de Ruiter	Grocery
ValueModelAgent	Bar Ilan University	Asaf Frieder Dror Sholomon Gal Miller	Car

over a number of time slots for a day-ahead market (the 24 hours in a day are discretised into 3 hourly time slots). For each issue, there is a demand reduction level possible from zero up to a maximum possible (specifically, 100 kW). In this scenario, the distributor obtains utility by encouraging consumers to reduce their consumptions. Participants set their energy consumption (in kWh) for each of 8 time slots. In each slot, they can reduce their consumption by 0, 25, 50, 75 or 100 kWh. This scenario is the largest in the 2011 competition (390,625 possible agreements) and has highly opposing utility functions, therefore, reaching mutually beneficial agreements requires extensive exploration of the outcome space by the negotiating agents.

3.2. Agent descriptions

The competition started with 18 agents, with the best 8 of these qualifying for the final (see Table 2). The competition rules allowed multiple entries from a single institution but required each agent to be developed independently. Furthermore it was prohibited to design an agent which benefits some other specific agent (cf. the work on collusion in the Iterated Prisoner's Dilemma competitions in 2004 and 2005 [36]). In the rest of this section, we provide, in alphabetical order, brief descriptions of the individual strategies of the finalists based on descriptions of the strategies provided by the teams.

AgentK2

This agent is identical to *Agent K* [20], winner of the ANAC 2010 competition. When creating a counter offer *Agent K* calculates a target utility U_t based on the previous offers made by the opponent and the time that is still remaining in the negotiation. *Agent K* then makes random bids above the target utility. If no such bid can be found, the target utility is lowered to allow for more offers. The target utility U_t at time t is calculated using the following formula:

$$U_t = 1 - \left(1 - E_{\max}(t)\right) \cdot t^{\alpha}, \tag{5}$$

where $E_{\max}(t)$ is the estimated maximum value the opponent will present in the future based on the average and variance of previous bids, and α is a parameter which controls the concession speed.

Agent K uses quite a sophisticated acceptance mechanism, where it will use the average and variations of the previous bid utilities presented by the opponent to determine the best possible bid it can expect in the future. It will either accept or reject the offer based on the probability that the opponent will present a better offer in the future. If it has already received an offer from the opponent with the same utility or higher, it will offer that bid instead.

BRAMAgent

This agent uses opponent modelling in an attempt to propose offers which are likely to be accepted by the opponent. Specifically, its model of the opponent stores the frequency with which each value of each issue is proposed. This information is maintained only over the 10 most recent offers received from the opponent. Therefore, the first 10 offers *BRAMAgent* makes will be its preferred bid (the one which maximises its utility), while it gathers initial data for its opponent model.

It also uses a time-dependent concession approach, which sets a threshold at a given time. In each turn, *BRAMAgent* tries to create a bid that contains as many of the opponent's preferred values as possible (according to its opponent model), with

a utility greater than or equal to the current threshold. If *BRAMAgent* fails to create such a bid, a bid will be selected from a list of bids that was created at the beginning of the session. This list contains all of the possible bids in the scenario (or all the bids it managed to create in 2 seconds), sorted in descending order according to the utility values. *BRAMAgent* chooses randomly a bid that is nearby the previous bid that was made from that list.

BRAMAgent will accept any offer with utility greater than its threshold. The threshold, which affects both acceptance and proposal levels, varies according to time. Specifically, the threshold levels are set as pre-defined, fixed percentages of the maximum utility that can be achieved (0–60 seconds: 93% of the maximum utility, 60–150 seconds: 85%, 150–175 seconds: 70%, 175–180 seconds: 20%).

Gahboninho

This agent uses a meta-learning strategy that first tries to determine whether the opponent is trying to learn from its own concessions, and then exploits this behaviour. Thus, during the first few bids, *Gahboninho* steadily concedes to a utility of 0.9 in an attempt to determine whether or not the opponent is trying to profile the agent. At the same time, the agent tries to assert selfishness and evaluate whether or not the opponent is cooperative. The degree of the opponent's selfishness is estimated based on the opponent's proposals. Then, the more the opponent concedes, the more competitive *Gahboninho*'s strategy becomes. The opponent's willingness to concede is estimated based on the size of variance of the opponent's proposals. After this phase, if the opponent is deemed concessive or adaptive, the agent takes a selfish approach, giving up almost no utility. However, if the opponent asserts even more hard-headedness, it adapts itself to minimise losses, otherwise it risks breakdown in the negotiation (which has very low utility for both parties). In generating the bids, the agent calculates its target, U_t at time t as follows:

$$U_t = U_{\max} - (U_{\max} - U_{\min}) \cdot t \tag{6}$$

where $U_{\max}$ and $U_{\min}$ are the maximum and minimum utilities (respectively) in the opponent's bidding history. $U_{\max}$ depends on the opponent's selfishness and the discount factor. Unlike many of the other agents, rather than using a model of the opponent to determine the offer to propose at a given utility level, *Gahboninho* uses a random search approach. Specifically, the agent proposes a random offer above the target utility $T(t)$. The benefit of this approach is that it is fast, therefore, given the format of the competition, a very large number of offers can be exchanged, allowing greater search of the outcome space. Moreover, the agent suggests using the opponent's best bid if the time is almost up.

HardHeaded

In each negotiation round, *HardHeaded* considers a set of bids within a pre-defined utility range which is adjusted over time by a pre-specified, monotonically decreasing function. A model of the opponent's utility function is constructed by analysing the frequency of the values of the issues in every bid received from the opponent. From a set of bids with approximately equal utility for the agent itself, the opponent model is used to suggest bids that are best to the opponent in order to increase chances of reaching an agreement in a shorter period of time.

The concession function specifies an increasing rate of concession (i.e. decreasing utility) for the utility of the agent's bids. The function has non-monotonic curvature with one inflection point, determined by the discount factor of the scenario. This function is determined by tuning the strategy based on the sample scenarios and data made available before the competition. For the scenarios with time discounting, the time-line is split into two phases over which the agent practices different strategies: it starts by using a *Boulware* strategy, and after a certain amount of time has passed (depending on the discount factor), it switches to a *Conceder* strategy [10].

IAMhaggler2011

This agent uses a Gaussian process regression technique to predict the opponent's behaviour [46]. It then uses this estimate, along with the uncertainty values provided by the Gaussian process, in order to optimally choose its concession strategy. In so doing, the concession strategy considers both the opponent's behaviour and the time constraints.

The concession strategy is then used to determine the target utility at a given time. In the concession strategy, the agent finds the time, t^*, at which the expected discounted utility of the opponent's offer is maximised. In addition, it finds the utility level, u^*, at which the expected discounted utility of our offer is maximised. The agent then concedes towards $[t^*, u^*]$, whilst regularly repeating the Gaussian process and maximisations.

Finally, having chosen a target, the agent proposes an offer which has a utility close to that target. In choosing the bids, *IAMhaggler2011* uses an approach similar to that of *Gahboninho*. Specifically, a random package, with utility close to the target is selected according to the concession strategy. This strategy is a fast process, which allows many offers to be made and encourages the exploration of outcome space.

Nice Tit-for-Tat Agent

This agent plays a tit-for-tat strategy with respect to its own utility. The agent will initially cooperate, then respond in kind to the opponent's previous action, while aiming for the Nash point in the scenario. If the opponent's bid improves its utility, then the agent concedes accordingly. The agent is nice in the sense that it does not retaliate. Therefore, when the opponent makes an offer which reduces the agent's utility, the *Nice Tit-for-Tat Agent* assumes the opponent made a mistake and does nothing, waiting for a better bid. This approach is based on [14]. *Nice Tit-for-Tat Agent* maintains a Bayesian model

[15] of its opponent, updated after each move by the opponent. This model is used to try to identify Pareto optimal bids in order to be able to respond to a concession by the opponent with a nice move. The agent will try to mirror the opponent's concession in accordance with its own utility function.

The agent detects very cooperative scenarios to aim for slightly more than Nash utility. Also, if the domain is large, if the discount factor is high, or if time is running out, the agent will make larger concessions towards its bid target. The agent tries to optimise the opponent's utility by making a number of different bids with approximately this bid target utility.

TheNegotiator

Unlike the other finalist agents, this agent does not model the opponent. Its behaviour depends on the mode it is using, which can be either: Discount or NoDiscount. A negotiation starts with the agent using its NoDiscount mode, which results in hardheaded behaviour. After a predetermined time period, the agent switches to its Discount mode, in which its behaviour becomes more concessive.

The main difference between the different modes is in the speed of descent of the minimum threshold for acceptance and offering. In the NoDiscount mode, most time is spent on the higher range of utilities and only in the last seconds are the remaining bids visited. The Discount mode treats all bids equally and tries to visit them all. An opponent's offer is accepted if it is above the current minimum threshold. An offer should also satisfy the minimum threshold, however a dynamic upper-bound is used to limit the available bids to offer in a turn. In 30% of the cases this upper-bound is ignored to revisit old bids, which can result in acceptance in later phases of the negotiation.

Finally, *TheNegotiator* attempts to estimate the number of remaining moves to ensure that it always accepts before the negotiation deadline.

ValueModelAgent

This agent uses temporal difference reinforcement learning to predict the opponent's utility function. The particular learning technique is focused on finding the amount of utility lost by the opponent for each value. However, as the bid (expected) utilities represent the decrease in all issues, a method is needed to decide which values should change the most. To achieve this, the agent uses estimations of standard deviation and reliability of a value to decide how to make the split. The reliability is also used to decide the learning factor of the individual learning. The agent uses a symmetric lower-bound to approximate the opponent's concession (if the opponent makes 100 different bids, and the 100th bid is worth 0.94, it is assumed the opponent conceded at least 6%). These parameters were determined in advance, based on average performance across a set of scenarios available for testing before the competition.

In more detail, *ValueModelAgent* starts by making bids which lie in the top 2% of the outcome space. It severely limits the concession in the first 80% of the time-line. If there is a large discount, the agent compromises only as much as its prediction of the opponent's compromise. If there is no discount, the agent does not concede as long as the opponent is compromising. If the opponent stops moving, the agent compromises up to two thirds of the opponent's approximated compromise. As the deadline approaches (80–90% of the time has elapsed), the agent compromises up to 50% of the difference, providing that the opponent is still not compromising. Once 90% of the time has elapsed, the agent sleeps and makes the "final offer", if the opponent returns offers the agent sends the best offer that has been received from the opponent (accepting his last offer only if its close enough). *ValueModelAgent* has a fixed lower limit on its acceptance threshold, of 0.7. Therefore it never accepts an offer with an undiscounted utility lower than this value.

4. Competition results

The 2011 competition consisted of two rounds: qualifying and final. We describe the results of these rounds in turn.

4.1. Qualifying round

The qualifying round consisted of the 18 agents that were submitted to the competition. For each pair of agents, under each utility function, we ran a total of 3 negotiations. By averaging over all the scores achieved by each agent (against all opponents and using all utility functions), eight finalists were selected based on their average scores. Note that these averages are taken over all negotiations, excluding those in which both agents use the same strategy (i.e. excluding self-play). Therefore, the average score $U_{\Omega}(p)$ of agent p in scenario Ω is given formally by:

$$U_{\Omega}(p) = \frac{\sum_{p' \in P, p \neq p'} U_{\Omega}(p, p')}{(|P| - 1)} \tag{7}$$

where P is the set of players and $U_{\Omega}(p, p')$ is the utility achieved by player p against player p' in scenario Ω.

It is notable that *Gahboninho* was the clear winner of the qualifying round (see Table 3). As we will discuss later (in Section 6), we believe its strong performance is partly due to the learning approach it adopts, in an attempt to determine whether the opponent is cooperative.

Table 3
Average scores of every strategy in the qualifying round.

Rank	Agent strategy	Mean utility
1	Gahboninho	0.756
2	HardHeaded	0.708
3	ValueModelAgent	0.706
4	AgentK2	0.702
5	IAMhaggler2011	0.701
6	BRAMAgent	0.690
7	Nice Tit-For-Tat Agent	0.686
8	TheNegotiator	0.685
9	GYRL	0.678
10	WinnerAgent	0.671
11	Chameleon	0.664
12	SimpleAgentNew	0.648
13	LYYAgent	0.640
14	MrFriendly	0.631
15	AgentSmith	0.625
16	IAMcrazyHaggler	0.623
17	DNAgent	0.601
18	ShAgent	0.571

Table 4
Tournament results in the final round.

Rank	Agent strategy	Utility		95% confidence interval	
		Mean	Standard deviation	Lower bound	Upper bound
1	HardHeaded	0.749	0.0096	0.745	0.752
2	Gahboninho	0.740	0.0052	0.738	0.742
3	IAMhaggler2011	0.686	0.0047	0.685	0.688
4[*]	AgentK2	0.681	0.0047	0.679	0.683
5[*]	TheNegotiator	0.680	0.0043	0.679	0.682
6[*]	BRAMAgent	0.680	0.0050	0.678	0.682
7[*]	Nice Tit-For-Tat Agent	0.678	0.0076	0.675	0.681
8	ValueModelAgent	0.617	0.0069	0.614	0.619

[*] Note that the competition ranking that was released immediately following the final round listed BRAMAgent as achieving 4th place, but in our extended runs, *BRAMAgent* achieves a lower ranking. This is due to the small differences of the scores of the agents in positions 4 to 7. Specifically, there is no statistically significant difference between the utilities achieved by *AgentK2*, *TheNegotiator*, *BRAMAgent* and the *Nice Tit-For-Tat Agent*.

Table 5
Detailed scores of every agent in each scenario in the final round. Bold text is used to emphasise the best score achieved in each scenario.

Agent	Nice or Die	Laptop	Company Acquisition	Grocery	Amsterdam Trip	Camera	Car	Energy
HardHeaded	**0.571**	0.669	0.749	0.724	0.870	**0.811**	**0.958**	0.637
Gahboninho	0.546	0.730	0.752	0.668	**0.929**	0.665	0.946	**0.682**
IAMhaggler2011	0.300	**0.750**	**0.813**	0.726	0.781	0.715	0.864	0.543
AgentK2	0.429	0.655	0.788	0.717	0.750	0.727	0.921	0.459
TheNegotiator	0.320	0.651	0.757	0.733	0.791	0.742	0.930	0.519
BRAMAgent	**0.571**	0.631	0.747	0.725	0.792	0.739	0.803	0.432
Nice Tit-For-Tat Agent	0.425	0.668	0.772	0.753	0.739	0.774	0.786	0.509
ValueModelAgent	0.137	0.641	0.764	**0.765**	0.857	0.781	0.951	0.037

4.2. Final round

For the final round, we matched each pair of finalists, under each utility function, a total of 30 times. Table 4 summarises the means, standard deviations, and 95% confidence interval bounds for the results of each agent, taken over the 30 iterations.[8] In common with the approach used in the qualifying round, all agents use both of the profiles that are linked to a scenario. The average score achieved by each agent in each scenario is given in Table 5, and presented visually in Fig. 2. In

---

[8] The standard deviations and confidence intervals are calculated based on the variance of the utilities across the 30 iterations of the tournament (after being averaged over all of the scenarios). Therefore they only measure the variance across complete tournaments, which may be due to intentional randomness within the agents' strategies or stochastic effects that are present in the tournament setup.
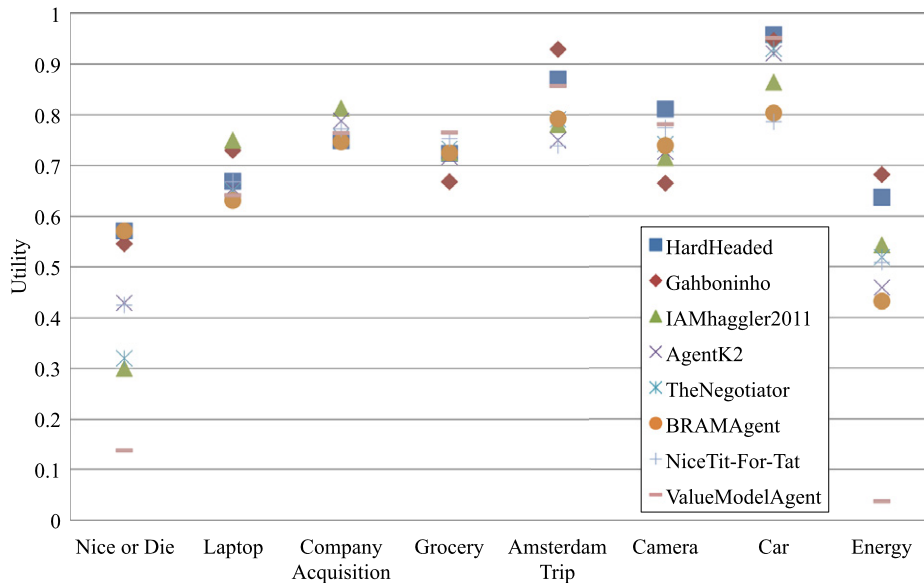
Fig. 2. Scores of every agent in each scenario in the final round.

the finals, *HardHeaded* proved to be the clear winner,[9] with a score of 0.749. Fig. 2 clearly shows that in most scenarios the margin between the worst and the best agents was minimal. Specifically, in 6 of the 8 scenarios, the worst agent achieved no less than 80% of the best agent's score. The remaining two scenarios that had a much greater range of results were also the scenarios with the greatest opposition between the two utility functions. The results presented in Table 5 and Fig. 2 also show that the winning *HardHeaded* agent did not win in the majority of the scenarios (it only did so in 3 out of 8). In some scenarios, especially those with a large size and high opposition, some of the agents performed poorly. However, *HardHeaded* and *Gahboninho* got the higher utility in such "tough" scenarios. Consequently, *HardHeaded* and *Gahboninho* won by a big margin in most scenarios. In addition, *IAMhaggler2011* won the *Company Acquisition* and *Laptop* scenarios where there is a low discount factor, therefore, *IAMhaggler2011* is well suited to cases where agreements need to be reached quickly. The differences among *BRAMAgent*, *AgentK2*, *TheNegotiator* and the *Nice Tit-For-Tat Agent* are very small, in fact they are statistically indistinguishable.

5. Evaluation methods

The competition results only give a fairly narrow view of the performance of the agents. In order to consider their behaviour in more detail, and also to explore the interactions between different strategies in a tournament setting, we use two further evaluation methods in this paper. As will become evident, the performance of an agent depends heavily on the opponent. Therefore, Section 5.1 focuses on the influence of the opponent on the negotiation outcome. Now, apart from individual performance, the tournament setting of ANAC demands that agents take into account their relative performance, i.e., that they are robust in the sense of not yielding too much to the other contestants. To this end, Section 5.2, explores the influence of the tournament pool by considering different mixes of opponent strategies.

5.1. Influence of the opponent

When looking at Table 4, it is natural to ask which agent characteristics were decisive factors in the final ranking of the agents. Which agents behaved very competitively and which ones were more cooperative? Were they successful because of it, and if so, against whom? In order to answer such questions, we characterise the ANAC agents by analysing their bidding behaviour and utility gain against different types of opponents. In particular, we focus on the amount they are willing to concede to the opponent, as this is a key determinant of an agent's bidding behaviour.

In more detail, making concessions is key to a successful negotiation: without them, negotiations would not exist [24]. Many of the classic negotiation strategies are characterised by the way they make concessions throughout the negotiation. For example, the time dependent tactics such as *Boulware* and *Conceder* [7] are characterised by the fact that they steadily

---

[9] There are a number of reasons why the winner in the final round was different to the qualifying round. Firstly, the set of scenarios used in the final was smaller than in the qualifying round, and it is possible that the final scenarios were more favourable to the *HardHeaded* agent. Secondly, the set of participating agents was smaller, and furthermore, due to the elimination of the lower scoring agents, those agents that remain were more competitive. Finally, it is possible that the agents were modified between the two rounds.
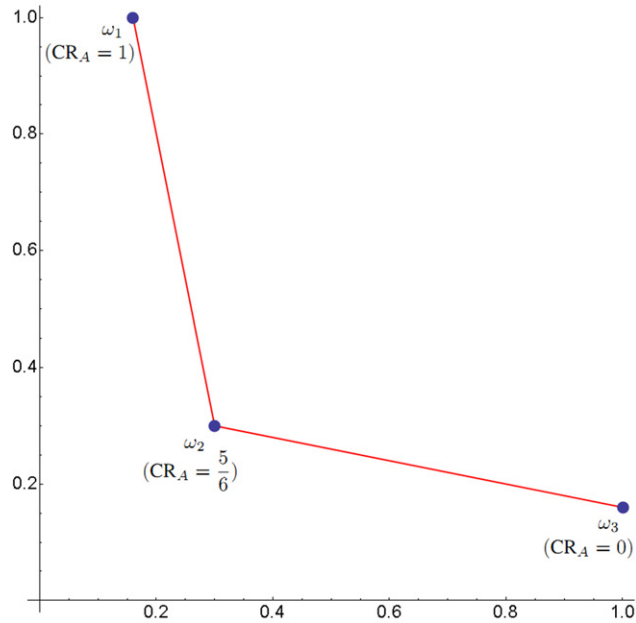
Fig. 3. The three possible outcomes of the *Nice or Die* scenario, with the utility of agent *A* plotted on the horizontal axis.

concede throughout the negotiation process. Concessions made by such tactics depend on various factors, such as the opening bid, the reservation value, and the remaining time. Where possible, it can be beneficial to use prior information about the market conditions in selecting a suitable concession strategy. However, in the ANAC setup, there is no such prior information available to the agents. Alternatively, behaviour dependent tactics (e.g. Tit for Tat) [4,7] base their decision to make concessions on the concessions of the other negotiating party.

Now, deciding what concessions to make depends in large part on the opponent. One can either choose to signal a position of firmness and stick to an offer. Alternatively, one can take a more cooperative stance, and choose to make a concession. Such a concession may, in turn, be reciprocated by the opponent, leading to a progression of concessions. On the other hand, the negotiation process can easily be frustrated if the opponent adopts a take-it-or-leave-it approach. Against this background, we study the ANAC negotiation strategies according to the way they concede towards different types of opponents.

5.1.1. Concession rate

We use the notion of *concession rate* (CR) to quantify the total amount an agent has conceded towards the opponent during a negotiation.[10] We will define the concession rate of an agent *A* as a normalised measure $CR_A \in [0, 1]$ with the following meaning: if $CR_A = 0$, then *A* did not concede at all during the negotiation, while $CR_A = 1$ means that player *A* yielded completely to the opponent by offering the opponent its best offer.

In doing so, it is generally not enough to simply consider the utility of the agreement as a measure for the concession rate. For instance, a negotiator may not get an agreement before the deadline. In that case, both parties receive zero utility, but this gives no information about the concessions that were made. Also, the last offer made by a negotiator is not necessarily the offer to which he was willing to concede the most. To capture the notion of concession rate, we define it in terms of the *minimum* utility *m* a negotiator has demanded during the negotiation, as this is a measure of the total amount the negotiator was willing to concede during a negotiation.

We illustrate the concept of CR by considering the *Nice or Die* scenario as described in Section 3.1, which has only three possible outcomes: $\omega_1 = \langle 0.16; 1 \rangle$, $\omega_2 = \langle 0.3; 0.3 \rangle$, and $\omega_3 = \langle 1; 0.16 \rangle$ (see Fig. 3).

Let us first consider the case where agent *A* sticks to the same offer ω_3 throughout the negotiation, demanding the highest possible utility for itself. In this case, the minimum utility *m* that player *A* has demanded is equal to 1, so *A* has not conceded anything, therefore the corresponding concession rate of agent *A* should be equal to zero:

$$CR_A = 0, \quad \text{for } m = 1. \tag{8}$$

---

[10] This technique was previously used in [5] to analyse the ANAC 2010 competition. However, in [5], a more comprehensive notion of concession rate is introduced, where the concession rate is measured at *arbitrary* time points, rather than an aggregate of the full negotiation. We adhere to the same naming convention, although in order to express an agent's concessive behaviour as a single number, we have employed a simplified version here that does not capture the temporal aspect.

On the other hand, when A concedes all the way to the opponent's best option (which is ω_1),[11] CR_A should be equal to 1. Note however, that conceding all the way does not necessarily mean demanding zero utility, or the lowest possible utility. For instance in this example, A would still receive 0.16 utility when bidding ω_1 (see Fig. 3). In general, there may also be bids with even lower utility for A that are, for example, also bad for its opponent. However, player A should be able to always obtain at least an agreement that is the best outcome for the opponent, as any rational player B will accept it. We shall refer to this utility as the *full yield utility* (FYU_A) of player A. Intuitively, it is the worst score player A can expect to obtain in the negotiation. Player A's concession rate is therefore maximal if he makes a bid on or below the full yield utility:

$$CR_A = 1, \quad \text{for } m \leqslant FYU_A. \tag{9}$$

Between the two extremes, we define CR_A to decrease linearly from 1 to 0. There is only one function that satisfies this constraint, along with Eqs. (8) and (9) above, namely:

$$CR_A(m) = \begin{cases} 1 & \text{if } m \leqslant FYU_A, \\ \frac{1-m}{1-FYU_A} & \text{otherwise.} \end{cases}$$

By using normalisation, it is guaranteed that if $CR_A = 0$, then A has not conceded at all, while for $CR_A = 1$, player A has conceded fully (i.e., up to its full yield utility). Normalising has the added benefit of reducing bias in the scenarios: in a typical scenario with strong opposition such as *Energy*, players may obtain utilities anywhere between 0.1 and 1, while in scenarios with weak opposition such as *Company Acquisition*, utility ranges are much more narrow. Normalisation ensures that the concession rate can be compared over such different scenarios.

Note that CR is a measure of the bidding strategy, so in particular it does not take into account the conditions under which an agent accepts an offer.

Example 1. Suppose player A has made the following bids: $\langle \omega_3, \omega_3, \omega_2, \omega_3 \rangle$ (see Fig. 3). Then its minimum demanded utility m is equal to the utility of ω_2, which is 0.3. The full yield utility FYU_A is equal to the utility of ω_1, which is 0.16. Therefore,

$$CR_A = \frac{1 - 0.3}{1 - 0.16} = \frac{5}{6}. \qquad \square$$

When the negotiation thread is given, the minimum demanded utility m of a player A is known, and then we shall denote the concession rate simply by CR_A. We also omit the subscript A when it is clear from the context.

5.1.2. Classifying the agents according to their concession rates

In order to classify the eight finalists according to their concession rate, we considered a negotiation setup with the following characteristics. To be able to compare the results, we need to fix the set of opponents that are used to measure the CR. This raises the question which agents can be used for this purpose. To test both sides of the spectrum, we let the finalists negotiate against both a very cooperative and a very competitive opponent. The opponent tactics that we use to measure concession rates are simple, non-adaptive negotiation tactics. We do so because we want to ensure that the concession rate results depend as much as possible on the agent's own negotiating tactic, and not on the opponent's. To be more precise, we aim for three opponent characteristics when measuring the concession rate:

1. **Simplicity:** The concession rate results of an agent are less sensitive to the opponent, and hence easier to interpret, if the opponent negotiation tactic is simple and easy to understand.
2. **Regularity:** We want to give the agent sufficient time to show its bidding behaviour; therefore, the opponent should not end the negotiation prematurely by either reaching an agreement too fast or breaking off the negotiation. Another issue here is that the opponent should generate sufficient bids. This requires computationally efficient agents that respond within a reasonable amount of time and excludes extreme agents that only make a limited number of offers.
3. **Deterministic behaviour:** In order to reduce variance in experimental results, we prefer deterministic agents to those that demonstrate random bidding behaviour.

For the *competitive opponent*, we chose *Hardliner* (also known as take-it-or-leave-it, or *Hardball* [24]). This strategy simply makes a bid of maximum utility for itself and never concedes. This is the most simple competitive strategy that can be implemented and it fits the other two criteria as well: it is deterministic and it gives the agent the full negotiation time to make concessions.

For the *cooperative opponent*, we selected *Conceder Linear*, i.e., the time dependent tactic adapted from [7,10] with parameter $e = 1$. Depending on the current time $t \in [0, 1]$, this strategy makes a bid with utility closest to

---

[11] Note that the best outcome for a player is not necessarily unique, but typical scenarios (including those considered in ANAC and hence, in this paper) all have unique optimal outcomes for both players, so that the full yield utility is well-defined.

Table 6

An overview of the concession rate and standard deviation of every agent in the experiments against *Conceder Linear* and *Hardliner*.

Agent	Concession rate vs conceder linear		Concession rate vs hardliner	
	Mean	Standard deviation	Mean	Standard deviation
AgentK2	0.232	0.029	0.435	0.019
BRAMAgent	0.141	0.042	0.642	0.002
Gahboninho	0.219	0.041	0.907	0.032
HardHeaded	0.041	0.027	0.622	0.039
IAMhaggler2011	0.419	0.058	0.987	0.005
Nice Tit-for-Tat Agent	0.169	0.033	0.876	0.031
TheNegotiator	0.071	0.011	1.000	0.000
ValueModelAgent	0.239	0.013	0.667	0.030
Random Walker	0.848	0.040	0.904	0.030
Hardliner	0.000	0.002	0.000	0.002
Boulware	0.162	0.009	1.000	0.000
Conceder Linear	0.414	0.034	1.000	0.000
Conceder	0.571	0.043	1.000	0.000

$$P_{\min} + (P_{\max} - P_{\min}) \cdot \big(1 - F(t)\big), \tag{10}$$

with

$$F(t) = k + (1 - k) \cdot t^{1/e}.$$

In our experiments, we selected the standard value $k = 0$, and $P_{\max}$, $P_{\min}$ are respectively set to the maximum and minimum utility that can be obtained in the scenario. With these values, and setting $e = 1$, we obtain a very simple conceding tactic. It reduces Eq. (10) to

$$P_{\min} + (P_{\max} - P_{\min}) \cdot (1 - t),$$

so that it linearly reduces its demanded utility from $P_{\max}$ to $P_{\min}$ as time passes.

Both strategies accept if and only if their planned offer has already been proposed by the opponent in the previous round.

There exist even simpler conceding tactics such as *Random Walker*[12] (which generates random bids), or an agent that accepts immediately. However, both opponent strategies are not *regular* in the sense that they do not give the agent enough time to show its bidding behaviour. *Random Walker* has the additional disadvantage of not being deterministic. Therefore, we believe *Random Walker* can serve as a useful baseline strategy to test the efficacy of a negotiation strategy, but not as a useful opponent strategy to measure an agent's willingness to concede. Consequently, we selected *Conceder Linear* as the cooperative opponent, as it fulfils the three requirements listed above.

5.1.3. Experiments

For our experimental setup we employed GENIUS using the same negotiation parameters as described in Section 2.2. We included all the negotiation tactics of the eight finalists. In addition to the ANAC agents, we included some well-known agents to explore some extreme cases. First, we included the *Hardball* strategy described above, which consistently makes the maximum bid for itself. We also studied three members of the time-dependent-tactics family [7] as defined above, namely: *Boulware* ($e = 0.2$), *Conceder Linear* ($e = 1$), and *Conceder* ($e = 2$). Finally, we included the *Random Walker* strategy, which randomly jumps through the negotiation space. We shall refer to these five strategies as our *benchmark strategies*.

We measured the concession rate of an agent *A* playing against *Conceder Linear*, and *Hardliner*. The two parties attain a certain outcome, or reach the deadline. In both cases, at the end of the negotiation, *A* has reached a certain concession rate as defined in Section 5.1.1. The concession rate is then averaged over all trials on all eight final scenarios (see Section 3.1), alternating between the two utility functions defined as part of that scenario. For example, in the *Energy* scenario, *A* will play both as the electricity distribution company and as the consumer. We then repeated every negotiation 30 times to increase the statistical significance of our results. We have taken the averages and standard deviations over these 30 data points, where each data point is the average taken over all scenarios.

We present the results of the concession rate experiments in Table 6, and its graphical representation is depicted in Fig. 4. In Fig. 5 through to Fig. 8, we present the connection between the concession rate of the agents and their average negotiation results against *Conceder Linear* and *Hardliner*. In all figures, the error bars represent one standard deviation from the mean.

In more detail, we start with the observations regarding the clustering of different strategies in Fig. 4. In particular, the *Hardliner* strategy and the *Random Walker* strategy are at the opposite sides of the concession rate spectrum. *Hardliner* will

---

[12] The *Random Walker* strategy is also known as the *Zero Intelligence* strategy [13].
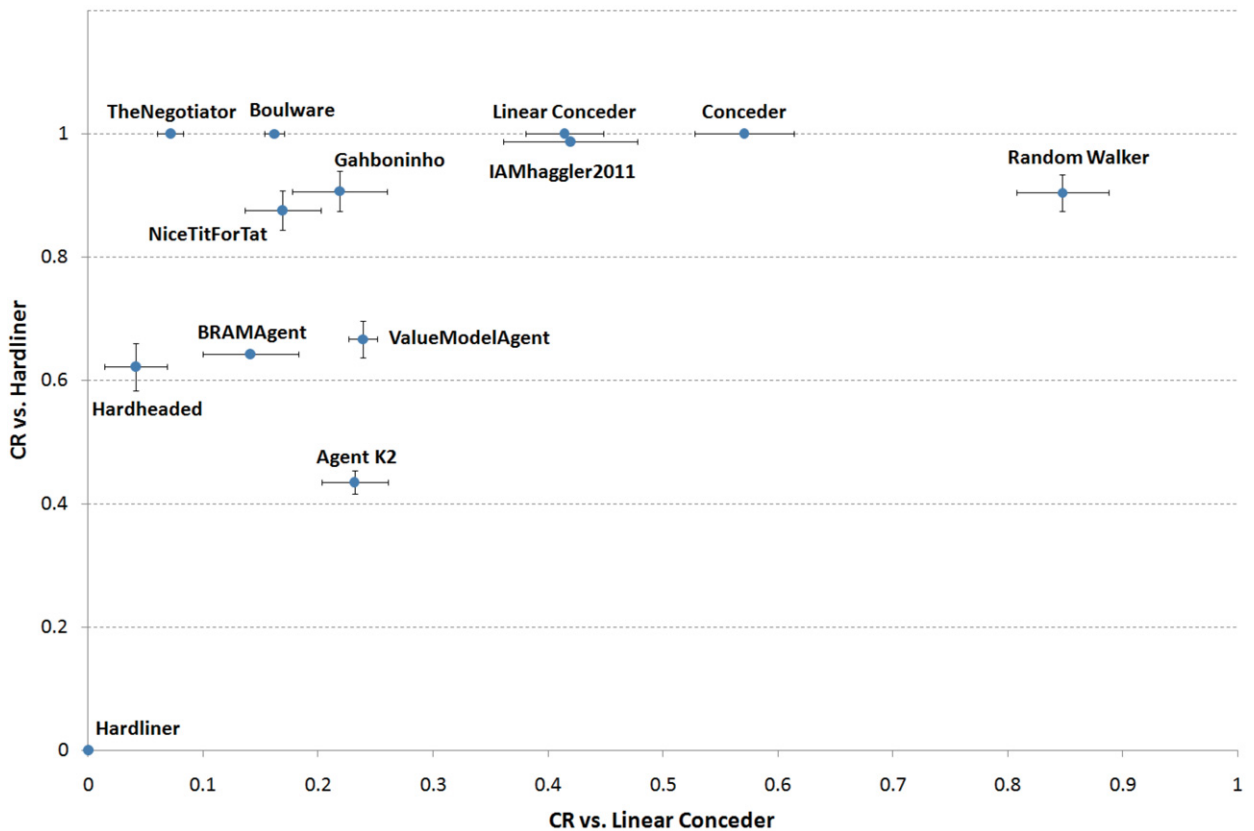
Fig. 4. A scatter plot of the concession rate and standard deviation of every agent in the experiments against *Conceder Linear* and *Hardliner*. The error bars represent one standard deviation from the mean.

not concede to any type of strategy, so by definition it has $CR = 0$ against both *Hardliner* and *Conceder Linear*. Consequently, Hardliner is the least conceding strategy possible.

On the other hand, *Random Walker* will make arbitrary concessions given enough time. This makes *Random Walker* one of the most concessive strategies possible. Against *Hardliner*, it is just a matter of time for *Random Walker* to randomly produce a bid with which it fully concedes, so *Random Walker* has $CR = 1$ against *Hardliner* when it manages to reach an agreement. In large domains, however, it may not be able to produce such a bid in time, and its score of 0.9 against *Hardliner* means that on average the deadline is reached after *Random Walker* has explored 90% of the bid space. Against *Conceder*, it may not have as much time to fully concede, but it generally will produce offers of very low utility in this case as well, resulting in a CR of 0.85.

We included three members of the time dependent tactics family in our experiments: *Boulware* ($e = 0.2$), *Conceder Linear* ($e = 1$), and *Conceder* ($e = 2$). They are all located in the top of the chart because they all share the same concession rate (equal to 1 versus Hardliner). This is to be expected, as any agent from the time dependent family will offer the reservation value when the deadline is reached [10], resulting in full concession to the opponent. In general, all time dependent tactics will lie on the line $CR = 1$ against *Hardliner*. In addition to the time dependent tactics, we also see *TheNegotiator* in the top of the chart. This strategy is very similar to a Boulware strategy as it adopts a very competitive time dependent strategy that makes larger concessions when the negotiation deadline approaches. All the strategies that lie on the line $CR = 1$ give in fully to Hardliner and are thus fully exploited by a strategy that does not give in at all. All of these four strategies have a very simple bidding strategy and are clearly not optimised to deal with minimally conceding strategies or strategies that do not concede at all.

Against *Conceder Linear*, the results are also intuitively clear: concessions of the agents get bigger when the parameter $e \in (0, \infty)$ gets bigger, so Boulware concedes the least, while *Conceder* concedes the most. More generally, when $e \rightarrow 0$, then $CR \rightarrow 0$. Conversely, when $e \rightarrow \infty$, then $CR \rightarrow 1$. An agent that has $CR = 1$ against both *Hardliner* and *Conceder Linear* is an agent that would jump to the opponent's best bid immediately.

Finally, all ANAC 2011 strategies reside in the left part of the chart, which means they play quite competitively against a conceding strategy. The top left of the chart is populated by the four 'nicer' strategies of ANAC: *TheNegotiator*, *Nice Tit for Tat*, *Gahboninho*, and *IAMhaggler2011*. The other four finalists: *HardHeaded*, *BRAMAgent*, *ValueModelAgent*, and *AgentK2* are more competitive. It is interesting to note that *HardHeaded*, the winner of ANAC 2011, concedes the least against a cooperating
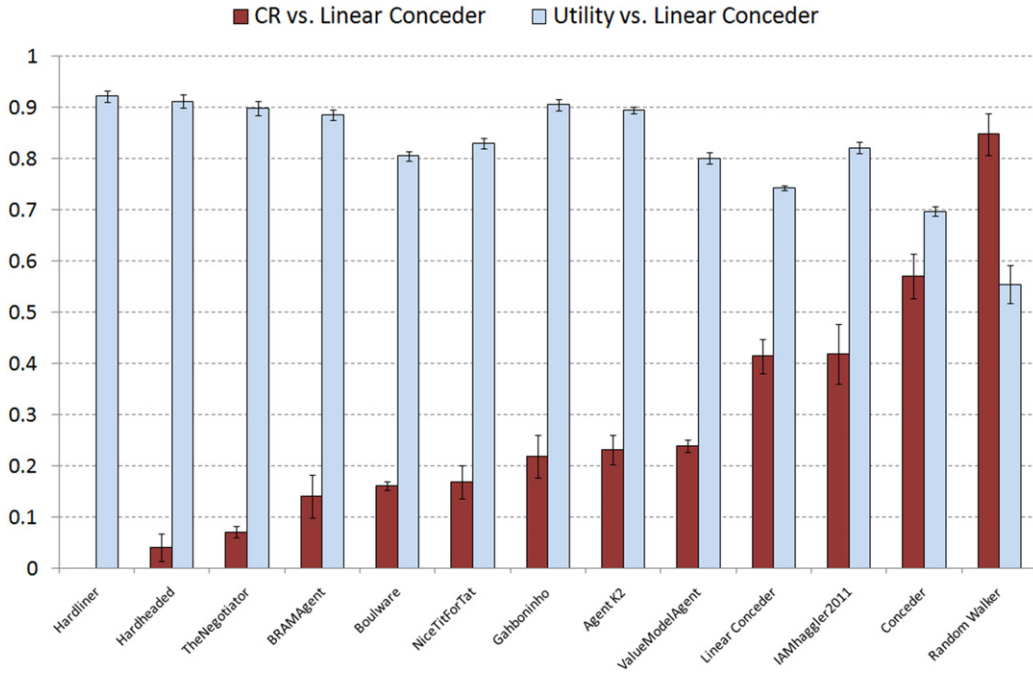
Fig. 5. Comparing the results versus the concession rate against the *Conceder Linear* opponent.
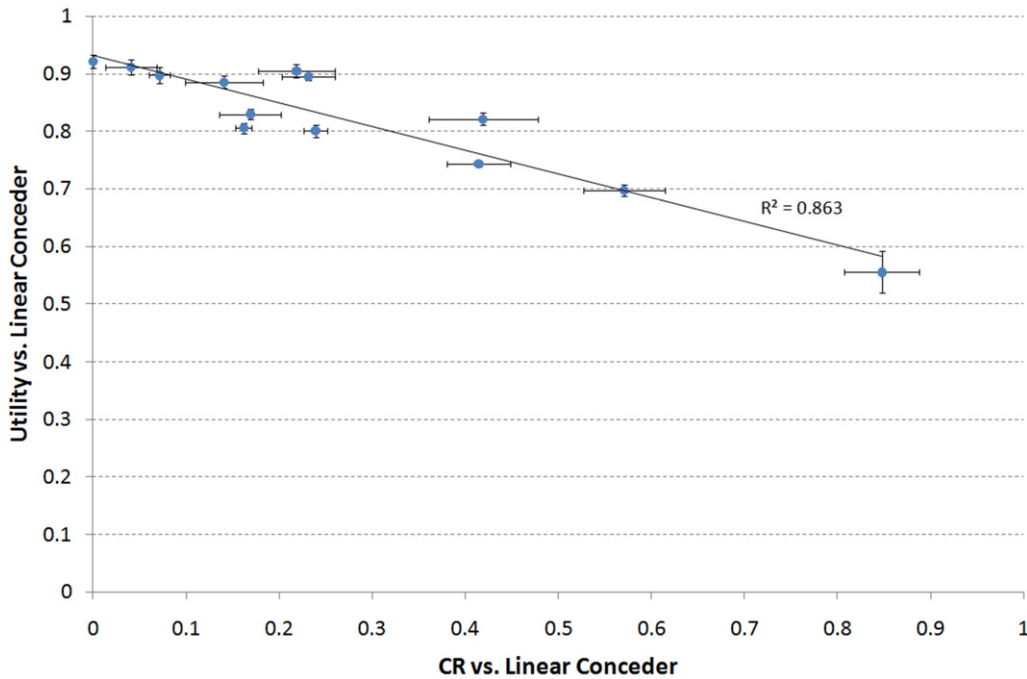

Fig. 6. The correlation between the concession rate and utility obtained against the *Conceder Linear* opponent of all ANAC and benchmark agents.

strategy. The fact that these strategies did very well in the competition seems to indicate that in order to be successful in such circumstances, an agent should behave competitively, especially against very cooperative strategies.

Indeed, it seems reasonable to conjecture that the nicer we play against a conceding opponent (meaning the higher our CR is against a conceding opponent), the less utility we make in the negotiation. To test this hypothesis, we selected *Conceder Linear* as the opponent, and then compared the CR of every agent against the average utility of the agreements being reached. The results are plotted in Fig. 5. It is readily observable that against this type of opponent, it pays to concede as little as possible. For example, *HardHeaded* has the lowest CR after *Hardliner*, which results in a very high score against
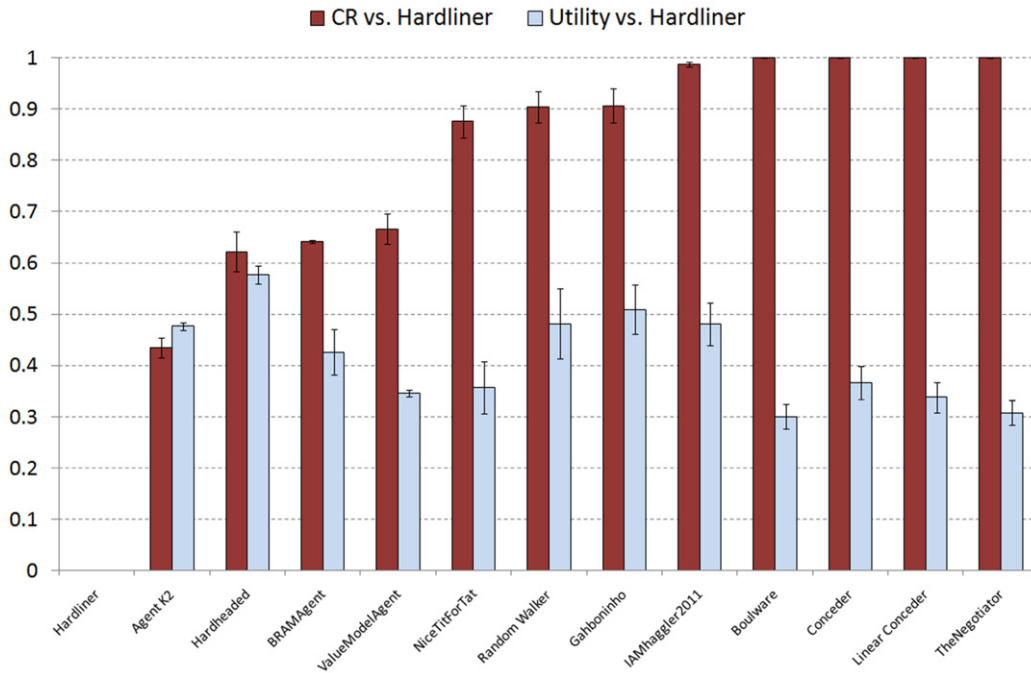
Fig. 7. Comparing the results versus the concession rate against the *Hardliner* opponent.
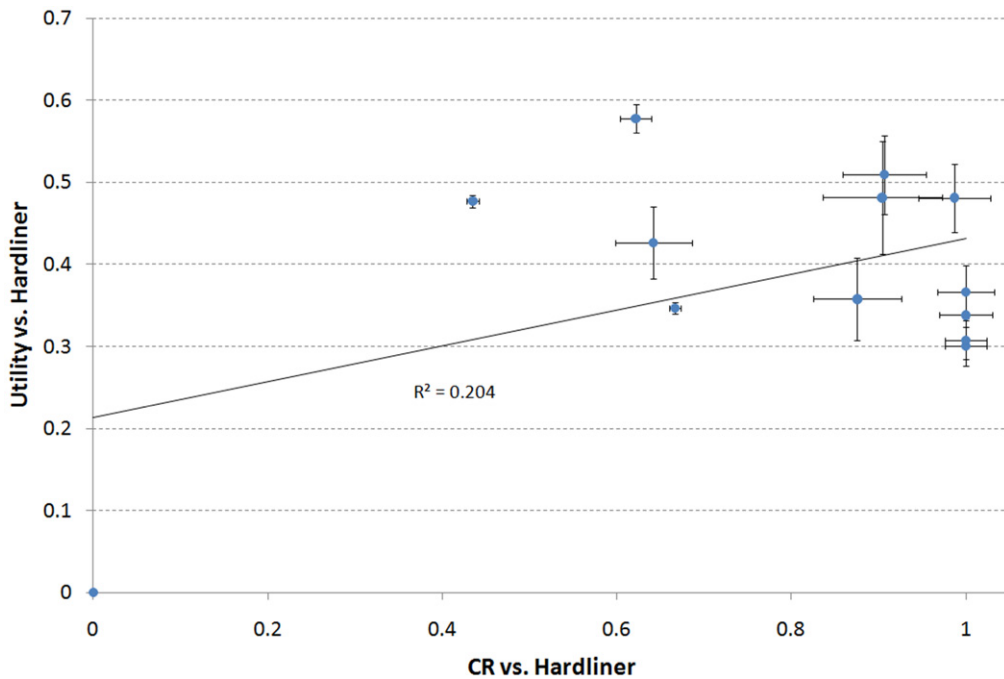

Fig. 8. The correlation between the concession rate and utility obtained against *Hardliner* of all ANAC and benchmark agents.

Conceder Linear. In general, the correlation is very high (coefficient of determination of $R^2 = 0.86$) between the CR and utility obtained against *Conceder Linear* (see Fig. 8). This means that CR is a good predictor of how an agent fares against a cooperative agent. As noted, the CR is a measure of the bidding strategy. Therefore, against such a cooperative opponent, the decisive factors in the success of an agent are the circumstances and timing under which the bidding strategy makes its concessions.

In contrast, we cannot make similar statements about how to play against a very competitive agent such as *Hardliner*. As can be seen in Figs. 5 and 6, there is a positive correlation between the CR and the average utility obtained when

playing *Hardliner*. This means that the nicer an agent plays, the more utility is obtained against *Hardliner*. This makes sense, as competitive play against *Hardliner* is sure to result in a break off. However, the correlation is very weak, with a coefficient of determination of $R^2 = 0.20$. Therefore in this case the CR is not a good predictor of the negotiation success. Because a break off of the negotiation is more likely against a very competitive opponent, the acceptance conditions of an agent play a much more important role than in the previous case.

Clearly, the only way to receive any utility in a negotiation against *Hardliner* is for an agent to give in fully. But in a tournament setting, an agent should not only take into account its own utility, but it should also make sure not to give away too much utility to the other contestants. Indeed, it is not entirely clear how one should successfully negotiate against the *Hardliner* (or similar strategies such as *HardHeaded*) without getting exploited in terms of tournament score. As a consequence, there exists a tension between the short-term negotiation incentives of a negotiating agent and its goal to get a high overall tournament ranking. Such considerations are related to the composition of the tournament pool, which we will cover more deeply in the next section.

5.2. Influence of the tournament pool

In this section, we begin by introducing the technique of empirical game theory as a method for analysing the results gathered from a series of negotiation encounters (Section 5.2.1). The technique was previously used successfully by [44] to provide insights into the strategies used in the Trading Agent Competition (TAC). Specifically, we use the approach in order to search for tournaments in which there is no incentive for any of the agents to change their strategy. Such tournaments are considered to be empirical equilibria. In more detail, Section 5.2.2 considers two player negotiations, in which each player has the freedom to choose any of the eight finalist strategies. Then, Section 5.2.3 extends the analysis to consider eight player negotiation tournaments, in which only the top three strategies are available to choose from. Finally, Section 5.2.4 removes this restriction, and therefore considers eight player tournaments in which all of the finalist strategies are available to choose from. We perform this for the average results achieved across the eight competition scenarios, and also for specific scenarios which have notable characteristics due to their size, discounting factor or mean distance to the Pareto frontier.

5.2.1. The empirical game theoretic analysis

As mentioned in Section 2, the method that is used to rank the agents which participated in the ANAC 2011 competition measures their average performance in a tournament setup, in which all agents negotiate against all other agents. Note that the goal of achieving the highest score in such a tournament is somewhat different to that of reaching the highest score in an individual negotiation. However, an agent can only control the outcomes of the negotiations it is involved in; it has no control over the negotiations between other agents. Thus, maximising the utility of each negotiation an agent participates in can be seen as a good approximation that an agent can take in order to maximise its tournament score.

The winner of the competition is the agent which obtains the highest score in the specific tournament where each agent uses a different strategy from a set of possible strategies. However, the scores achieved in this tournament, by themselves, do not reveal much about the *robustness* of the negotiation strategies submitted in different scenarios. For example, we may be interested to know how the winning strategy would change if the tournament size were chosen differently, and especially, if the mix of opponent strategies were different. Moreover, it is natural to ask whether the agents participating in the competition have an incentive to switch to a different strategy (from the set of strategies used in ANAC 2011) in order to improve their score.

In order to answer such questions, we perform an EGT analysis of the tournament results, using the techniques first developed by [44]. The aim of an EGT analysis is to search for pure Nash equilibria using a set of empirical results. It is not feasible to use a standard game theoretic approach (in which all possible strategies are considered) for analysing ANAC 2011 (or other similar competitions), since there can be an infinite number of strategies that an agent could take.[13] In contrast, EGT analysis uses the assumption that the strategy used by each player is selected from a fixed set of strategies and that the results achieved using any combination of those strategies can be determined through simulation. Similar techniques have been used to analyse continuous double auctions [43]. The technique has been shown to be a useful tool in addressing questions about robustness of trading strategies and, following [46], in this paper, we apply it to a bilateral negotiation setting as follows.

In an EGT analysis, a profile specifies the strategy that is used by each player in the game. Furthermore, in the ANAC tournaments, the empirical payoff of any given strategy in any profile can be determined by averaging the utility it obtains in bilateral negotiations against all the other opponent strategies in that profile, where some of these opponents may use the same strategy.

In the particular bilateral encounters that we consider in this paper, although each tournament consists of up to eight players, due to the way their scores are calculated, it is possible to determine the results of any tournament without directly simulating every 8-player profile. In contrast, in the TAC SCM game, which was analysed by [44] using a similar

---

[13] To automatically generate strategies using an approach such as genetic programming would require the strategy space to be represented by a set of parameters. Due to the complexity of the strategies that could be used, it would be highly complex to generate such a representation, and this aspect is beyond the scope of this work.

technique, each game consisted of 6 players, with each having the freedom to choose their own strategy and interact with each other, and therefore it was necessary to obtain the results, by simulation, of every possible 6-player profile. In more detail, using Eq. (7), we can compute the results of any profile consisting solely of strategies in set S, provided that $\forall s_1, s_2 \in S, U(s_1, s_2)$ is known. This vastly reduces the number of games that need to be simulated, as there are only $|S|^2$ values for $U(s_1, s_2)$, compared to $\binom{|P|+|S|-1}{|P|}$ total strategy profiles (where $|P|$ is the number of players). For $|P| = 8, |S| = 8$, as in the tournaments we consider later, $\binom{|P|+|S|-1}{|P|} = \binom{15}{8} = 6435$, whereas $|S|^2 = 8^2 = 64$.

Using the payoffs achieved by each agent in a given profile (according to the strategies chosen by those agents), we consider the best single-agent deviation available to an agent in that profile. Here, a deviation is defined as the incentive of one agent to unilaterally change its strategy, assuming that all other agents maintain their current strategies. We consider that an agent has such an incentive to switch to another strategy if and only if this switch will bring a statistically significant improvement in its own utility. In contrast, if, in a given profile, no agent has an incentive to unilaterally deviate to another strategy (keeping the strategies of the other agents fixed), then that profile is said to be a pure Nash equilibrium. Specifically, since we use empirical results in our analysis, such a profile is an empirical pure Nash equilibrium. A game may have none, one or several such equilibria. Moreover, in some games, there may be a subset of profiles, each of which are not an equilibrium by themselves, but for which there exists a path of best deviations which connect them, and there is no best deviation which leads to a profile outside of the subset.[14] Such a subset is referred to as a *best reply cycle* [48]. We limit our analysis to pure equilibria, that is, where each agent plays a single strategy, rather than mixed equilibria in which an agent may choose a strategy probabilistically. This fits with the practical aspect to this work in that, in many real life negotiations, people would like to be represented by a strategy with a predictable behaviour rather than a probabilistic one, due to their risk aversion [3,33]. To summarise, we refer to two types of empirical *stable sets* (which we define as a set of one or more strategy profiles where there are no best deviations leading to profiles outside that set):

- *Single profile pure Nash equilibria*: a single strategy profile in which there is no statistically significant positive effect for any single agent in a profile to deviate.
- *Best reply cycle*: a set of strategy profiles from which there is no statistically significant positive effect for any single agent in a profile to deviate to a strategy outside that set, although there exist such deviations within the set.

Since a game may have more than one stable set, in order to assess the relative importance of the stable sets for a given game, we use the notion of *basin of attraction*. The basin of attraction of a stable set is defined as the number of profile states which, through a series of single-agent best deviations, will lead to that stable set. The size of the basin of attraction gives a measure of the likelihood of reaching that stable set given that initially each agent is equally likely to select any of the strategies. Similar to [43], we argue that the existence of a particular strategy in a stable set with a large basin of attraction suggests that the strategy is robust. For instance, if for most profile states, the best deviation available to the agents leads to a particular stable set containing a large proportion of strategy s_e, this shows that strategy s_e is robust in the long run (based on a number of repeated tournaments in which the players deviate to more favourable strategies) starting from a wide range of initial setups.

As indicated above, we use the empirical results from a set of simulations in order to compute an individual strategy payoff in each strategy profile, and to identify the best deviations. To do this, we first identify deviations which have a statistically significant positive effect on the utility of the deviating agent. From this set of deviations, we take the one which provides the largest increase in utility.

In this evaluation, we consider the set of strategies, $S \in \{B, G, H, I, K, N, T, V\}$, to be the strategies used in ANAC 2011 (see Table 7 for the mapping). Our evaluation is based on the results of the ANAC 2011 negotiations. These results allow us to determine the utility, $U(s_1, s_2)$, for any pair of strategies, $s_1, s_2 \in S, s_1 \neq s_2$, in each of the 8 negotiation scenarios. To complete our evaluation, it was also necessary to perform negotiations in which both agents use the same strategy, to allow us to determine $U(s, s)$ for each strategy $s \in S$, as such negotiations were not required for the actual tournament. These additional negotiation sessions were carried out in the same way, using the IRIDIS high-performance cluster, and repeated 30 times. These payoffs in any negotiation between any pair of strategies $s_1, s_2 \in S$ can then be averaged across all of the 8 scenarios in the final to get the joint payoff matrix presented in Table 7, which will be used in our analysis.

In what follows, we will apply this EGT analysis technique in two contexts:

- A single negotiation between two players.
- A negotiation tournament, in which a larger number of players take part in a set of two-player negotiations.

In so doing, we analyse both the underlying bilateral negotiation encounters, as well as the tournament format used in the competition. For the second analysis, we consider two versions of increasing complexity. First, we begin by considering tournaments of size 8, but in which agents can be one of three possible strategies, similar to the work presented in [46]. Specifically, we use the top three strategies from the ANAC 2011 final, as found in Section 4. We use this reduced set

---

[14] An intuitive instance of such a game is the well-known "rock, papers, scissors" game.

Table 7

Payoff matrix giving the scores averaged over all scenarios (showing the score of the row player). The letters in bold are the identifiers for each strategy, as used in subsequent figures.

	Agent **K**	**B**RAMAgent	**G**ahboninho	**H**ardHeaded	**I**AMhaggler2011	**N**ice Tit-For-Tat Agent	**T**heNegotiator	**V**alueModelAgent
Agent **K**	0.710	0.635	0.633	0.556	0.860	0.793	0.740	0.549
BRAMAgent	0.696	0.615	0.563	0.621	0.850	0.668	0.749	0.613
Gahboninho	0.744	0.707	0.713	0.574	0.912	0.728	0.824	0.687
HardHeaded	0.732	0.674	0.706	0.576	0.913	0.764	0.783	0.669
IAMhaggler2011	0.725	0.710	0.634	0.611	0.776	0.718	0.693	0.714
Nice Tit-For-Tat Agent	0.683	0.629	0.716	0.642	0.836	0.634	0.662	0.578
TheNegotiator	0.729	0.661	0.656	0.651	0.793	0.639	0.688	0.633
ValueModelAgent	0.613	0.587	0.619	0.609	0.733	0.550	0.607	0.593

of strategies, as this setup has a significantly smaller number of possible profile combinations and is therefore easier to visualise, without requiring extensive pruning of the state space.

Next, we turn our analysis to the full ANAC 2011 setup of eight agents. Here, each agent chooses a single strategy from the set of strategies used by the eight finalists. We performed this analysis for the scores achieved in each individual scenario, and also for the average scores achieved across all competition scenarios. However, in this work, we report on the analysis of the more interesting cases, specifically:

- The average scores achieved across all competition scenarios.
- The scores achieved in two selected scenarios which contain more extreme characteristics. Specifically, we selected the *Energy* scenario (which has the largest number of potential agreement points) and the *Grocery* scenario (which has the largest mean distance to the Pareto frontier). The purpose for this part of the analysis is to highlight interesting features which are not observed in the average-case analysis.

5.2.2. Two player negotiations among eight possible strategies

We start by considering the payoffs of the $|P| = 2$ player game, in which each player can select one of the $|S| = 8$ strategies of the competition finalists. In this case, there are $|S|^{|P|} = 8^2 = 64$ possible games that could be played, although some of them are equivalent due to symmetry. For example, the tournament in which player 1 uses strategy s_1 and player 2 uses strategy s_2, is equivalent to the one in which the strategies are swapped (player 1 uses strategy s_2 and player 2 uses strategy s_1). As a result, we can eliminate these, reducing the number of tournaments to $\binom{|P|+|S|-1}{|P|} = \binom{9}{2} = 36$.

In more detail, Fig. 9 shows the deviation analysis for this two-player game, according to the results in Table 7. In particular, each node represents a strategy profile, with the first row indicating the pair of strategies that form the profile. Here, each letter identifies a single strategy, as shown in bold text in Table 7. The second row gives the average utility achieved by the pair of agents. This is proportional to, and therefore a measure of, the social welfare achieved by the players in the profile. The social welfare is the sum of the utilities achieved by all agents in a profile and is a measure of the benefit of a negotiation to the entire set of agents, rather than to a specific agent. Each edge represents a deviation with a statistically significant positive effect. This figure shows that, from every strategy profile, there is an incentive for one agent to deviate and that, therefore, no pair of strategies is in equilibrium. However, it also shows that, from all strategy profiles, there exists a path of single-agent deviations which leads to a stable cycle of best deviations containing three profiles (denoted GN, KN, GK). The basin of attraction of this cycle includes 100% of the profiles, as the chain of best deviations from all strategy profiles will lead to one of the strategies in this cycle.

This two-player negotiation analysis provides some interesting insights. Surprisingly, no single competition strategy is in equilibrium against another agent which uses the same strategy, and more generally, there are no single profile pure Nash equilibria. Furthermore, although the winner of the competition, *HardHeaded* (H), achieves the highest score in negotiations against any other competition strategy, none of the profiles present in the best reply cycle include this strategy. However, this analysis still does not give us a full insight of how these strategies will perform in a tournament setup, where the average utility obtained across a range of opponents is taken. We address this issue in the next section.

5.2.3. Eight player tournaments composed of three strategies

In this section, we turn our attention to analysing tournament settings consisting of more than two players. In more detail, the standard ANAC competition consists of 8 agents, each of which negotiates independently against all other agents. We used Eq. (7) to determine the tournament score of each player. The values used in the equation were taken from a payoff matrix such as the one in Table 7. Therefore, it was easy to compute the scores for any tournament formed entirely from strategies present in the payoff matrix.

We begin our analysis by considering all tournaments consisting of $|P| = 8$ agents, each using one of the top $|S| = 3$ strategies found in the results in Section 4. There are $|S|^{|P|} = 3^8 = 6561$ such tournament profiles, but since the ordering of the agents is irrelevant, this can be reduced to $\binom{|P|+|S|-1}{|P|} = \binom{10}{8} = 45$ distinct tournaments. Given this, Fig. 10 shows these 45 tournaments, and the deviations that exist between them. In each node, the top row of the table gives the set of
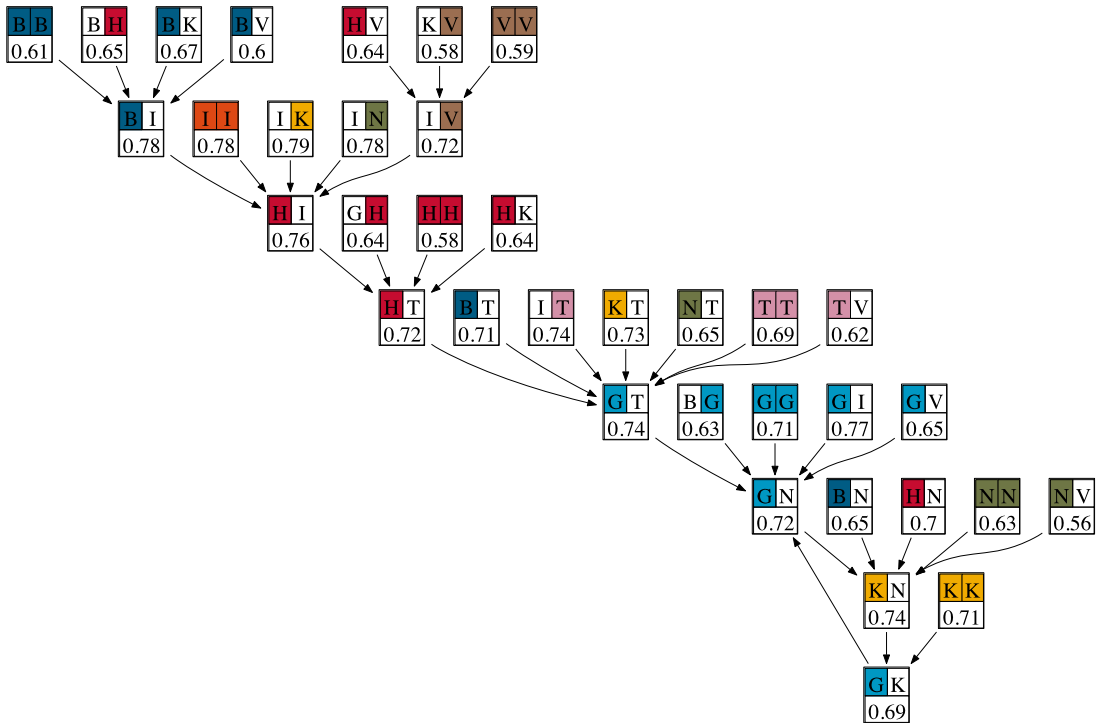
Fig. 9. Deviation analysis (by averaging the results over the eight scenarios) for the two-player negotiations. Arrows indicate statistically significant positive deviations between strategy profiles. At each node, the highest scoring agent is marked by a shaded background. (In the web version of this article, each agent is represented by a different colour.)

strategies present in the tournament. The second row of the table gives the number of agents using each strategy. Equilibria are indicated by nodes which have no outgoing arrow and which have a thicker border. The nodes in the corners of the triangle represent strategy mixtures in which all agents play the same strategy.

Using this analysis technique we find two pure Nash equilibria as follows:

  I. All eight agents use the *Gahboninho* strategy.
 II. Six agents use the *HardHeaded* strategy, whilst the remaining two agents use the *IAMhaggler2011* strategy.

The first equilibrium is the one with the largest basin of attraction, with 92% of the tournaments having paths of best deviation which lead to this equilibrium. In contrast, only the remaining 8% of tournaments have paths of best deviations which lead to the second equilibrium. This already highlights an interesting effect: the strategy *Gahboninho* (G), although not the actual winner of the competition[15] seems a very robust strategy, as the basin of attraction of the strategy profile containing only this strategy represents most of the profile space, at least in this restricted 3-strategy tournament.

5.2.4. Eight player tournaments composed of eight strategies

We now extend our analysis to consider tournaments in which all eight finalist strategies are available for an agent to choose from. We cannot represent such a tournament space using a triangular diagram, such as the one presented in Fig. 10. This is because in a tournament setting where $|P| = 8$ agents can each choose one of $|S| = 8$ strategies, there are $|S|^{|P|} = 8^8 = 16.8 \times 10^6$ possible combinations of tournaments. Since the ordering of the agents in a strategy profile does not matter, this can be reduced to $\binom{|P|+|S|-1}{|P|} = \binom{15}{8} = 6435$ distinct combinations. While we have in fact constructed the full graph corresponding to the EGT analysis for this case, such graphs still contain far too many nodes to visualise.

Therefore, in order to construct meaningful visualisations for this case, we prune the full graph containing all possible tournaments to just a few nodes that capture the most interesting information present in the full graph. More specifically, the nodes corresponding to the following tournaments were retained in our visualisation:

---

[15] We note, though, that *Gahboninho* was the winner of the qualifying round. The reason for the change in position can be explained by the effect of a full pairwise matching of all agents at each stage. Specifically, the qualifying round contained a number of weaker agents (those which were eliminated), compared to the final. Thus *Gahboninho* was able to achieve a higher average score against a wider range of strategies present in the qualifying round than against the stronger finalists.
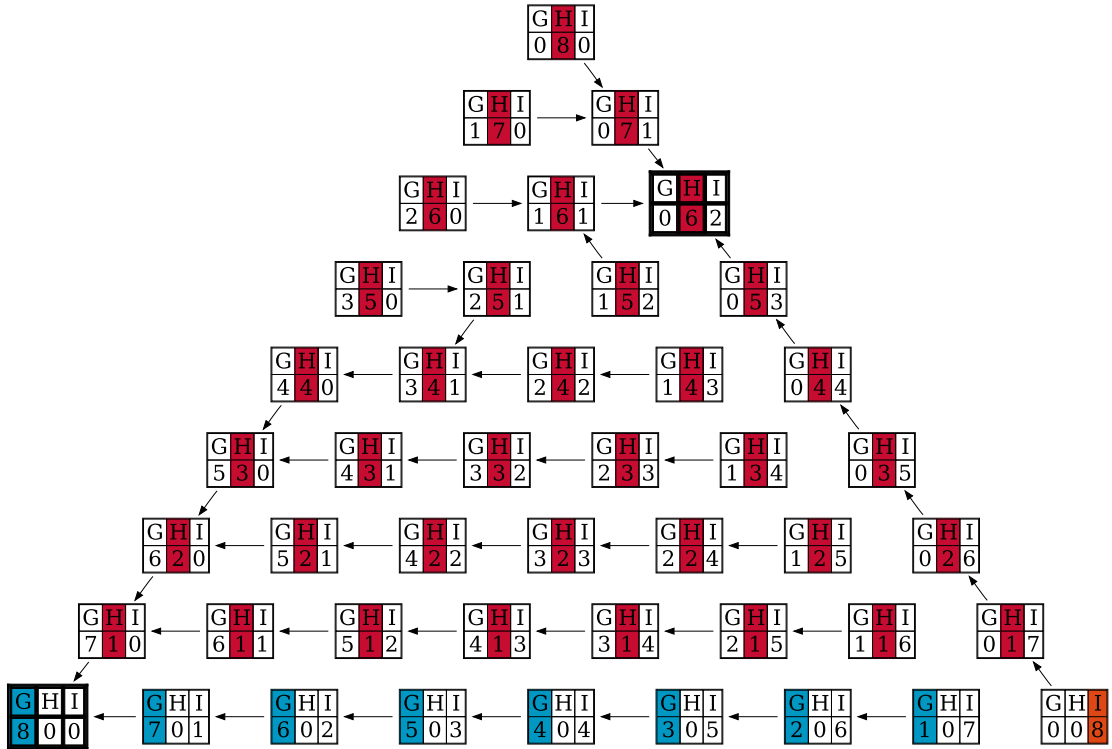
Fig. 10. Deviation analysis (by averaging the results over the eight scenarios, for top three strategies in the ANAC 2011 competition: *Gahboninho*, *HardHeaded*, *IAMhaggler2011*). Arrows indicate statistically significant positive deviations between strategy profiles. Equilibria are marked with a thicker border. At each node, the highest scoring agent is marked by a shaded background. (In the web version of this article, each agent is represented by a different colour.)

Type I. All profiles in which all agents use the same strategy. These are the most extreme profiles, represented by the corners of the triangular diagram in Fig. 10.

Type II. The profile which represents the actual ANAC 2011 setup, in which each of the 8 agents uses one of the finalist strategies. Thus, this is the only profile in which all 8 finalist strategies are present and no two agents use the same strategy. This profile is marked with a thick border.

Type III. All pure Nash equilibria and best reply cycle profiles. All such profiles are marked with a thick border.

Type IV. All profiles which are on the path of best deviations from a profile of type I or II to a profile of type III.

Therefore, the pruning of the graph was done as follows. Any nodes (except for those representing profiles of type I or II) which did not have an incoming arrow were progressively removed, as these nodes could not possibly represent any of the four types of profile that were to be retained. This process was repeated until no further nodes could be removed. By performing this pruning, the size of the graphs was significantly reduced, whilst retaining all nodes which belong to one of the four types detailed above.

In some cases, to further reduce the size of the graphs, we replaced paths of repeated deviations (where a single strategy, s_a, deviates to another, s_b, repeatedly) with a single dashed edge. This additional pruning further reduces the size of the graphs, without removing any significant detail, and again maintains any pure Nash equilibria and best reply cycles.

The above procedure was used to construct the 8 player, 8 strategy visualisations both in the average case (shown in Fig. 11), and for two scenarios in which special features of the competition can be highlighted (Figs. 12 and 13).

Analysis of results averaged over all scenarios. The deviation analysis using the average scores over all scenarios reveals a single pure equilibrium in which 7 of the agents use the *Gahboninho* strategy, and the remaining agent uses the *Nice Tit-for-Tat* strategy. This equilibrium has a basin of attraction consisting of 37% of the tournament space, shown in the top half of Fig. 11. The remaining 63% of the space form a basin of attraction which leads to a best reply cycle, shown in the bottom half of Fig. 11. In all tournaments which form this cycle, it is the agents which use the *HardHeaded* strategy which achieve the highest score. Overall, this analysis seems to confirm that *Gahboninho* is a robust strategy, that appears in most stable set tournament profiles, although often in combination with the winner, *HardHeaded*, and *Nice Tit-for-Tat*.

However, it is worth pointing out that this average case analysis, while useful, also hides a lot of detailed information, since we average over all scenarios. As mentioned in Section 3.1, in the ANAC 2011 competition format, the scenarios across which the strategies were tested were also submitted by the teams themselves (each team submitting one scenario, leading
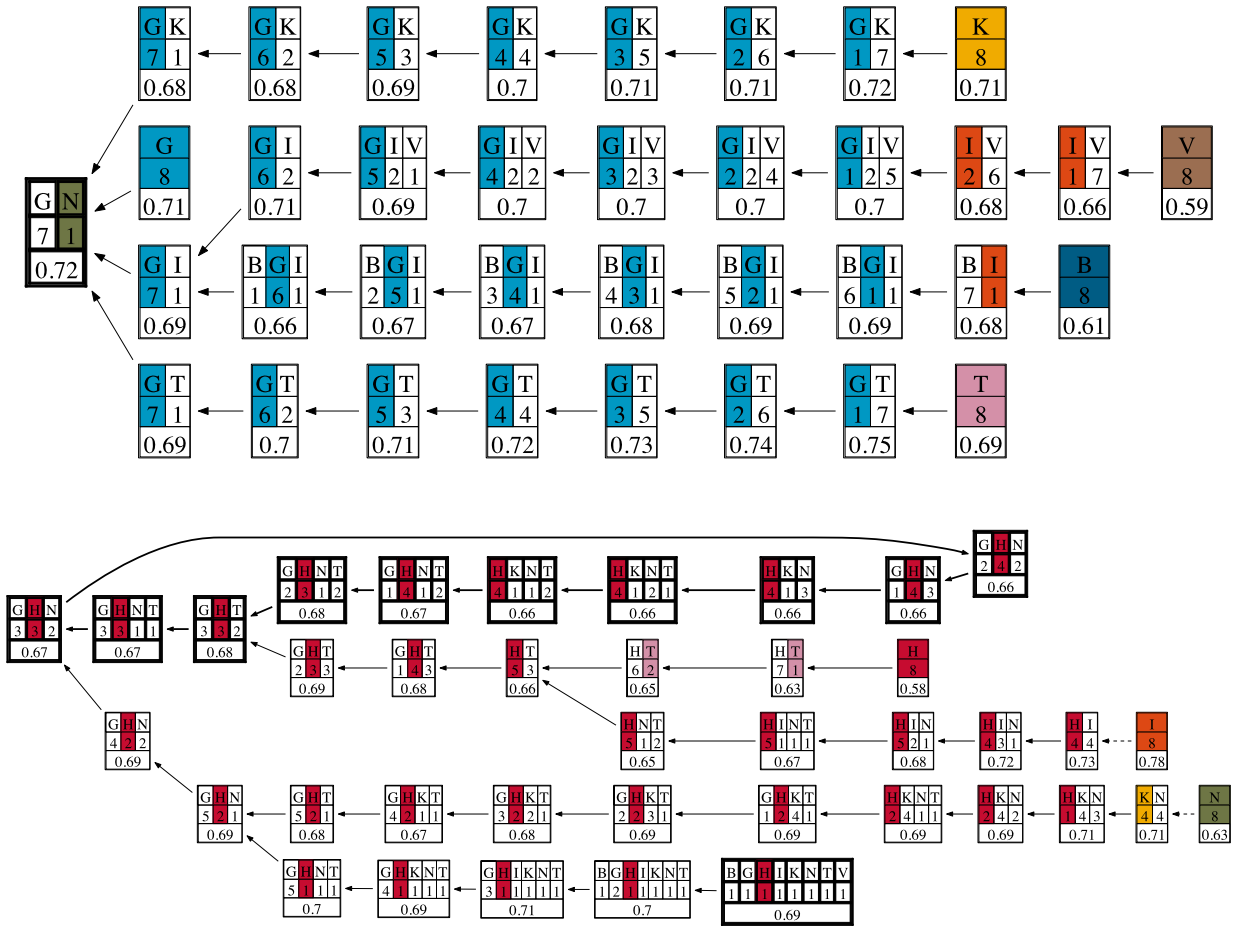
Fig. 11. Deviation analysis (by averaging the results over the eight scenarios). Arrows indicate statistically significant positive deviations between strategy profiles. Dashed arrows indicate a series of repeated deviations. Profiles which belong to stable sets (and also the original ANAC 2011 tournament) are marked with a thicker border. At each node, the highest scoring agent is marked by a shaded background. (In the web version of this article, each agent is represented by a different colour.)

to the 8 scenarios from the final presented here). While this made the competition less vulnerable to possible bias on the part of the organisers, it also meant that most teams submitted relatively small, cooperative scenarios and only a few of which included time discounting (presumably, this is partially because smaller scenarios are easier to design). In such scenarios, it is relatively easy for both agents to get a high utility, which advantages strategies exhibiting less concessive behaviour. Specifically, we would like to understand how different strategies perform in larger, more complex scenarios, which have stronger opposition and have outcomes which are far from the Pareto frontier, in the sense that agents need to explore the search space more in order to reach good agreements. To this end, we have chosen to present in more detail the 8 agent, 8 strategy analysis in two scenarios from the final, which we believe exhibit these more challenging features.

Analysis of the energy scenario. This scenario was the largest scenario from the competition, with over 390,000 possible agreements. Its opposition is also quite strong, with the two agents' utilities in the Nash bargaining solution (the one which maximises the product of the agents' scores) having a geometric mean[16] of 0.687. In contrast, most of the other competition scenarios have a Nash bargaining solution with a geometric mean greater than 0.84. Fig. 1(h) shows the outcome space of this scenario. In a scenario with strong opposition (such as the *Energy* scenario), it is necessary for at least one of the agents to concede a significant amount of utility in order to reach an agreement. Due to this, we found that there are two strategies, *HardHeaded* and *ValueModelAgent*, which fail to reach agreement in self-play (or against each other), indicated by a zero score in the tournaments in which all agents use one of those strategies. This is due to their tough behaviour in making any concessions. Since this scenario is quite competitive, if both agents take a tough approach, there is no part of the outcome space that both agents can reach and therefore, they are unable to reach any agreement. In contrast, in a

---

[16] The geometric mean of outcome ω is given by $\sqrt{U_1(\omega) \times U_2(\omega)}$, where $U_1(\omega)$ is the utility to agent 1, $U_2(\omega)$ is the utility to agent 2. Provided that all utilities are positive, the solution which maximises the product of the utilities is also the one which maximises the geometric mean.
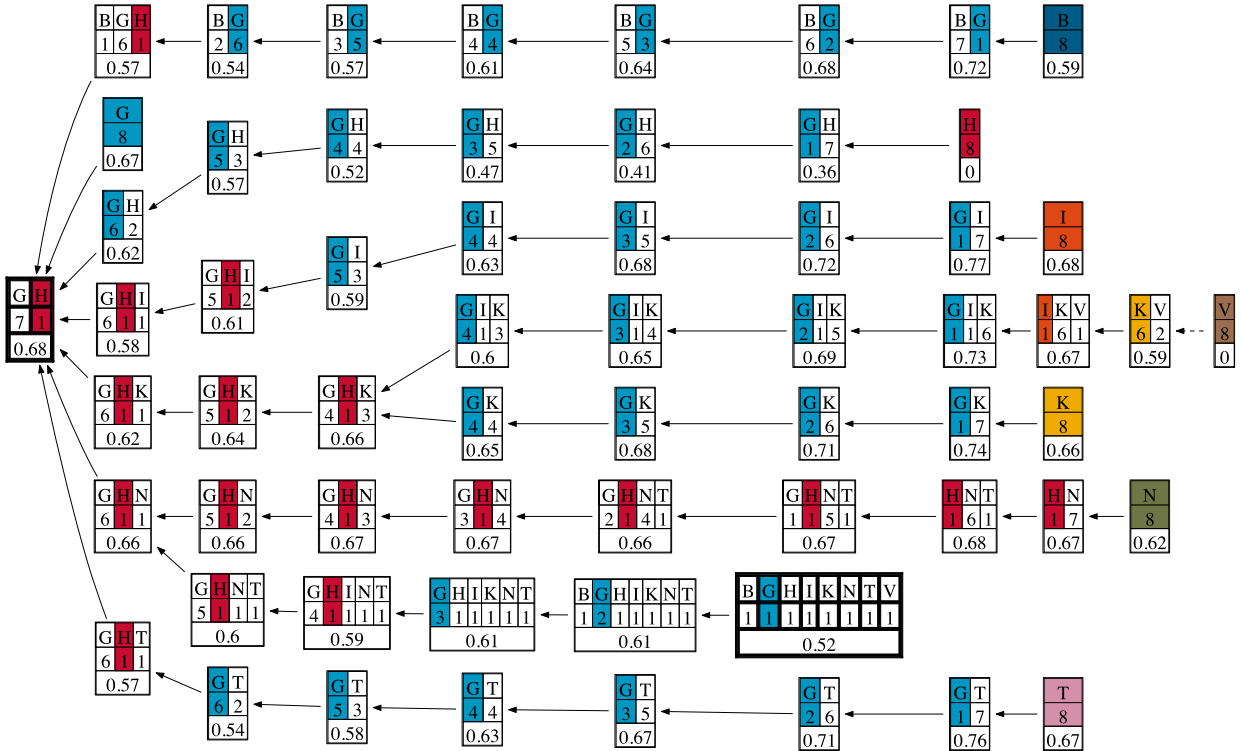
Fig. 12. Deviation analysis for the *Energy* scenario. Arrows indicate statistically significant positive deviations between strategy profiles. Dashed arrows indicate a series of repeated deviations. Equilibria (and also the original ANAC 2011 tournament) are marked with a thicker border. At each node, the highest scoring agent is marked by a shaded background. (In the web version of this article, each agent is represented by a different colour.)

scenario with a very high scoring Nash bargaining solution (such as the *Car* scenario, which has a Nash bargaining solution with geometric mean 0.935), it is possible for an agreement to be reached with very little concession from either agent. The effects seen in this scenario are similar to that of the classic Hawk–Dove game [39] in that there are two asymmetric two-player equilibria (where one player uses the hawk-like, *HardHeaded* strategy and the other uses the dove-like, *Gahboninho* one), but there is no dominant strategy. However, if a player knows that its opponent is using the *HardHeaded* strategy, it can benefit by playing the *Gahboninho* one (and vice versa).

Especially in a tournament setting, there may be an incentive for a small proportion of the agents to take a tough approach, in order to take advantage of the more concessive opponents. To this end, Fig. 12 presents the deviation analysis for this scenario, which reveals a single pure equilibrium in which 7 of the agents use the *Gahboninho* strategy and the remaining agent uses the *HardHeaded* strategy. From all non-equilibrium tournaments, there exists a path of best deviations which leads to this equilibrium. In this equilibrium, it is the tough, *HardHeaded* strategy which wins the tournament, even though the majority of agents use the *Gahboninho* strategy. However, there is no incentive for any more agents to deviate to using the *HardHeaded* strategy, since if any agent playing the *Gahboninho* strategy in the "7 Gahboninho – 1 HardHeaded" equilibrium unilaterally switches to playing *HardHeaded*, this will decrease its payoff.

Analysis of the grocery scenario. This scenario is one of the discounted scenarios from the competition. The opposition is weaker than the *Energy* scenario, in that it is possible to reach an agreement which gives the two agents a mean utility as high as 0.867. However, a special feature of this scenario (in comparison to all the other scenarios entered in the competition) is that the outcome space is made up of points that, on average, lie far from the Pareto frontier. That is, the outcomes are not clustered around the Pareto frontier (as in most other scenarios submitted), but rather are widely distributed throughout the utility space (Fig. 1(d) provides a graphical illustration of the outcome space of this scenario). This means that there are many agreements that can be reached which are far from optimal for both parties, so agents really need to explore to reach a more profitable part of the outcome space. Specifically, the average geometric mean of the two agents' utilities for all 1600 outcomes in this space is just 0.425, so in this scenario, the utility that can be achieved through effective negotiation by both agents (i.e. 0.867) can be more than twice what could be expected from simple random selection.

Interestingly, the stable set profiles revealed by the EGT analysis in this scenario are rather different than in the general case. There is one pure equilibrium, in which all 8 agents use the *IAMhaggler2011* strategy, which has a basin of attraction
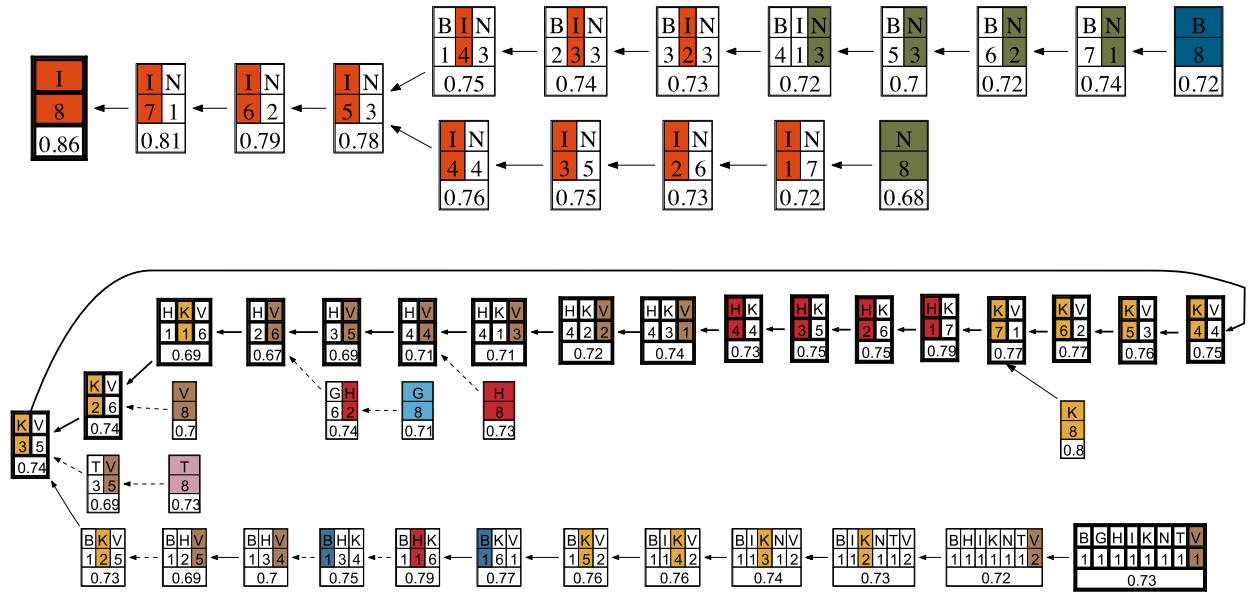
Fig. 13. Deviation analysis for the *Grocery* scenario. Arrows indicate statistically significant positive deviations between strategy profiles. Dashed arrows indicate a series of repeated deviations. Profiles belonging to stable sets (and also the original ANAC 2011 tournament) are marked with a thicker border. At each node, the highest scoring agent is marked by a shaded background. (In the web version of this article, each agent is represented by a different colour.)

consisting of 54% of all the tournament profiles, as shown in the top half of Fig. 13. The remaining 46% of the tournament space leads to a best deviation cycle, shown in the bottom half of Fig. 13.

The profiles leading to these stable sets actually form subgraphs which are disconnected from each other with the strategies *IAMhaggler2011*, *BRAM-Agent* and *Nice Tit-for-Tat* eventually deviating to the pure *IAMhaggler2011* profile, and the profiles containing the rest of the strategies deviating to the best reply cycle. The reason for this separation is that *IAMhaggler2011*, *BRAM-Agent* and *Nice Tit-for-Tat* are more cooperative (or adaptive), and will explore the space more to ensure a jointly profitable outcome. This matches the observations made in Section 5.1, although the EGT analysis in this section refers only to the *Grocery* scenario (which has an outcome space with a higher mean distance to the Pareto frontier than the others).

By contrast, the strategies appearing in the best reply cycle in Fig. 13 aim more at exploiting the opponent by playing tough, but in a scenario where parts of the outcome space are far from the Pareto frontier, this may backfire, as the mutually profitable part of the search space may be rarely reached. This can also be seen by comparing the average social welfares of the agents in two sets of stable sets, which are shown in Fig. 13, as a number for each state.

6. Discussion

The 2011 competition was a successful event, enriching the research field on practical automated negotiation in line with the aims as set out in Section 1. In particular, the widespread availability of efficient, general and domain-independent automated negotiators, which this tournament has achieved, has the advantages of minimising the effort required for adaptation of a general automated negotiator to a new domain. Furthermore, the availability of the different agents allows researchers to have an objective measure to assist them in validating and testing the effectiveness of future automated negotiators. Against this background, in this section we summarise the main insights gained from the competition process and the agents submitted.

We incorporated a number of features into the competition environment to increase realism and to encourage the development of flexible and practical negotiation agents. In particular, in contrast to the first competition, ANAC 2010, we introduced a discount factor for some of the scenarios and a shared time-line. The latter means that, if one agent causes a delay, this will affect both agents equally, both in terms of the discounting and getting closer to the deadline. In addition, and similarly to the previous year, the scenarios were unknown to the agents prior to the tournament and the agents received no information about their opponent's preferences during the tournament. Another degree of uncertainty is the strategies used by the opponents. Thus, although it was possible to learn from the agents and domains in ANAC 2010, successful agents submitted to ANAC 2011 had to be flexible, domain-independent and operate in both discounted and undiscounted settings.

If we consider the agents submitted, most of them were indeed adaptive to some degree, in that they take into account the strategy of their opponent. Several of these (such as *IAMHaggler2011*, *Gahboninho*, *Nice Tit-for-Tat* or *ValueModelAgent*)

make use of machine learning techniques to try and predict the concession or preferences of the opponent. Another important impact of a bidding strategy is how it *affects* subsequent offers by the negotiating counterpart. For example, *Nice Tit-for-Tat* tries to entice the opponent into cooperative behaviour by reciprocating the opponent's concessions. Such behaviour is a form of second-level adaptivity, in the sense that it not only adapts to the opponent, but also attempts to influence the opponent's behaviour. It will be interesting to see if other principles than tit-for-tat can elicit such behaviour in future competitions.

Despite being adaptive, a possible limitation of several of the agents is that they rely on fixed, pre-determined parameters in their strategies, such as the time elapsed before the agent becomes more concessive or the utility an agent should concede to at a given time. Such fixed parameters are used both by less successful agents, such as *ValueModelAgent*, *BRAM-Agent* and *TheNegotiator*, but also by the winner *HardHeaded*. Strategies that try to be more generic, and avoid relying too much on hand-tuned parameters include *IAMhaggler2011*, *Gahboninho* and *Nice Tit-for-Tat*.

Now, it is interesting to see that, while the environment encourages flexibility, being adaptive does not necessarily benefit the agents in terms of winning the competition. In particular, it may come as a surprise that *HardHeaded*, a relatively simple and inflexible strategy, relying heavily on hard-coded parameters, achieved first place in the finals. There are several reasons for this. First of all, such a strategy relies on other, more complex strategies, to eventually adapt to its demands. This is seen in the EGT analysis, specifically in the *Energy* scenario, where the *HardHeaded* agent scores zero if all other agents use that strategy, but against a set of more adaptive opponents (and *Gahboninho* at the most extreme) the *HardHeaded* agent performs well. Thus, if there are sufficient adaptive strategies, being hard-headed is a good strategy to use. Second, we see that winning the competition does not require the agent to win in all or even most scenarios. In fact, the *HardHeaded* agent only had the highest utility in 3 scenarios out of the 8, and the runner-up, *Gahboninho*, had the highest utility in 2 of the scenarios in the finals. As long as it wins by a large margin in those games where it comes first, it can win the entire competition.

Clearly, therefore, setting the right criterion for winning the competition is important. An alternative criterion to further encourage the development of flexible negotiators would be the total number of games won, instead of the average utility obtained in these games. However, such a criterion would encourage agents to simply beat their opponent, rather than maximise their own utility. This means that agents would be encouraged to get more utility than their opponent at all cost, even if this means reducing their own utility (e.g. by delaying the agreement in the case of a discounted scenario). Such spiteful behaviour has also been found in auction bidding [6,42] and is realistic in some cases, but obviously encouraging such strategies should not be the goal of the competition.

Our analysis provided further insights into the proficiency and adaptability of the submitted agents. In particular, the analysis of self-play found that, especially in settings with strong opposition, such as the *Energy* scenario, *HardHeaded* failed to reach any agreement in self-play, achieving a zero score. While self-play results were not considered when computing the competition ranking (because, arguably, this would have introduced some bias), such results, as well as robustness to a wide range of opponent strategies, are likely to be important in the practical success of automated negotiation agents.

When it comes to robustness and flexibility, the *Gahboninho* agent, which achieved the highest average utility in the qualifying round, was found to be particularly robust in the EGT analysis. The main reason for this is its special type of adaptivity: *Gahboninho* has a meta-learning strategy which first tries to establish the learning behaviour of the negotiation opponent, and then uses this to exploit its opponent. For example, if the opponent is adapting to the strategy of the agent, then *Gahboninho* will be less flexible. If, on the other hand, the opponent does not seem to adapt, *Gahboninho* will be more flexible. In other words, the strategy tries to establish whether its opponent is a teacher or learner, and adapts accordingly (note that this teacher/learner dilemma has been identified in other game-theoretic learning competitions, such as the Lemonade Game [41]). Such an approach has been shown to be successful in the ANAC competition, but also provides useful insights for practical negotiations in general, and is likely to be useful in future research on automated negotiation.

Other strategies, notably *IAMhaggler2011*, also proposed a very flexible and generic learning heuristic. Its high degree of adaptivity enables the agent to reach efficient agreements, even in large domains, or in scenarios that are subject to considerable time discounting. However, while *IAMhaggler2011* performed well in general, this did not secure it a winning position overall, nor in many of the specific scenarios. *Nice Tit-for-Tat* is also a strategy that could provide substantial insights for future research, as it provides an interesting extension of the tit-for-tat concept to multi-issue negotiation spaces. The failure of these strategies to win in the competition is due to their relative "niceness", i.e. willingness to adapt to their opponent's demands.

Interestingly, if we analyse how the submitted agents deal with the multi-issue aspect, only a few explicitly modelled the utility space, in order to try to get solutions close to the Pareto frontier (*Nice Tit-for-Tat* and *HardHeaded* are examples of an agent which did model the utility space). This is probably due to the fact that agents could potentially exchange a very large number of offers (up to several million in some cases), allowing for exhaustive exploration of all but the largest utility spaces. Several of the agents (such as *IAMhaggler2011*, *AgentK2* or *HardHeaded*) used learning to determine the appropriate concession level, but they used random search to sample the space for offers with utilities above that level. Such strategies did surprisingly well in the competition, because the number of offers that can be exchanged between the agents in the span of 3 minutes easily reaches tens of thousands. Such a large number of offers clearly makes it easier for software agents to explore the utility space, yet one could argue whether this is entirely realistic. Even though software agents are able to compute many more offers than humans, in practice there may be other constraints, such as network delays, that would

limit the number of offers bargaining agents could exchange. Thus, introducing a minimum time delay between offers may be an issue to consider in future competitions.

Finally, we discuss the submitted scenarios. The results show that, in both the *Energy* and *Nice or Die* scenarios, the agents obtain lower utility on average, compared to the other scenarios. Moreover, the performance of the agents in these scenarios are more diverse, with some agents having much higher utilities than others. In other scenarios, these utilities are much closer. The reason is that, even though the scenarios are quite distinct (with the *Energy* being the largest and *Nice or Die* being the smallest) both of these scenarios have strong opposition (i.e., are relatively competitive), whereas in the other scenarios it is possible to achieve close to the maximum score for both agents. Due to the diversity, the performance in these scenarios has a greater impact on the overall utility, and therefore the strategies which do well in these scenarios have a definite advantage over the other strategies. Therefore, it is important to maintain a variety of scenarios with different characteristics to properly evaluate the performance of an agent. Another parameter that can potentially affect the diversity of the outcomes is the discount factor. Interestingly, however, the results show no clear connection between the discount factor and the diversity in performance. This could be due to the fact that the most discounted scenarios (i.e. *Laptop* and *Company Acquisition*, which have the lowest discount factor) also have weak opposition and small domains, meaning that win–win agreements can be easily found. In future, it would be interesting to apply the discount factor to other types of scenarios as well (i.e., with larger domains and stronger opposition), to see the impact of the discount factor in more challenging settings.

7. Conclusions and future work

This paper describes and analyses the Second International Automated Negotiating Agents Competition (ANAC 2011). The main purpose of ANAC is to motivate research in the area of bilateral multi-issue negotiations, with an emphasis on the practical design and development of successful automated negotiating agents. Additional goals include: collecting and objectively evaluating different state-of-the-art negotiation strategies and opponent models, defining a wide variety of benchmark negotiation scenarios, and making them available to the negotiation research community. Based on the submissions and the process of running the competition, as well as the post-tournament analysis reported in this paper, we believe that this competition has served its purposes. Eighteen teams from seven different institutes participated, and we expect the trend of increasing participation to continue. Moreover, ANAC 2011 has also been successful in collecting 18 new negotiation scenarios and has provided a number of new strategies which can be used as benchmarks to test the efficacy of subsequent work in this area.

The competition has also advanced the development of Genius. In particular, we have released a new, public build which can be used to run negotiation tournaments. Furthermore, we have made available all negotiation scenarios, agents and results of the competition. This will make it possible for the negotiation research community to do a complete re-run of ANAC and to perform subsequent in-depth analysis of other facets of negotiation encounters.[17]

In addition to describing the competition and its results, the aim of this paper is an in-depth analysis of the strategies and the tournament results. To this end, we introduce two different evaluation methods in this paper. In the first part of our analysis, we focus on the influence of different opponents on the negotiation outcome. Specifically, we characterise negotiation strategies by how they conceded throughout the negotiation, and we do so by using a quantitative measure called the concession rate of an agent. From this, we show that it pays for negotiating agents to cooperate as little as possible against an agent that concedes steadily through time. In general, the correlation is very high between the concession rate and utility obtained against such a conceding strategy. This means that a low concession rate is a good predictor for high performance against a conceding agent. The same behaviour would be inappropriate against a very competitive agent that does not concede at all during the negotiation. We have also demonstrated that, in general, the nicer an agent plays against this type of opponent, the more utility it obtains. This is because, against such a hard-headed opponent, there is a high chance the negotiation would break down if no concessions are made.

This then leads to a dilemma for the agents to be either a teacher (who tries to entice their opponent to adapt by employing a tough strategy) or a learner (who tries to adapt to maximise its own utility, given the behaviour of the opponent), which lies at the heart of many bilateral negotiation problems. If the opponent is flexible and adapting to one's demands, there is little point in conceding. However, if the opponent is being strictly hard-headed (even appearing irrationally so), reaching some agreement is typically preferable to no agreement. The importance of learning strategies that try first to detect the adaptivity of the opponent (such as *Gahboninho*, which proved very robust in ANAC 2011) is an important insight which could be taken up in further research in bilateral negotiations beyond the competition.

The focus of our concession-based analysis of the agents has been on bidding behaviour, and not on acceptance strategy. In general, this is also an important part of a negotiator's strategy, which is a key determinant of the outcome of a negotiation, not least due to the fact that acceptance ends a negotiation. We believe the same issues of competitiveness and adaption to the opponent should play a role when agents decide when and whether to accept. This could provide an interesting direction for future research.

In the second part of our analysis, we ascertain the robustness of the agents by considering different tournament sizes and compositions. Specifically, by using techniques from empirical game theory (EGT), we analyse when and whether agents have an incentive to deviate (i.e. switch) to a different strategy. In particular, we argue that, since an agent never knows which opponent it faces in practice, strategies which are part of an equilibrium, and have a large basin of attraction (i.e., where sequences of deviations often lead to that strategy), are more robust than strategies which only appear in few games or are not in equilibrium. This analysis highlights several interesting conclusions. For instance, we found that the *Gahboninho* strategy (which ranked second overall) was actually the most robust, across the 8 scenarios from the final. Moreover, we see that, depending on the characteristics of the scenario, such as the discount factors used, there are considerable differences in the empirical equilibria that are formed and the basin of attractions corresponding to those equilibria. Our results demonstrate the importance of a comprehensive analysis which considers the robustness of agents across a variety of settings, rather than just considering the average utility achieved by negotiation outcomes in a single scenario, which is the standard metric in existing negotiation literature. We argue that such EGT techniques should be used more widely by the community in order to build more robust and practical bargaining agents in the future.

As with many competitions, ANAC is continually evolving to address new challenges and issues. In particular, for future work, we would like to include domains with a combination of continuous and discrete issues. Our current results are based on discrete issues, where each issue takes a value from a finite set. In contrast, a continuous issue takes any value in a defined range. The GENIUS platform already handles continuous issues, but all the domains and strategies submitted by the participants in 2011 focused on the discrete case. Also having continuous issues may force the participants to develop strategies specifically geared to such settings, which would generalise the agents even further and benefit application domains where continuous issues naturally occur (such as the allocation of continuous resources, like electricity or time). Another important extension is to consider agent utility functions with inter-dependencies between the issues being negotiated, such as those considered in [12,28,35]. So far, all the agent utility functions considered in the competition were additive. However, this may be a limitation since in many real-life scenarios the utility functions of different agents exhibit complex interdependencies between issues.

Given the lessons learnt from running ANAC 2011, we can identify several directions for extending the format of this international competition in the future. In more detail, in the upcoming years, we intend to introduce several tracks to the competition, that model different aspects of the automated negotiation problem, similar to the tracks of the Trading Agents Competition (TAC). We see these parallel competition tracks as naturally supporting the different strands of ongoing research, that are considered important in the automated negotiation community.

Firstly, we intend to have a competition track in which participants have to design agent strategies that deal not only with bilateral, but also with concurrent, one-to-many negotiation threads. In this type of negotiation, the concession strategy in each thread may be considerably influenced by the offers made and received in the parallel threads. Such issues have already been explored by the negotiation community, e.g. [1,2,30,47], but there is no universally accepted benchmark to compare such approaches.

Another important direction is to include a track in which participants are asked to submit agents that perform well not only against other automated strategies, but also against human opponents. Recent research has demonstrated that agents designed to be matched with other automated agents do not necessarily scale when matched with people [25], and we believe this to be the case with the agents present in this year's competition. Emotions, culture and computability are some of the issues which need to be considered when designing such an agent; issues that were not tackled when designing current agents. Specifically, different approaches are required, since negotiations with humans need to be much shorter, in terms of the number of offers that can be exchanged in a short timespan.

Yet another direction of further work could be around *mediated* negotiation scenarios, in which a mediator agent has the task of eliciting the complex utilities of the two (or more) negotiating agents suggesting mutually agreeable outcomes.[18] In such a setting, which is somewhat similar to the TAC market design competition [38], the negotiating agents themselves would be provided as part of the platform, while participants would be asked to design the policy of the mediating agents. A further track could consider a repeated series of negotiations in which the agents can learn how to negotiate with an opponent based on previous interactions with that opponent.

Finally, we believe our work has influence outside the framework of the competition. For example, the relative success of a meta-learning strategy such as *Gahboninho* shows that the teacher vs. learner dilemma (already observed in other competitions that study game-theoretic learning, such as the Lemonade Stand Game [29,41]) may be an important one to also explore in electronic negotiation. In fact, the study of learning and concession behaviour of agents in bilateral negotiations provides a natural framework to explore such issues. Moreover, as we discussed above, we believe other measures for evaluating negotiating agents, such as performance in self-play or robustness against different pools of strategies (such as highlighted by EGT analysis) could be used in other bilateral negotiation settings as well.

---

[18] Note that mediated negotiation is different from combinatorial auctions, since players do not have to reveal their full utility space to the mediator in advance, but can do so incrementally.

Acknowledgement

References

[1] B. An, N. Gatti, V.R. Lesser, Extending alternating-offers bargaining in one-to-many and many-to-many settings, in: Proc. of the 2009 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2009, pp. 423–426.

[2] B. An, V.R. Lesser, K.M. Sim, Strategic agents for multi-resource negotiation, Autonomous Agents and Multi-Agent Systems 23 (1) (2011) 114–153.

[3] K. Arrow, The Theory of Risk Aversion, 1965.

[4] R. Axelrod, The Evolution of Cooperation, Basic Books, New York, NY, 1984.

[5] T. Baarslag, K. Hindriks, C.M. Jonker, Towards a quantitative concession-based classification method of negotiation strategies, in: D. Kinny, J.-j. Hsu, G. Governatori, A. Ghose (Eds.), Agents in Principle, Agents in Practice. Lecture Notes of The Fourteenth International Conference on Principles and Practice of Multi-Agent Systems, in: Lecture Notes in Computer Science, vol. 7047, Springer-Verlag, 2011, pp. 143–158.

[6] F. Brandt, T. Sandholm, Y. Shoham, Spiteful bidding in sealed-bid auctions, in: Proc. of Twentieth International Joint Conference on Artificial Intelligence, 2007, pp. 1207–1214.

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

[8] P. Faratin, C. Sierra, N.R. Jennings, Using similarity criteria to make trade-offs in automated negotiations, Artificial Intelligence 142 (2) (2002) 205–237.

[9] S.S. Fatima, M. Wooldridge, N.R. Jennings, Multi-issue negotiation under time constraints, in: Proc. of the First International Joint Conference on Autonomous Agents and Multiagent Systems, ACM, New York, NY, USA, 2002, pp. 143–150.

[10] S.S. Fatima, M. Wooldridge, N.R. Jennings, Optimal negotiation strategies for agents with incomplete information, in: Revised Papers from the Eighth International Workshop on Intelligent Agents, Agent Theories, Architectures, and Languages'01, Springer-Verlag, 2002, pp. 377–392.

[11] S.S. Fatima, M. Wooldridge, N.R. Jennings, A comparative study of game theoretic and evolutionary models of bargaining for software agents, Artificial Intelligence Review 23 (2) (2005) 187–205.

[12] K. Fujita, T. Ito, M. Klein, Approximately fair and secure protocols for multiple interdependent issues negotiation, in: Proc. of the Eighth International Conference on Autonomous Agents and Multiagent Systems, 2009, pp. 1287–1288.

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

[14] K. Hindriks, C.M. Jonker, D. Tykhonov, The benefits of opponent models in negotiation, in: Proc. of the 2009 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, vol. 2, 2009, pp. 439–444.

[15] K. Hindriks, D. Tykhonov, Opponent modelling in automated multi-issue negotiation using bayesian learning, in: Proc. of the Seventh International Joint Conference on Autonomous Agents and Multiagent Systems, 2008, pp. 331–338.

[16] K. Hindriks, D. Tykhonov, Towards a quality assessment method for learning preference profiles in negotiation, in: Proc. of the Tenth International Workshop on Agent Mediated Electronic Commerce, 2008, pp. 5–17.

[17] T. Ito, H. Hattori, M. Klein, Multi-issue negotiation protocol for agents: Exploring nonlinear utility spaces, in: Proc. of the Twentieth International Joint Conference on Artificial Intelligence, 2007, pp. 1347–1352.

[18] N.R. Jennings, P. Faratin, A.R. Lomuscio, A.D. Parsons, M. Wooldridge, C. Sierra, Automated negotiation: Prospects, methods and challenges, Group Decision and Negotiation 10 (2) (2001) 199–215.

[19] C.M. Jonker, V. Robu, J. Treur, An agent architecture for multi-attribute negotiation using incomplete preference information, Journal of Autonomous Agents and Multi-Agent Systems 15 (2) (2007) 221–252.

[20] S. Kawaguchi, K. Fujita, T. Ito, Compromising strategy based on estimated maximum utility for automated negotiating agents, in: T. Ito, M. Zhang, V. Robu, S. Fatima, T. Matsuo (Eds.), New Trends in Agent-Based Complex Automated Negotiations, in: Series of Studies in Computational Intelligence, Springer-Verlag, 2010.

[21] S. Kraus, Strategic Negotiation in Multiagent Environments, MIT Press, Cambridge, MA, 2001.

[22] S. Kraus, P. Hoz-Weiss, J. Wilkenfeld, D.R. Andersen, A. Pate, Resolving crises through automated bilateral negotiations, Artificial Intelligence 172 (1) (2008) 1–18.

[23] S. Kraus, J. Wilkenfeld, G. Zlotkin, Multiagent negotiation under time constraints, Artificial Intelligence 75 (2) (1995) 297–345.

[24] R.J. Lewicki, D.M. Saunders, J.W. Minton, Essentials of Negotiation, McGraw-Hill, Boston, MA, 2003.

[25] R. Lin, S. Kraus, Can automated agents proficiently negotiate with humans? Communications of the ACM 53 (1) (2010) 78–88.

[26] R. Lin, S. Kraus, D. Tykhonov, K. Hindriks, C.M. Jonker, Supporting the design of general automated negotiators, in: Proc. of the Second International Workshop on Agent-Based Complex Automated Negotiations, 2009.

[27] R. Lin, Y. Oshrat, S. Kraus, Investigating the benefits of automated negotiations in enhancing people's negotiation skills, in: Proc. of the Eighth International Conference on Autonomous Agents and Multiagent Systems, 2009, pp. 345–352.

[28] I. Marsa-Maestre, M.A. Lopez-Carmona, J.R. Velasco, T. Ito, M. Klein, K. Fujita, Balancing utility and deal probability for auction-based negotiations in highly nonlinear utility spaces, in: Proc. of the Twenty-First International Joint Conference on Artificial Intelligence, 2009, pp. 214–219.

[29] Y. Mazliah, Y. Gal, Coordination in multi-player human–computer groups, in: Proc. of the First Human–Agent Interaction Design and Models Workshop (HAIDM), 2005.

[30] T.D. Nguyen, N.R. Jennings, Coordinating multiple concurrent negotiations, in: Proc. of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, vol. 3, 2004, pp. 1064–1071.

[31] M.J. Osborne, A. Rubinstein, Bargaining and Markets, Academic Press, San Diego, CA, 1990.

[32] M.J. Osborne, A. Rubinstein, A Course in Game Theory, MIT Press, Cambridge, MA, 1994.

[33] J.W. Pratt, Risk aversion in the small and in the large, Econometrica 32 (1964) 122–136.

[34] H. Raiffa, The Art and Science of Negotiation, Harvard University Press, Cambridge, MA, 1982.

[35] V. Robu, D.J.A. Somefun, J.A. La Poutré, Modeling complex multi-issue negotiations using utility graphs, in: Proc. of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, 2005, pp. 280–287.

[36] A. Rogers, R. Dash, S. Ramchurn, P. Vytelingum, N.R. Jennings, Coordinating team players within a noisy iterated prisoner's dilemma tournament, Theoretical Computer Science 377 (1–3) (2007) 243–259.

[37] A. Rubinstein, Perfect equilibrium in a bargaining model, Econometrica 50 (1) (1982) 97–109.

[38] B. Shi, E.H. Gerding, P. Vytelingum, N.R. Jennings, An equilibrium analysis of market selection strategies and fee strategies in competing double auction marketplaces, Autonomous Agents and Multi-Agent Systems (2012), in press, http://dx.doi.org/10.1007/s10458-011-9190-5.

[39] J. Smith, Evolution and the Theory of Games, Cambridge University Press, Cambridge, United Kingdom, 1982.

[40] I. Sofer, D. Sarne, A. Hassidim, Negotiation in exploration-based environment, in: Proc. of the Twenty-Sixth AAAI Conference on Artificial Intelligence, 2012.

[41] A.M. Sykulski, A.C. Chapman, E.M. de Cote, N.R. Jennings, EA^2: The winning strategy for the inaugural lemonade stand game tournament, in: Proc. of the Nineteenth European Conference on Artificial Intelligence, Lisbon, Portugal, 2010, pp. 209–214.

[42] I. Vetsikas, N. Jennings, Outperforming the competition in multi-unit sealed bid auctions, in: Proc. of the Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems, 2007, pp. 702–709.

[43] P. Vytelingum, D. Cliff, N.R. Jennings, Strategic bidding in continuous double auctions, Artificial Intelligence 172 (14) (2008) 1700–1729.

[44] M.P. Wellman, J. Estelle, S. Singh, Y. Vorbeychik, C. Kiekintveld, V. Soni, Strategic interactions in a supply chain game, Computational Intelligence 21 (1) (2005) 1–26.

[45] M.P. Wellman, P.R. Wurman, K. O'Malley, R. Bangera, S.d. Lin, D. Reeves, W.E. Walsh, Designing the market game for a trading agent competition, IEEE Internet Computing 5 (2) (2001) 43–51.

[46] C.R. Williams, V. Robu, E.H. Gerding, N.R. Jennings, 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, 2011, pp. 432–438.

[47] C.R. Williams, V. Robu, E.H. Gerding, N.R. Jennings, Negotiating concurrently with unknown opponents in complex, real-time domains, in: Proc. of the Twentieth European Conference on Artificial Intelligence, 2012, pp. 834–839.

[48] H. Young, The evolution of conventions, Econometrica 61 (1) (1993) 57–84.