Challenges and Main Results of the Automated Negotiating Agents Competition (ANAC) 2019

Reyhan Aydoğan[1,2(✉)], Tim Baarslag[3,4], Katsuhide Fujita[5,9], Johnathan Mell[6], Jonathan Gratch[6], Dave de Jonge[7], Yasser Mohammad[8,9], Shinji Nakadai[8,9], Satoshi Morinaga[8,9], Hirotaka Osawa[10], Claus Aranha[10], and Catholijn M. Jonker[2]

[1] Özyeğin University, Istanbul, Turkey
reyhan.aydogan@ozyegin.edu.tr
[2] Delft University of Technology, Delft, The Netherlands
[3] Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
[4] Utrecht University, Utrecht, The Netherlands
[5] Tokyo University of Agriculture and Technology, Fuchu, Japan
[6] USC Institute for Creative Technologies, Los Angeles, USA
[7] IIIA-CSIC, Bellaterra, Spain
[8] NEC Corporation, Minato City, Japan
[9] National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan
[10] University of Tsukuba, Tsukuba, Japan

Abstract. The Automated Negotiating Agents Competition (ANAC) is a yearly-organized international contest in which participants from all over the world develop intelligent negotiating agents for a variety of negotiation problems. To facilitate the research on agent-based negotiation, the organizers introduce new research challenges every year. ANAC 2019 posed five negotiation challenges: automated negotiation with partial preferences, repeated human-agent negotiation, negotiation in supply-chain management, negotiating in the strategic game of Diplomacy, and in the Werewolf game. This paper introduces the challenges and discusses the main findings and lessons learnt per league.

1 Introduction

Negotiation has become a well-established research field within the area of Artificial Intelligence and multi-agent systems. The research has focused on formalization of negotiation process (i.e., domain and preference representation, and protocols) and the design of intelligent negotiating agents (i.e., bidding strategies, opponent models, and acceptance strategies) in order to automate this complex process. Automated negotiation dates back to the 1980's when e-commerce took flight, see e.g., [29,37]. The field was formalized in the 1990's (see e.g., [34,36,38]). Over the years negotiating agents have been developed for automated negotiation, human-agent negotiation, and negotiation support.

N. Bassiliades et al. (Eds.): EUMAS 2020/AT 2020, LNAI 12520, pp. 366–381, 2020.
https://doi.org/10.1007/978-3-030-66412-1_23

In automated negotiation all negotiation parties are automated agents, while in human-agent negotiation, some of them are human [31]. Negotiation support agents form a team with one or more humans to play together as one negotiation party in any kind of negotiation (automated, human-human, or human-agent) [18].

With the growing number of proposed negotiation agents, the need for comparison and rigorous evaluation of the quality of the negotiating agents increases as well. This led to formal evaluation metrics [12,14,21,23], the open-source negotiation platform GENIUS to enable benchmarking [20], and in 2010 it initiated the annual ANAC (Automated Negotiation Agents Competition) [5].

The competition turned into a mechanism for the field to organize itself as the researchers use the yearly meetings to jointly set the research agenda. Over the years, the negotiation problems studied in the context of ANAC [3] span bilateral [5], non-linear [2], multilateral [11], and repeated [1] negotiations. As an added advantage, by now GENIUS holds a host of agents, negotiation domains and preference profiles.

Since 2017, ANAC has added two new leagues: a Human-Agent league and a league for the game Diplomacy. In the Human-Agent league, which is based on the IAGO framework [24], agents exchange partial offers, preferential information and through emoji's some emotional information with their human opponents, see e.g., [26]. In the game Diplomacy the agents have to negotiate on the basis of heuristics, as there is no explicit utility function available [15]. In 2019, two more leagues were added: the Supply Chain Management league (SCM) [28] and the Werewolf League [30]. The SCM league allows researchers to study negotiation and partner selection in a recurring setting of trade. In the Werewolf game the essence of negotiation studied is that agents need to assess the utility functions of the other players and convince others to play a successful voting strategy. The challenges for the ANAC 2019 competition were as follows (organised per league):

- **Automated Negotiation Main League: preference uncertainty.** Human negotiators do not necessarily know their own utility function explicitly, and there are practical limits on the preference elicitation process. Therefore, the challenge is to design an agent that can do bilateral negotiations receiving only partial qualitative preference information.
- **Human-Agent League: building cooperation.** The challenge is to establish a cooperative relationship with the human participant in repeated negotiations with the same human opponent. Successful agent strategies capture human behavior. While an aggressive strategy in the first negotiation may prove effective, it could have such a backfire effect by the last negotiation that it is not the right choice overall.
- **Diplomacy: beat the basic agent.** Like last year, the challenge was to beat the standard agent provided by the BANDANA framework. No participating agent managed this in 2018.
- **Supply Chain Management: recurrent chain negotiations.** The challenges are to decide on their own utility function, when and with whom to

negotiate and how to negotiate in a supply chain in order to maximize their overall profit.
- **The Werewolf game.** The challenge for the agents is to estimate possible allies and deceivers (estimated utility), to communicate strategy and information to other agents, and to negotiate a voting pattern that is beneficial to one's own team.

This paper consists of sections for each league in which the challenges and main competition results are discussed. The last section presents some of the upcoming challenges.

2 Automated Negotiation Main League

There are still many open challenges for automated negotiation [6,7], such as strategy learning, opponent recognition, domain learning, preference elicitation and reasoning. The Automated Negotiation league in 2019, informally known as the GENIUS league, focused on negotiating agents that receive partial preference information. This challenge is part of the larger research problem of domain learning and preference elicitation. The motivating idea is that when a negotiating agent represents a user in a negotiation, it cannot know exactly what the user wants due to practical limits on the preference elicitation process [4].

For ANAC 2019, the preferences of the agent were given in the form of a ranking of a limited number of possible agreements ω_i; i.e. $\omega_1 \leq \cdots \leq \omega_d$. The rankings were generated randomly from existing negotiation scenarios in which full utility information was available from a standard additive utility function u. Intuitively, the number of rankings d that the agent receives is inversely correlated to the preference uncertainty of the agent. The agent has to negotiate based on these ordinal preference information, and if it manages to reach a certain outcome ω^*, then the score the agent receives for this agreement is based on the original utility function, i.e., $u(\omega^*)$. An overview of this procedure is presented in Fig. 1. In short, the agent receives ordinal information only, but is evaluated based on the underlying cardinal ground-truth.

Table 1 shows the average individual utility and the average product of utilities gained by all participants in a tournament in which each agent negotiated with all other agents five times for each negotiation scenario. When evaluating on individual utility, `AgentGG` won the competition with an average of 0.76, the agents `KakeSoba` and `SAGA` were awarded second and third place. When evaluating on fairness (i.e. the product of the utilities of the negotiated agreements), `winkyAgent` won the competition with an average utility of 0.56, and agents `FSEGA2019` and `AgentGP` were awarded second and third place respectively.

As intended, the key to win this league is for agents to predict both their own and their opponent's utility accurately from uncertain information. The top agents were able to obtain high individual utilities even under high preference certainty, using a variety of preference estimation techniques. In estimating the preferences, the top ranking agents used techniques such as batch gradient

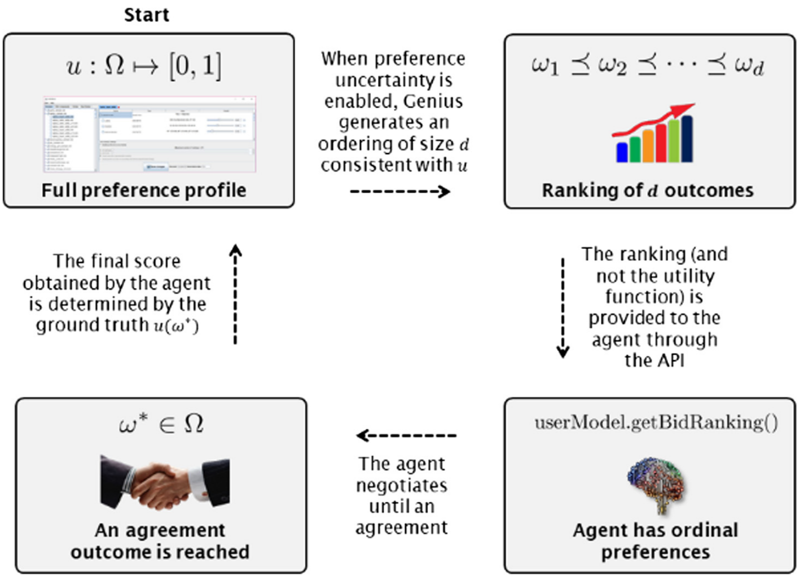

Fig. 1. Uncertainty challenge in ANAC 2019

descent (e.g. `winkyAgent`), genetic algorithms using spearman's rank correlation (e.g. `SAGA`), and statistical frequency modelling (e.g. `AgentGG`).

The performance of the top ranking agents suggests that it is possible to reconstruct enough of the true utility function based on partial information about the ranking of a number of bids. The next question is of course, *how much* partial information is required. Therefore, the ANAC community decided to formulate the next challenge, which incentivizes agent designers to use as little information as possible to still get good performance: next year, the agents initially will receive very sparse preference data, and will be allowed to ask for more preference information against an elicitation cost.

3 Human-Agent League

The human-agent league focuses on the myriad social effects present in mixed human-agent interactions. Indeed, understanding how humans negotiate has been a key question in business and psychological literature for many years—it is a complex social task [19,25,32,33]. But as automated agents are used more frequently in real-life applications (e.g., chatbots), we should design agents that are capable of interacting with humans in social settings. As human-agent negotiation is fundamentally different from agent-agent negotiation, the Human-Agent League (HAL) was added in 2017 to promote research into this promising area.

HAL utilizes the IAGO Negotiation platform, which was proposed and designed by Mell et al. [24]. IAGO provides a front-facing GUI for the human participants. This feature allows subjects to be recruited using online platforms, such as Amazon's Mechanical Turk (MTurk). Additionally, IAGO provides the features necessary for simulating the characteristics of human negotiation. These include an expanded set of channels for communication between both sides of negotiation, such as by sending text, expressing preferences, and transmitting emotions. Text is transmitted through a set of pre-selected utterances, and emotions are transmitted by selecting from a variety of prototypical "emojis" within the interface. Furthermore, IAGO allows "partial offers" (i.e., offers not containing values for all negotiation issues) and implements a flexible, human-inspired protocol: few enforcement mechanisms for incomplete deals, and no explicit turn-taking.

These features of IAGO mean that it provides a platform to address the basic features that intelligent agents require to negotiate with humans. It provides information that allows for human-opponent modeling, and for agents to pursue more complex strategies that require specific features (such as partial offers), and the information from the multiple channels for communication.

Results from the first and second human-agent leagues (see [25,26]) show that while certain strategies may be effective in the short term, there is a trade-off between agent likeability and agent success. To further examine this, the structure of the repeated negotiations were changed.

In this year's competition, three back-to-back negotiations were conducted. Similar to previous competitions, the negotiation involved a 4-issue bargaining task. Each issue could take from 4 to 8 items, e.g., offering 4 to 8 bananas. Each of the three negotiations took up to 7 min, and a BATNA was available for those who could not reach an agreement. Each agent negotiated against at least 25 human participants using Amazon's Mechanical Turk subject pool, and those participants were subject to *attention checks and filtering*. All participants were US-residing, and English-speaking. Per standard practice, incentives were scaled with performance, so participants were encouraged to do well. Data was collected on demographics, performance metrics, and subject-reported likeability measures of the agent. All procedures were approved by University of Southern California's Institutional Review Board, including the informed consent process.

In contrast to previous years, the negotiations were not identical in structure. Instead, while there were integrative opportunities to "grow the pie" within each negotiation, there was a larger, cross-negotiation possibility to find integrative potential between negotiations #1 and #3. This higher performance opportunity is reflected in negotiation #3, where agents generally have more points due to structural differences.

Regardless of this effect, we did find a variety of performance differences across the submitted agents. In particular, we had two standout agents in terms of performance: agents `Dona` and `Draft` (See Fig. 2). The `Draft Agent` was submitted by Bohan Xu, Shadow Pritchard, James Hale, & Sandip Sen from the University of Tulsa, while the `Dona Agent` was submitted by Eden Shalom Erez, Inon Zuckerman, and Galit Haim of Ariel University and The College of Management Academic Studies. These agents took unique approaches to the challenges of negotiation by making agents be guided by the "meta-rules" of negotiation.

Table 1. Results of automated negotiation league

	Individual utility		Nash product	
Agent	♯	Mean	♯	Mean
AgentGG	**1**	**0.7574**	5	0.5398
AgentGP	10	0.6948	3	0.546
AgentLarry	15	0.5984	11	0.5082
AgentPeraltaV2	18	0.5528	17	0.4012
AuthenticAgent	20	0.3882	20	0.1947
dandikAgent	9	0.6995	13	0.4628
EAgent	14	0.6553	16	0.4286
FSEGA2019	8	0.7002	2	0.5495
GaravelAgent	13	0.6571	15	0.4365
Gravity	17	0.5782	18	0.361
Group1_BOA	11	0.6927	6	0.5392
HardDealer	5	0.7245	9	0.5172
IBasic	21	0.32	21	0.136
KAgent	16	0.5943	14	0.4569
KakeSoba	2	0.7433	7	0.5259
MINF	4	0.728	10	0.5133
SACRA	19	0.4901	19	0.3166
SAGA	3	0.7315	4	0.5423
SolverAgent	6	0.7126	8	0.5257
TheNewDeal	12	0.6863	12	0.4873
winkyAgent	7	0.7093	**1**	**0.5625**

`Dona agent` customized the interface to instruct the user to answer questions using the emoji buttons. `Draft agent` enforced strict protocols for the humans to follow; it required human participants to describe their preferences in a set order. The success of these agents speaks to the importance of setting a clear protocol in negotiations that cannot be manipulated by the agents. Furthermore, we learned that humans are inclined to adhere to changes in protocol made by their automated counterparts.

For the next Human-Agent competition, we have decided to adapt the task beyond the 2019 competition. Firstly, while the novelty of the agents that modified the interface led to some unexpected yet interesting results, we will be returning to a competition in which the interface protocols are set at the beginning of the interaction. The lessons learned from this competition have led to new insights in UI design which have been integrated into the IAGO platform. Secondly, we will be allowing the human users to set their own preferences in the negotiation. This is both more realistic to the real world, and will also ensure that the agent designers have to contend with a set of potential negotiation structures. We hope that this next competition will continue to push the envelope in designing more realistic and useful social agents.

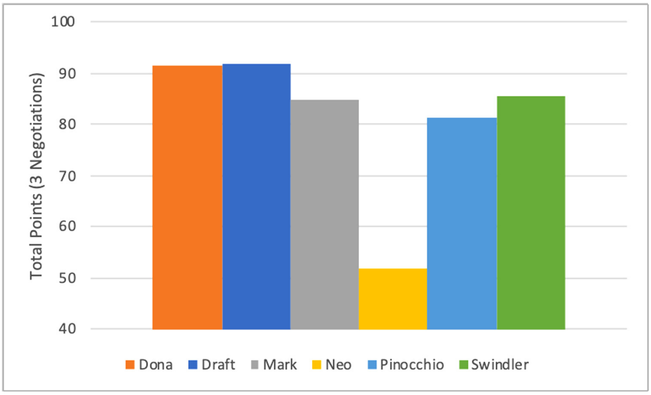

Fig. 2. Total agent score (summed over all negotiations)

4 The Diplomacy League

Diplomacy [9] is a deterministic board game for seven players, with no hidden information. It is designed such that players need to form coalitions and negotiate with each other. The interesting aspect of Diplomacy as a test case for Automated Negotiations is that there is no explicit formula to calculate utility values. Just as in games like Chess or Go, it is simply too complex to calculate such values exactly, so agents have to apply heuristics to estimate the values of their deals. Although Diplomacy has been under attention of the Automated Negotiations community for a long time [8,10,17,22,35], to date few successful negotiating Diplomacy players have been developed.

In the previous two editions of the ANAC Diplomacy League ANAC [15] none of the submitted agents was able to beat the challenge and outperform even a non-negotiating agent. Therefore, we decided to make the challenge slightly easier by making sure that negotiating agents were always assigned to 'Powers' that are known to work well together [15]. Other than this, the setup of the 2019 competition was kept practically identical to the previous years.

Participants had to implement a negotiation algorithm on top of the existing `D-Brane agent` [17], which by itself does not negotiate. Negotiations took place under the Unstructured Negotiation Protocol [16]. The competition consisted of two rounds. In Round 1, each agent only played against three copies of itself and three non-negotiating instances of `D-Brane`, while in Round 2, all submitted agents played against each other. In order to *beat the challenge*, an agent had to satisfy two criteria: it would have to outperform the non-negotiating `D-Branes` in Round 1, as well as beat all opponents in Round 2. In case no agent was able to beat the challenge (as in previous years) the *Backup Rule* would come into effect, which states that the agent that made the most proposals in Round 2 that were eventually accepted by other agents would be declared the winner.

The competition received five submissions. Unfortunately, none of them was able to outperform `D-Brane` in Round 1 (Table 2). This suggests that the agents were not cooperative enough to be able to strike a good deal even when their opponents are identical to themselves. On the plus side, in Round 2 we did see that for the first time in the history of the ANAC Diplomacy league one agent, namely `Oslo_A`, by Liora Zaidner *et al.*, was able to clearly outperform all other agents (Table 3). However, according to the rules of the competition it was `Saitama`, by Ryohei Kawata that was declared to be the winner, by virtue of the Backup Rule (Table 4).

Table 2. Results of round 1. None of the agents outperformed `D-Brane`

Agent	Score	Result
D-Brane	15.15	
Saitama	14.75	FAIL
Oslo_A	14.62	FAIL
DipBrain	14.56	FAIL
Biu3141	14.48	FAIL
BackstabAgent	14.47	FAIL

Table 3. Results of round 2. `Oslo_A` outperforms all other agents, and is the only one that outperforms the non-negotiating `D-Brane`. `Biu3141` could not participate in this round because it was too slow. `GamlBot` and `M@stermind` are submissions from previous years that were added to complete the field.

	Agent	Score
1	**Oslo_A**	6.68 ± 0.31
	D-Brane	5.56 ± 0.27
2	**DipBrain**	5.06 ± 0.24
3	**Saitama**	4.88 ± 0.23
	GamlBot	4.79 ± 0.21
4	**BackstabAgent**	4.20 ± 0.25
	M@sterMind	2.83 ± 0.17

Analyzing the source code of `Oslo_A` we noticed that its bidding strategy was surprisingly simple. At the beginning of each round it simply asks the underlying `D-Brane` module which moves it would make if no agreements are made. Then, for each of these moves, it asks the other players to support those moves. The authors

Table 4. Results according to the Backup Rule. This table displays the number of proposals proposed by each agent in Round 2 that were eventually accepted by the other agents. Saitama was declared the winner.

	Agent	Accepted proposals
1	**Saitama**	9091
2	**BackstabAgent**	6585
3	**Oslo_A**	4393
4	**DipBrain**	4373

also intended their agent to react to incoming proposals by either accepting them or making counter proposals, but due to bugs in the code, these components did not work. This also explains why Oslo_A failed in Round 1: if all players are copies of Oslo_A then no proposal is ever accepted, so no deals are made at all.

From this competition (as well as its predecessors in 2017 and 2018) we learn that it is still very hard to implement successful negotiation algorithms for domains as complex as Diplomacy. So far, no submission has been able to beat the challenge. Specifically, we make the following observations:

1. Most agents never make any proposals for any of the future turns. They only make proposals for the current turn.
2. Many agents do not outperform the non-negotiating D-Brane, or even score worse. This means that the deals they make often have a detrimental effect.
3. Many of the agents seem to have bugs in their code.

Regarding the first point, we remark that in Diplomacy it is essential to plan several steps ahead, because it does not often occur that two players can both directly benefit from cooperation. Players should therefore be willing to help another player while only expecting the favor to be returned at a later stage. However, most submissions do not seem to exhibit this behavior. The second point might explain the success of Oslo_A. After all, this agent only asked the other agents to support the orders that it was already going to make anyway. Therefore, its agreements can never have any detrimental effect. Furthermore, these observations suggest that Diplomacy is so complex that it requires a long time to design sophisticated agents for the game. This may explain that within the design time given in the ANAC competition, none of the participating teams managed to develop an agent that can beat D-Brane.

5 The Werewolf League

Werewolf, also known as *Mafia*, is a communication game where an uninformed majority team (the village) plays against an informed minority team (the werewolves). The goal of the game is to eliminate all players from the opposing team through a voting process: at each turn, the players must agree on one player to eliminate. This takes the shape of a discussion, followed by a vote, and a simple majority eliminates one player from the game.

From an Automated Negotiation point of view, the challenge for an agent in Werewolf is to successfully engage in coalition-building. In other words, the agent must identify other players in the game that share the same utility values as itself, at the same time that it must avoid deceitful agents. This requires the agent to communicate its own utility to the other agents, and engage in discussion to obtain the necessary information.

In the past six years, the AIWolf Project has proposed the Werewolf game as a benchmark for AI research [39] and organized four national competitions on the game. Compared to other AI benchmark games such as Go, Starcraft or Poker, Werewolf is unique in that the communication between agents is the key skill that must be mastered to obtain high levels of play. A successful Werewolf agent must be able to build a model of the other players' beliefs, identify allied players, and exchange this information through communication [13].

The 2019 Automated Negotiating Agents Competition was the first time that the AIWolf Project competition was held for an international audience. The participants were tasked to implement an agent capable of playing the Werewolf game against other automated players. The interaction of the agents is governed by a communication protocol[1]. This protocol uses a formal grammar, a fixed set of keywords, basic logic and causal expressions [30]. The keywords enable the players to express intent, beliefs about the game state, requests for information, and requests for action. For example, to express the following sentence:
"I vote for agent 3 because agent 3 did not vote for agent 4, and agent 4 was a werewolf",

An agent would have to use the following protocol sentence:
```
BECAUSE (AND (NOT <agent3> VOTE <agent4>)
             (COMINGOUT <agent4> <werewolf>))
(VOTE <agent3>)
```
The competition happens in two stages. In the *preliminary* stage, all agents play in a large number of trials. Each trial is composed of one hundred 15-player games, where the players are chosen randomly from the competition pool, and the roles are also distributed randomly. These trials are repeated until all agents have played a minimum number of games. The agents are ranked by their victory rate, and the 15 highest agents advance to the next stage. In the *finalist* stage, the participating design teams are allowed to modify their agents and submit source code and a description document. Then, the agents play several games against each other in 15-player games. The agent with the highest victory rate is declared the winner of the competition.

A total of 94 people registered to the competition, and 74 submitted agents. Out of those agents, 43 were disqualified due to bugs. Many of these disqualified competitors submitted a single version of their agents, which indicates that they did not review their agent based on the feedback from the testing server. Among the 15 finalists, 8 submitted agents in Java, 6 in Python, and 1 in C-Sharp. The

[1] AIWolf Protocol Version 3.6.

winning agent, "Takeda", had a 0.6 overall win rate, and a 0.68 villager win rate. Two of the finalists had to be disqualified due to bugs in their code.

Most of the finalist agents were forks of the agents that won the 4th Japanese AIWolf competition. Here we highlight the "Fisherman" agent, which used three different winners from the previous competition as basis, and chose which winner to play based on a multi-armed bandit strategy. Five agents, including the grand winner "Takeda", used some form of machine learning to estimate the team allegiance of the other players. The other nine agents used hand-crafted scripts to define the actions of the agents. Among these hand crafted rules, we highlight trying to remove agents with the highest or lowest winning rate on previous game, and rules for estimating the best timing for revealing role information.

We were satisfied with the high number of participants in this competition. However, the large number of disqualified agents shows that there is still a lot of work necessary in providing good quality English translations of the reference materials in Japanese, as well as better guidance on the use of the training server. In fact, all of the 15 finalists were from Japanese institutions, indicating that much work needs to be done for the internationalization of the Werewolf competition.

Regarding the strategies of the finalist agents, this year's protocol had many new features compared to last year's competition, in particular the introduction of logical and causal statements to the communication protocol. However, none of the winning agents made heavy use of these new features. In fact, it seems that the current winning strategy is to fine tune the ability of the agent to estimate the role of the other players based on their output, with very little back and forth happening between the players. This indicates that the best agents in werewolf are stuck in "wait and detect" local optima for their strategy. We hope that participants in future competition will find ways to exploit this fixed strategy.

With this in mind, the ANAC 2020 werewolf challenge will focus on refining the development environment by providing more documentation, example code, and tools, so that the participants can spend less time finding bugs in their agents, and more time developing interesting and diverse strategies for the Werewolf game.

6 Supply Chain Management

The SCM league models a real-world scenario characterized by profit-maximizing agents that inhabit a complex, dynamic, negotiation environment [28]. A distinguishing feature of the SCM league is the fact that agents' utility functions are endogenous. The agents are responsible for devising their utilities for various possible agreements, given their unique production capabilities, and then to negotiate with other agents to contract those that are most favorable to them.

The world modeled by SCML2019 consists of four types of entities: factories, miners, consumers, and an insurance company. In more detail:

Factories. Entities that convert raw materials and intermediate products into intermediate and final products by running their manufacturing processes.

Different factories are endowed with different capabilities, specified as *private* production profiles, known only to the factory's manager.

Miners. Facilities capable of mining raw materials as needed to satisfy their negotiated contracts. Miners act only as sellers in the SCM world.

Consumers. Companies interested in consuming a subset of the final products to satisfy some predefined consumption schedule. Consumers act only as buyers in the SCM world.

Insurance Company. A single insurance company that can insure buyers against *breaches of contract* committed by sellers, and vice versa.

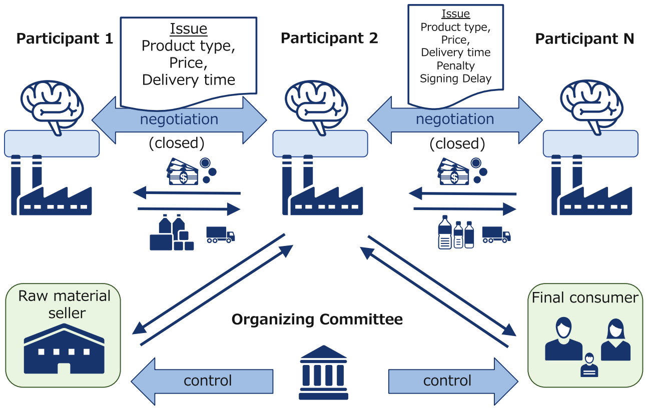

Fig. 3. SCML Organization. Factory managers controlled by participants negotiate with each other and the organization-committee provided miners, consumers and factory mangers.

In the SCM world, each type of entity is run by a *manager* agent. The organizing committee provided manager agents for miners, consumers, and the insurance company. Figure 3 shows the organization of SCML. The organizing committee provided a description of the behavior of these agents, including the miners' and consumers' (exact) utility functions, the factory managers' scheduling algorithm, and an estimation method for the factory managers' utility functions to all participants [28]. The simulation used NegMAS as the negotiation platform [27].

The committee also provided a default agent: i.e., a *greedy factory manager*, instances of which participated in the competition to ensure sufficiently many trading opportunities. The goal of each factory manager agent is to accrue as much wealth (i.e., profit) as possible.

Participants needed to write and submit code (in Java or Python) for an autonomous agent that acts as a factory manager trying to maximize its total profit on multiple simulations with varying world configurations.

The competition was conducted in three tracks: basic, collusion and sabotage. In the basic and collusion tracks, agents were tasked with maximizing their own profit. In the sabotage track, they were tasked with *minimizing* everyone else's profits. The difference between the basic and collusion tracks is that in the former at most one instance of every submitted agent was running in any simulation. In the collusion tracks, participants were encouraged to find ways for their agents to collude together to maximize their profit (e.g. by cornering the market). The sabotage track was introduced to find problems in the league design that could be exploited to block trade in the market.

After disqualifying agents that did not conform to the rules of the competition, six agents ran in the basic and collusion tracks and two agents ran in the sabotage competition.

7 Future Directions

This paper presents the challenges and discusses the results of the competition leagues. Future directions for research are determined by the participants in ANAC's leagues after the lessons learned have been shared. These directions per league are as follows.

For the Automated Negotiation Main league, the challenge for 2020 is to design a negotiating agent that can elicit preference information from a user during the negotiation. The idea is that when a negotiating agent represents a user in a negotiation, it does not know exactly what the user wants, and therefore the agent needs to actively improve its user model through a preference elicitation process. To improve the user model, the agent may elicit further information about the ranking against an elicitation cost.

For Diplomacy, we have concluded that the challenge requires a long-time effort beyond the possibilities for the current competitors, which may also explain the low number of competitors. Therefore, we decide to discontinue this league for the time being.

For the Werewolf league, we will focus in providing a more complete suite of manuals and sample code to participants, and extend the communication with organizers during the initial agent testing phase, with the objective of reducing the number of agents rejected due to bugs and crashes.

For the next Human-Agent competition, we have decided to continue to expand the problem by allowing human users to specify their own preferences. We hope this may help increase participant investment in the scenario, as well as encourage agent designers to respond to a variety of negotiation structures.

For the SCM league, we plan to strengthen the competition while reducing its complexity. This will be achieved by removing the insurance company, avoiding any sources of external funds from being inducted into the system, removing built-in agents from the simulation, having a larger variety of built-in agents and decomposing the agent into easy to manage components allowing participants to focus all of their efforts on the main challenge of situated negotiation.

Acknowledgements. This work is part of the Veni research programme with project number 639.021.751, which is financed by the The Dutch Research Council (NWO). This work was partially funded by project LOGISTAR, under the E.U. Horizon 2020 research and innovation programme, Grant Agreement No. 769142. This research was also sponsored by the U.S. Army Research Office and was accomplished under Cooperative Agreement Number W911NF-20-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

1. Aydogan, R., Fujita, K., Baarslag, T., Jonker, C.M., Ito, T.: ANAC 2018: repeated multilateral negotiation league. In: The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, Japan (2019)
2. Aydoğan, R., et al.: A baseline for non-linear bilateral negotiations: the full results of the agents competing in ANAC 2014 (2016)
3. Baarslag, T., Aydoğan, R., Hindriks, K.V., Fuijita, K., Ito, K., Jonker, C.M.: The automated negotiating agents competition, 2010–2015. AI Mag. **36**(4), 115–118 (2015)
4. Baarslag, T., Gerding, E.H.: Optimal incremental preference elicitation during negotiation. In: Proceedings of the Twenty-fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, pp. 3–9. AAAI Press (2015)
5. Baarslag, T., Hindriks, K., Jonker, C., Kraus, S., Lin, R.: The first automated negotiating agents competition (ANAC 2010). In: Ito, T., Zhang, M., Robu, V., Fatima, S., Matsuo, T. (eds.) New Trends in Agent-Based Complex Automated Negotiations, vol. 383, pp. 113–135. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-24696-8_7
6. Baarslag, T., Kaisers, M., Gerding, E.H., Jonker, C.M., Gratch, J.: Computers that negotiate on our behalf: major challenges for self-sufficient, self-directed, and interdependent negotiating agents. In: Sukthankar, G., Rodriguez-Aguilar, J.A. (eds.) AAMAS 2017. LNCS (LNAI), vol. 10643, pp. 143–163. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-71679-4_10
7. Baarslag, T., Kaisers, M., Gerding, E.H., Jonker, C.M., Gratch, J.: When will negotiation agents be able to represent us? The challenges and opportunities for autonomous negotiators. In: Proceedings of the Twenty-sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, pp. 4684–4690 (2017)
8. Fabregues, A.: Facing the challenge of automated negotiations with humans. PhD thesis, Universitat Autònoma de Barcelona (2012)
9. Fabregues, A., Sierra, C.: DipGame: a challenging negotiation testbed. Eng. Appl. Artif. Intell. **24**, 1137–1146 (2011)
10. Ferreira, A., Cardoso, H.L., Reis, L.P.: DipBlue: a diplomacy agent with strategic and trust reasoning. In: Proceedings of the 7th International Conference on Agents and Artificial Intelligence (ICAART 2015), pp. 398–405 (2015)
11. Fujita, K., Aydoğan, R., Baarslag, T., Hindriks, K., Ito, T., Jonker, C.: The sixth automated negotiating agents competition (ANAC 2015). In: Fujita, K., Bai, Q., Ito, T., Zhang, M., Ren, F., Aydoğan, R., Hadfi, R. (eds.) Modern Approaches to Agent-based Complex Automated Negotiation. SCI, vol. 674, pp. 139–151. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-51563-2_9

12. Hindriks, K., Jonker, C.M., Tykhonov, D.: Let's dans! an analytic framework of negotiation dynamics and strategies. Web Intell. Agent Syst. Int. J. **9**(4), 319–335 (2011)
13. Hirata, Y., et al.: Werewolf game modeling using action probabilities based on play log analysis. In: Plaat, A., Kosters, W., van den Herik, J. (eds.) CG 2016. LNCS, vol. 10068, pp. 103–114. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-50935-8_10
14. Jennings, N.R., Faratin, P., Lomuscio, A.R., Parsons, S., Sierra, C., Wooldridge, M.: Automated negotiation: prospects, methods and challenges. Int. J. Group Decis. Negot. **10**(2), 199–215 (2001)
15. de Jonge, D., Baarslag, T., Aydoğan, R., Jonker, C., Fujita, K., Ito, T.: The challenge of negotiation in the game of diplomacy. In: Lujak, M. (ed.) AT 2018. LNCS (LNAI), vol. 11327, pp. 100–114. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-17294-7_8
16. de Jonge, D., Sierra, C.: NB3: a multilateral negotiation algorithm for large, non-linear agreement spaces with limited time. Auton. Agent. Multi-Agent Syst. **29**(5), 896–942 (2015)
17. de Jonge, D., Sierra, C.: D-brane: a diplomacy playing agent for automated negotiations research. Appl. Intell. **47**, 1–20 (2017)
18. Jonker, C.M., et al.: An introduction to the pocket negotiator: a general purpose negotiation support system. In: Criado Pacheco, N., Carrascosa, C., Osman, N., Julián Inglada, V. (eds.) EUMAS/AT -2016. LNCS (LNAI), vol. 10207, pp. 13–27. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59294-7_2
19. Kelley, H.H.: A classroom study of the dilemmas in interpersonal negotiations. In: Strategic Interaction and Conflict, p. 49 (1966)
20. Lin, R., Kraus, S., Baarslag, T., Tykhonov, D., Hindriks, K., Jonker, C.M.: Genius: an integrated environment for supporting the design of generic automated negotiators. Comput. Intell. **30**(1), 48–70 (2014)
21. Lomuscio, A.R., Wooldridge, M., Jennings, N.R.: A classification scheme for negotiation in electronic commerce. Group Decis. Negot. **12**(1), 31–56 (2003)
22. Marinheiro, J., Lopes Cardoso, H.: Towards general cooperative game playing. In: Nguyen, N.T., Kowalczyk, R., van den Herik, J., Rocha, A.P., Filipe, J. (eds.) Transactions on Computational Collective Intelligence XXVIII. LNCS, vol. 10780, pp. 164–192. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-78301-7_8
23. Marsa-Maestre, I., Klein, M., Jonker, C.M., Aydoğan, R.: From problems to protocols: towards a negotiation handbook. Decis. Support Syst. **60**, 39–54 (2014)
24. Mell, J., Gratch, J.: IAGO: interactive arbitration guide online. In: Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems, pp. 1510–1512. International Foundation for Autonomous Agents and Multiagent Systems (2016)
25. Mell, J., Gratch, J., Aydoğan, R., Baarslag, T., Jonker, C.M.: The likeability-success tradeoff: results of the 2nd annual human-agent automated negotiating agents competition. In: Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction (ACII 2019), pp. 1–7. IEEE (2019)
26. Mell, J., Gratch, J., Baarslag, T., Aydoğan, R., Jonker, C.M.: Results of the first annual human-agent league of the automated negotiating agents competition. In: Proceedings of the 18th International Conference on Intelligent Virtual Agents, pp. 23–28. ACM (2018)
27. Mohammad, Y., Greenwald, A., Nakadai, S.: NegMAS: a platform for situated negotiations. In: Twelfth International Workshop on Agent-based Complex Automated Negotiations (ACAN 2019) in Conjunction with IJCAI (2019)

28. Mohammad, Y., Viqueira, E.A., Ayerza, N.A., Greenwald, A., Nakadai, S., Morinaga, S.: Supply chain management world. In: Baldoni, M., Dastani, M., Liao, B., Sakurai, Y., Zalila Wenkstern, R. (eds.) PRIMA 2019. LNCS (LNAI), vol. 11873, pp. 153–169. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33792-6_10
29. Narayanan, V., Jennings, N.R.: An adaptive bilateral negotiation model for e-commerce settings. Econometrica **50**, 97–110 (1982)
30. Osawa, H., Otsuki, T., Aranha, C., Toriumi, F.: Negotiation in hidden identity: designing protocol for werewolf game. In: Agent-Based Complex Automated Negotiation Workshop (2019)
31. Oshrat, Y., Lin, R., Kraus, S.: Facing the challenge of human-agent negotiations via effective general opponent modeling. In: AAMAS (1), pp. 377–384. IFAAMAS (2009)
32. Patton, B.: Negotiation. In: The Handbook of Dispute Resolution, pp. 279–303 (2005)
33. Raiffa, H.: The Art and Science of Negotiation. Harvard University Press, Cambridge (1982)
34. Rosenschein, J.S., Zlotkin, G.: Rules of Encounter: Designing Conventions for Automated Negotiation Among Computers. MIT Press, Cambridge (1994)
35. Ephrati, E., Kraus, S., Lehman, D.: An automated diplomacy player. In: Levy, D., Beal, D. (eds.) Heuristic Programming in Artificial Intelligence: The 1st Computer Olympia, pp. 134–153. Ellis Horwood Limited (1989)
36. Sandholm, T.: Automated negotiation. Commun. ACM **42**(3), 84–85 (1999)
37. Sycara, K.P.: Resolving goal conflicts via negotiation. In: AAAI, vol. 88, pp. 21–26 (1988)
38. Sycara, K.P.: Negotiation planning: an AI approach. Eur. J. Oper. Res. **46**(2), 216–234 (1990)
39. Toriumi, F., Osawa, H., Inaba, M., Katagami, D., Shinoda, K., Matsubara, H.: AI wolf contest — Development of game AI using collective intelligence —. In: Cazenave, T., Winands, M.H.M., Edelkamp, S., Schiffel, S., Thielscher, M., Togelius, J. (eds.) CGW/GIGA -2016. CCIS, vol. 705, pp. 101–115. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57969-6_8