The Fourth
Automated Negotiating
Agent Competition
Outline

• Introduction (for AAMAS audience)
• Agent presentations
• Results
• Challenges and next steps for ANAC/GENIUS
Background

• Started in 2010, as a joint project of the universities of Delft (group of Prof. Catholijn Jonker, Dr. Koen Hindriks, Dr. Dmytro Tykhonov, Tim Baarslag) and Bar-Ilan (Prof. Sarit Kraus, Dr. Raz Lin)

• In 2011, organized by Nagoya Institute of Technology (Prof. Takayuki Ito, Dr. Katsuhide Fujita)

• In 2012, organized by University of Southampton (Colin Williams, Dr. Valentin Robu, Dr. Enrico Gerding, Prof. Nick Jennings)

• In 2013, organized by Ben Gurion University of the Negev (Litan Ilany, Dr. Yaakov (Kobi) Gal)

• Aim: to provide a platform to compare and benchmark different state-of-the-art heuristics developed for automated, bilateral negotiation
Competition Setup

• Bi-lateral Negotiation
• Alternating Offers Protocol
• Real-time, 60 second Deadline
• Discounting Factor
• Reservation value
Domains and Preferences

• Each domain consists of pair of preference profiles.
• Each preference profile specified as linearly additive utility function.

• Between 1 and 7 issues.
• Domains with between 3 and 56,700 possible outcomes.
Example Domain

• Property Rental
  - Rent Price per month
    • $1,800, $2,000, $2,400, $2,700
  - Number of Payments
    • 1, 2, 3
  - Advance Payment
    • 0.5%, 1%, 2%, 2.5%
  - Contract Period
    • 3 months, 6 months, 9 months, 12 months
## Example Preferences

<table>
<thead>
<tr>
<th>Rent Price per month</th>
<th>Landlord</th>
<th>Tenant</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>weight</em></td>
<td>0.350</td>
<td>0.353</td>
</tr>
<tr>
<td>$1,800</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>$2,000</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>$2,400</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>$2,700</td>
<td>80</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Payments</th>
<th>Landlord</th>
<th>Tenant</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>weight</em></td>
<td>0.2</td>
<td>0.129</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>
Previous Competitions

• 2010
  – 7 Entries

• 2011
  – 18 Entries (6 institutions)

• 2012
  – 17 Entries (8 institutions)

• 2013
  – 19 Entries (7 institutions)
New Feature

• Learning between negotiations
  – Ability to save information during and after negotiation session, and load it at the beginning of new session on the same domain and profile.
  – Can model the opponent’s profile, but cannot model the opponent itself.
Participants

• 19 Teams Entered
• 8 Institutions from 3 countries
  – Delft University of Technology, The Netherlands
  – Maastricht University, The Netherlands
  – Nagoya Institute of Technology, Japan
  – Shizuoka University, Japan
  – Tokyo University of Agriculture and Technology, Japan
  – Bar Ilan University, Israel
  – Ben Gurion University of the Negev, Israel
Preliminary round
Preliminary round

• Negotiations carried out for every combination of:
  – 19 Agents
  – 18 Opponents
  – 11 Domains (randomly selected from submissions)
• Each repeated 10 times to establish statistical significance and to allow learning.
  – Every pair played 20 times in each domain, once for each set of profiles
• Total of 75,240 negotiations.
<table>
<thead>
<tr>
<th>Position</th>
<th>Agent</th>
<th>Rank</th>
<th>Mean</th>
<th>Variance (Per run)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AgentKF</td>
<td>1</td>
<td>0.562</td>
<td>0.00019</td>
</tr>
<tr>
<td>2</td>
<td>TheFawkes / Agent Slinkhard</td>
<td>2-3</td>
<td>0.522</td>
<td>0.00132</td>
</tr>
<tr>
<td>3</td>
<td>TMFAgent</td>
<td>2-4</td>
<td>0.516</td>
<td>0.00163</td>
</tr>
<tr>
<td>4</td>
<td>MetaAgent</td>
<td>3-4</td>
<td>0.495</td>
<td>0.00252</td>
</tr>
<tr>
<td>5</td>
<td>G-Agent</td>
<td>5-8</td>
<td>0.457</td>
<td>0.00241</td>
</tr>
<tr>
<td>6</td>
<td>InoxAgent</td>
<td>5-8</td>
<td>0.455</td>
<td>0.00235</td>
</tr>
<tr>
<td>7</td>
<td>SlavaAgent</td>
<td>5-11</td>
<td>0.447</td>
<td>0.00018</td>
</tr>
<tr>
<td>8</td>
<td>VAStockMarketAgent</td>
<td>5-11</td>
<td>0.446</td>
<td>0.00520</td>
</tr>
<tr>
<td>9</td>
<td>RoOAgent</td>
<td>7-11</td>
<td>0.432</td>
<td>0.00313</td>
</tr>
<tr>
<td>10</td>
<td>AgentTalex</td>
<td>7-11</td>
<td>0.431</td>
<td>0.00285</td>
</tr>
<tr>
<td>11</td>
<td>AgentMRK2</td>
<td>7-11</td>
<td>0.430</td>
<td>0.00344</td>
</tr>
<tr>
<td>12</td>
<td>Elizabeth</td>
<td>12-14</td>
<td>0.387</td>
<td>0.00443</td>
</tr>
<tr>
<td>13</td>
<td>ReuthLiron</td>
<td>12-15</td>
<td>0.374</td>
<td>0.00416</td>
</tr>
<tr>
<td>14</td>
<td>BOAconstrictorAgent</td>
<td>12-15</td>
<td>0.373</td>
<td>0.00141</td>
</tr>
<tr>
<td>15</td>
<td>Pelican</td>
<td>13-18</td>
<td>0.359</td>
<td>0.00434</td>
</tr>
<tr>
<td>16</td>
<td>Oriel_Einat.Agent</td>
<td>15-18</td>
<td>0.350</td>
<td>0.00534</td>
</tr>
<tr>
<td>17</td>
<td>MasterQiao</td>
<td>15-18</td>
<td>0.345</td>
<td>0.00214</td>
</tr>
<tr>
<td>18</td>
<td>Eagent</td>
<td>15-18</td>
<td>0.338</td>
<td>0.00707</td>
</tr>
<tr>
<td>19</td>
<td>ClearAgent</td>
<td>19</td>
<td>0.315</td>
<td>0.00109</td>
</tr>
</tbody>
</table>
Preliminary round results

0.31
0.34
0.35
0.35
0.36
0.37
0.37
0.37
0.39
0.43
0.43
0.43
0.43
0.45
0.45
0.45
0.46
0.49
0.52
0.52
0.56
Agent Presentations

ANAC 2013
Agent KF

Katsuhide Fujita
Tokyo University of Agriculture and Technology, JAPAN
katfuji@cc.tuat.ac.jp
Opponent Modeling

- Average and Variance of opponent’s bids in the utility function of our agent
  \[
  \mu(t) > \mu_h : \text{Uncooperative} \\
  \mu(t) = \mu_h : \text{Neutral} \\
  \mu(t) < \mu_h : \text{Cooperative}
  \]

- \(\nu > \nu_h\) : Seeking Step
- \(\nu \leq \nu_h\) : Compromising Step

\(\mu(t), \nu(t)\) : Average and Variance of opponent’s bids in time \(t\)
\(\mu_h, \nu_h\) : Average and Variance of opponent’s bids in the all history

<table>
<thead>
<tr>
<th></th>
<th>Uncooperative</th>
<th>Neutral</th>
<th>Cooperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compromising</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Responses of opponent’s bids

- (Utility of opponent’s bids) > Bid(t): Accept the opponent’s bid
- Opponent’s Model is “Cooperative” and “Compromising Step” Reducing the *alpha* of Bid(t) (= more cooperative strategy)
Proposals of our bids

• Our agent searches the bids which utility is almost same as $Bid(t)$ in the previous slide
• For finding the pareto-optimal bids, we select the most similar bid between the opponent’s first bid and our alternatives
  – Opponent’s first bid = The best bid for the opponent
  – Similarity Measure:
    
    $$\sum_{i=1}^{N} w_i \text{Bool}(b_t, b_0)$$

    $\text{Bool}(b_t, b_0) : \text{if}(b_t == b_0) \text{ then return 1 else return 0}$
GAgent

Yuta Iwama
Nagoya Institute of Technology
Takayuki Ito Laboratory
iwama.yuta@itolab.nitech.ac.jp
Outline of GAgent

• We use two functions, **sigmoid function** and **opponent modeling** function to determine a threshold value.

\[
y = \frac{1 - \text{Low}}{1 + e^{ax}} + \text{Low}
\]

- **Low**: our max concession
- **a**: constant
- **x**: time
- **y**: threshold

• The threshold, **y**, tends to get closer to agreement while time passes
Proposals/Responses

- If opponents don’t concede, our agents also not concede according to opponent modeling function
- We tend to agree about passes over time according to sigmoid function
- Since use DOUS, we can analyze movement of opponent in domain scale
- Since this threshold is very high value versatility, we can use this threshold when accepting from opponent’s bid

Sigmoid function

Opponent modeling function
Opponent modeling function analyzes opponent’s concession condition without save-to-desk function

**Opponent modeling function**

DOUS: Distance of the Opponent in my Utility Space (DOUS)

\[ \text{DOUS} = \text{ourMaximumUtility} - \text{OpponentFirstBid} \]

1. Calculate DOUS
2. Calculate the mean and variance by opponent’s concession condition in DOUS
3. Use value that multiplied by the mean and variance as a threshold value
ANAC 2013 agents by Delft University of Technology

Tim Baarslag, Kees Boon, Mariana Branco, Alex Dirkzwager, Madalin Dumitru-Guzu, Mark Hendrikx, Koen Hindriks, Catholijn Jonker, Vincent Koeman, Joris van den Oever, Catalin Stanculescu, and Ruben van Zessen

Interactive Intelligence Group
Department of Intelligent Systems
Faculty Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
For our agents, we use a framework that distinguishes different negotiation components:
There are two important questions we can ask about the opponent:

• **Preference modeling**: “what does the opponent want?”

• **Strategy prediction**: “what will the opponent do?”
Opponent Modeling

Frequency Model
J. Hao et al., CUHKAgent, 2012.

Bayesian Models
K. Hindriks et al., 2008. and
C. Williams et al., 2012.
Two very promising techniques:

Gaussian Process Regression

Discrete Wavelet Prediction

C. Williams et al., 2011.

S. Chen et al., 2012.
Popular methods:

• Time dependent concessions

• Tit for tat

• Aiming for certain optima (Nash/Kalai)
Acceptance Strategies

$AC_{\text{next}}$  
Accept when the opponent’s bid is better than our upcoming bid.

$AC_{\text{time}}(T)$  
Accept based on the remaining time.

• Optimal stopping  
Compute the optimal time to accept using stochastic decision techniques.

$T. \text{Baarslag et al., 2013.}$
Applying the BOA Framework

Combining components

Frequency model

Bayesian learning

No opponent model

Tit for Tat strategy

Conceding strategy

Accept the best offer so far

Accept very good bids only

Accept above a fixed threshold

Accept as late as possible

Accept early

Accept at the last moment

Never accept

Accept at the last moment

Accept as late as possible
Applying the BOA Framework

<table>
<thead>
<tr>
<th>Agent</th>
<th>Bidding Strategy</th>
<th>Opponent Model</th>
<th>Acceptance Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BOA Constrictor</strong></td>
<td>Time dependent, aims for Kalai</td>
<td>Frequency model</td>
<td>Optimal stopping with GPR</td>
</tr>
<tr>
<td><strong>Inox agent</strong></td>
<td>Time dependent + TFT, aims for Kalai</td>
<td>Frequency model</td>
<td>$A_{C_{\text{next}}}$</td>
</tr>
</tbody>
</table>
| **Slinkhard**        | Time dependent, based on strategy prediction         | Discrete Wavelet Prediction + Frequency model | $A_{C_{\text{time}}}(T)$ }
Inox Agent

R. van Zessen & M. Branco
Delft University of Technology
Different components for:

- **One-issue** (modelling is useless) – conceding strategy approach to the median utility;

- **Multi-issue** – more complex, explained in following slides.
Bidding Strategy

- Reset for bad opponent offers (no improvement from last 4 offers);
- Aims for Kalai-Smorodinsky point (fairest outcome).
- We used round estimation instead of using the normalized time, to be more flexible when dealing with slower agents or different times;
Opponent Model

- Simple frequency model; only updates weights when the bid change often enough and value calculation is scaled:
Acceptance Strategy

- Acceptance strategy breaks when reservation value seems to be better, or accepts with $AC_{\text{worst}}$ (with a small margin from discount factor).

Saving & Loading

- Update of the “median utility” used as minimum in bidding and acceptance strategy;
- Depends on the outcome of the negotiation: calculated using the average of all outcomes;
- In case of no outcome uses the real median utility.

Notice that: Both bidding and acceptance strategy use the median utility from our and the opponent first offer as a minimum.
Meta-Agent

Litan Ilany

Ben Gurion University of the Negev
Not a regular agent...

Not a single negotiation strategy code line has written....
Algorithm selection method using only ANAC 2012’s algorithms.

1. Feature selection
2. Machine learning
3. Selecting best algorithm
4. Learning and improvement
1. Feature Selection

- Used to describe a negotiation environment (domain & profile):
  - Domain features:
    - size
    - Number of issues
    - Average utility of all bids
    - ...
  - Profile features:
    - Discount factor
    - Opponent first bid’s utility
    - Average utility of relevant bids
    - ...
2. Machine learning

• Dependent variable:
  – Performance: Distance of the utility from the average of all agents.
  – Boolean variable: win/loose

• Learning methods:
  – Linear regression
  – Logistic regression
  – CART (selected for ANAC 2013)
  – Neural networks
"3. The “Meta Agent

• Given a new domain
  − Compare all performance predictions of all agents on the domain
  − Use the agent with the highest predicted utility.
4. Learning Process

- Using multi armed bandit approach (UCB)

\[
\arg\max_a \left( \hat{U}_a + \sqrt{\frac{2 \ln(N)}{n_a}} \right)
\]

- Initialize with 5 “instances” of the predicted utility per agent
Slava Agent

Moshe Hazoom, Guy Dubrovski, Slava Bronfman
Bar-Ilan University
Proposals

• Split into two modes: Exploration and Exploitation.

• At exploration mode - find the bid which maximizes the agent’s utility.
  – with probability of 0.5, offer it to the opponent
  – with probability of 0.5, offer to him a bid that is “good enough” for us randomly (i.e. one that has utility more than 0.95).

• At exploitation part, offer always the best bid for us. Accept:
  – If the opponent offers a bid that has utility bigger than 0.7 for us, and has value higher or equals to all opponent its previous offers.
  – Otherwise, again offer the best bid for us.

➢ Exploration defined to be 95% of the time and exploitation defined to be on the other 5% of the time.
• We define a variable called GOOD_ENOUGH.Utility(~0.8) that stands for a threshold that if the opponent agent offers us a bid that has a utility higher than it, accept immediately.
Opponent modeling

• with probability of 0.5 we offer some random bid that satisfies us.

• With this kind of randomness, and not a strict policy, we make the opponent agent offers us different bids than before (hopefully)
  − That way we can maximize the bid with highest utility for us that the opponent agent offer to us.
  − We encountered couple of negotiation sessions on which this randomness made the opponent agent offer a bid with utility that is more than GOOD_ENOUGH_UTILITY and we accepted it and won.
Save-to-disk

• Save at the end of the negotiation the best bid with highest utility we found from the domain.
• At the beginning of the session, we load this value from the disk, and initialize the bid with the highest utility for us to be that value.
• That way, we can learn about the specific domain from session to session and improve agent’s performance.
TMF-Agent

Yedidya Bar-Zev
Ben Gurion University of the Negev
Opponent Modeling

• Estimating the utility of each value for all issues
  – When opponent offers some value in an issue, it added to the relevant “value bucket” with the current offering time (descending = 1-t)
  – Earlier offers gets higher scores
  – Normalize values’ utilities to range [0,1]

• Equal weights was assumed for all issues
  – Allows to convert a offer to the opponent estimated utility

<table>
<thead>
<tr>
<th>Value</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>0.41</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
</tr>
</tbody>
</table>
Define $\alpha$ as hardness parameter that depends on time and discount factor.

- The agent is hard-headed as $\alpha$ is higher
Bidding + Acceptance model

• Select the bid $B$ that maximize:

$$
\alpha \times B.\text{utility} + (1 - \alpha) \times B.\text{OpponentEstimateUtility}
$$

- If the opponent bid's utility is higher than $B$, accept.
- If the reservation value is lower than $B$, end negotiation.

• Timeout avoidance:
  - Estimating negotiation “round trip time” between two offers
  - When arrives to last two actions before timeout,
    • offer back the best opponent offer that suggested so-far
    • If the opponent rejects (non-rational), accept his final offer (only if it’s better than the Reservation-Value)
Learning Function

• After ending a negotiation:
  − Enlarge the weight of the chosen values in the *Opponent model* and normalize the values’ utilities
  − At the start of the next negotiation the agent uploads *Opponent model* (doesn’t model from scratch)
Final round
Final Round

• 7 Agents
• 18 Domains
  • (12 submitted this year, 6 from 2012)
  • Total of 15120 negotiations
Final Round

Domain Size

Discount factor & Reservation value

Discount Factor
Reservation Value
## Prizes

<table>
<thead>
<tr>
<th>Prize</th>
<th>Prize Money</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Place</td>
<td>$500</td>
</tr>
<tr>
<td>2nd Place</td>
<td>$400</td>
</tr>
<tr>
<td>3rd Place</td>
<td>$300</td>
</tr>
<tr>
<td>Most Social Agent</td>
<td>$150</td>
</tr>
<tr>
<td>Best Learning Agent</td>
<td>$150</td>
</tr>
</tbody>
</table>

With thanks to our sponsors:
- Prof. Dr. Catholijn Jonker
- Prof. Dr. Sarit Kraus
- Prof. Dr. Takayuki Ito / Makoto Lab., Inc.
## Overall Ranking

<table>
<thead>
<tr>
<th>Variance</th>
<th>Mean</th>
<th>Rank</th>
<th>Agent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000011</td>
<td>0.606434</td>
<td>1</td>
<td>???</td>
</tr>
<tr>
<td>0.000083</td>
<td>0.600209</td>
<td>2</td>
<td>???</td>
</tr>
<tr>
<td>0.000012</td>
<td>0.583094</td>
<td>3</td>
<td>???</td>
</tr>
<tr>
<td>0.000069</td>
<td>0.568215</td>
<td>4-5</td>
<td>???</td>
</tr>
<tr>
<td>0.000055</td>
<td>0.564908</td>
<td>4-5</td>
<td>???</td>
</tr>
<tr>
<td>0.000147</td>
<td>0.534514</td>
<td>6</td>
<td>???</td>
</tr>
<tr>
<td>0.000023</td>
<td>0.484973</td>
<td>7</td>
<td>???</td>
</tr>
</tbody>
</table>
Overall Ranking
3rd place

TMF Agent

Automated Negotiating Agent Competition
2013
First and Second place

• No significant difference in score between first and second place.
• Awarded second place to the agent with the higher variance among the two.
Meta Agent
The Fawkes
Overall Ranking

ANAC 2013

Automated Negotiating Agent Competition 2013
Comparison with ANAC 12

The Fawkes
Most Social Agent

• Agent which maximises the sum of its own utility and its opponent’s.
Best Learning Agent

7 measures for improvement was used (Thanks to Tim):

- % of Pareto bids during negotiation
- Exploration rate (% of new bids made out of the total bids in the negotiation session)
- Undiscounted utility distance from Pareto frontier
- Undiscounted utility distance from Kalai-Smorodinsky solution
- Undiscounted utility distance from Nash solution
- Average utility
- Time of agreement
Best Learning Agent

- Calculated the average slope for each of the measures (normalized for maximum)

- Best learning agent was the agent which achieved the highest average slope over all negotiation sessions.
Best Learning Agent

- (6/7)
- (6/7)
- (3/7)
- (2/7)
- (4/7)
- (1/7)
- (2/7)
Learning comparison with 2012
Challenges and next steps

• Learning of other opponents?
• Putting people in the mix?
• Challenges
  – How to breakthrough to other researchers?
  – Difficult to run competitions.
  – GENIUS GUI?