3rd Annual Human-Agent Negotiation track, ANAC 2019

Motivation:

The Human-Agent Negotiation (HAN) competition is conducted to explore the strategies, nuances, and difficulties in creating realistic and efficient agents whose primary purpose is to *negotiate with humans*. Previous work on human-agent negotiation has revealed the importance of several features not commonly present in agent-agent negotiation, including retractable and partial offers, emotion exchange, preference elicitation strategies, favors and ledgers behavior, and myriad other topics. To understand these features and better create agents that use them, this competition is designed to be a showcase for the newest work in the negotiating agent community. This year's challenge will focus primarily on the nuances of *repeated* negotiations, which require more complex strategies than one-shot negotiations.

Please be sure and read carefully, as several details may have changed. These details have been written in red.

Early submission deadline: **Monday, April 22nd, 2019**. (optional, for code review)
Final submission deadline: **Monday, June 15th, 2019**. (extended).
We encourage you to submit early to test your code compilation.
Notification of finalists: **Saturday, June 19th, 2019**.
Competition special session: **Saturday, August 10th – Friday, August 16th @ IJCAI (exact date TBD)**

Please note that we will be offering limited code review to test for compilation errors and/or bugs for those who submit by the early deadline in April, as a new courtesy this year. As always, we encourage you to test extensively locally, as this service is not a substitute for debugging.

Summary:

The HAN competition requires each author or group of authors to submit an agent that will be tested in competition against human subjects in a study run through the University of Southern California. Based on the performance of the agent, we will determine which agent strategies are most effective. The subject pool will be taken from the standard populace available on Amazon's Mechanical Turk (MTurk) service, with normal filtration done for participants who are ineligible (see "Subject Selection", below).

All agents must be compliant with the IAGO (Interactive Arbitration Guide Online) framework and API, which will allow standardization of the agents and efficient running of subjects on MTurk. The most up-to-date version of IAGO is required, and will be available for download in **February**. The version of IAGO used in the 2nd Annual HAN is available for download now, for researchers interested in getting a head-start.

Agents will all be run on the same set of multi-issue bargaining tasks, examples of which are included below ("Domain Example"). Agents will be allowed to communicate on several channels, including a set of natural language utterances that have been pre-selected and curated by the ANAC committee. Other channels include the exchange of offers through visual cues and natural language, preference statements, and emotional displays.

2019 Challenge:

This competition is intended to foster agents that can maximize their individual utility when engaged in repeated negotiations with the same human opponent. Human participants will compete against each submitted agent in three back-to-back negotiations. One difference from ANAC2018 is that the structure of each negotiation may be different (i.e., the agent and the human participant may have different preferences in each negotiation). [1] The reason for this change is to reward agents that can establish a cooperative relationship with the human participant, as opposed to rewarding agents that are simply good at learning the human's preferences.

Like the 2018 challenge, these preferences will be <u>unknown</u> to the opposing side at the beginning of the three negotiations. In this way, agents that do a good job of learning opponent's preferences will likely outperform agents that do not.

More fundamentally, this approach allows us to capture which agent strategies successfully account for 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. This year's challenge will provide insight into these and more choices when designing agents whose primary purpose is to negotiate with humans over time.

IAGO API:

IAGO is a platform developed by Mell and Gratch at the University of Southern California. It serves as a testbed for Human-Agent negotiation specifically. IAGO is a web-based servlet hosting system that provides data collection and recording services, a human-usable HTML5 UI, and an API for designing human-like agents.

A full documentation of IAGO is available from the download site, available at https://myiago.com.

---

[1] Note: The total possible points for each side, over all three negotiations, will equal the same amount. See "Additional Rules".

, including guides to getting started, tips and tricks, and live example agents (as well as up-to-date Javadocs). This website will have additional information added through January 2019.

Submission Details:

The Submission Portal will open in April, and will allow for initial submissions for code review. Additional materials **will be required** this year, including a 2-page maximum description of agents, which should describe any algorithms (novel or otherwise) in use that drive the agent's strategy. Further information on these requirements (as well as formatting guides), will be made available on the IAGO website (https://myiago.com).

Subject Selection and Data Treatment:

Competition subject participants will be selected from the MTurk subject pool. Subjects will be adults in the US (18 years or older), and will assert that they are permanent residents of the US (this will be verified with IP address tracking). Restriction to the US will be done to reduce cross-cultural effects. Each agent will be tested against 25 participants. Participants will not be re-used or be matched against more than one agent. All testing will be approved by USC's Institutional Review Board (an ethics committee).

Due to the fact that MTurk participants will be US-restricted and natural language statements are used in the utterance set of the competition, participants will also be asked to affirm that their first language is English.

Basic demographic information of subjects will be collected, and the subjects may be asked a set of verification questions/attention checks to ensure they comprehend and are engaged in the negotiation. Subjects who fail these questions will be removed from the competition and the resulting data set. If a substantial amount of subjects are removed due to failing an attention check, additional subjects may be run. Subjects whose data is not captured due to agent malfunctions will generally not be rerun (see "Testing", below).

The data set collected by the competition organizers may be released to the organizing committee, and all agent developers/researchers may request access to the data if the organizing committee releases it. All submitted source code may be released and/or reused by the organizing committee. Researchers not wishing to release source code should contact the organizers directly.

Competition Winners and Evaluation:

A set of prizes will be awarded to the winners of the competition according to the highest score achieved by the agent. The winner will be the researcher whose agent has achieved the highest score at the end of the bargaining time. Non-significant differences will be tie-

broken by the highest score. All differences, including differences between the control agent and submitted agents, will be reported.

Note that since there will be a series of negotiations, aggressive strategies may backfire.

Note: The 2019 Challenge does not have a Likeability Prize. Please see "2019 Challenge" for how likeability indirectly affects outcome.

We maintain the opportunity to examine other categories for "bonus" prizes.

Testing

Like the 2018 competition, we will provide automated compilation testing for all submissions. Additionally, as a new service, we will provide limited <u>runtime</u> testing for those who submit by the early submission deadline. We will also provide a guide for manual runtime testing with internal subjects (so that you may test your own agents before submitting). Note that our runtime testing is purely a courtesy, and is not intended as a guarantee that the agent will run "bug-free".

Important: Agents that experience malfunctions during runtime will have incomplete data excised, and additional subjects will not be re-run.

Domain Example:

We present here an example domain. A domain similar to this will be used in the actual competition.

This negotiation is a multi-issue bargaining task, which means both the agent and the human participant will negotiate over the same set of items. Items may have differing values to each side. A "full offer" means that all items are assigned to either the agent or the human participant. A "partial offer" means that some items remain on the table and undecided. <u>*No offer is considered binding until both players accept the same full offer.*</u>

A negotiation will only end when such a full offer is accepted, or the 7-minute time limit for the negotiation has expired. Human participants will have a warning shown when there is only 1 minute remaining. Agents will have continuous access to the current negotiation time, accurate within approximately 5 seconds. In the case that time expires with no full offer, each player will take points equal to their respective Best Alternative To Negotiated Agreement (BATNA).

Note that the 2019 version of IAGO (3.0) allows for new agent/human actions. These include direct discussion of BATNA as well as the ability to terminate a negotiation early "walking away".

Note that the IAGO API allows agent designers to read the natural language descriptions of the issues at runtime (e.g., "Issue1" can be understood to be something like "Lumber" or "Luxury Cars"). However, agents will make use of domain-agnostic calls.

The following example challenge is a simple multi-issue bargaining task over resources between two countries. There will be four distinct resources, with five items in each category. The items will have images and descriptions identifying them as either "Oil", "Iron", "Foodstuffs" or "Lumber". The human player is assigned a value of 4 points to each Oil, 3 points to each Iron, 2 points to each Lumber, and 1 point to each Foodstuff. The agent player is assigned a value of 4 points to each Foodstuff, 3 points to each Lumber, 2 points to each Iron, and 1 point to each Oil. Each player's BATNA is equal to 4, i.e., the value of a single one of their highest item.

In the second and third negotiations, these values may change.

Note than in both domains, the human's point values and BATNA will NOT be revealed to the agent designers prior to the competition.

Natural Language Utterances:

Please see the IAGO website for the most up-to-date version of the pre-selected utterances that humans can use with the agent. Agents may use any dialogue they wish.

New dialogue options for discussion player BATNA have been added!

Additional Rules:

Competition participants will be given a test scenario to practice with their agents. However, to prevent hard-coding preference data into agents, a different set of preferences will be used for the actual competition.

There will be no fewer than 3 distinct issues, and no greater than 5. Each issue will have fewer than 20 items. There will be three back-to-back negotiation games.

Issue utilities will adhere to the following new rule:

$$\sum_{i=1}^{k}\sum_{g=1}^{j} \text{Agent\_utility}(i,g) * (\#levels(i,g) - 1) = \sum_{i=1}^{k}\sum_{g=1}^{j} \text{Human\_utility}(i,g) * (\#levels(i,g) - 1)$$

where k is the total number of issues and j is the total number of negotiation games.

Succinctly, this relationship means that the total for each side would be the same if that side obtained every item, over all the games. Note that the total for a SINGLE game may not be even, but over all games, it will be.

Note that no single researcher may appear on more than one submission. Exceptions in the case of an advisor wishing to appear on two student papers will be granted on a case-by-case basis, and inquiries should be made to iago@ict.usc.edu.

It is highly encouraged that researchers use any technique by which an agent can successfully store information within the three negotiations for a given participant. This includes methods by which the agent may learn preferences in one negotiation and then subsequently passes that information back to itself in future negotiation. IAGO's agents persist across all three rounds, so data may be stored within your agent Class, assuming it respects GAME_START events. However, the intent of this competition is not to learn an entire domain, and therefore data may not be stored across participants--all 25 participants are to be treated as fresh instances against which the same agent will be run.

Note: Participation in this competition is done in good spirit and for the furtherance of academic knowledge. Attempts to circumvent the rules described herein or as they are described by the ANAC organizers will not qualify for prizes.

Reference:

Mell, J., Gratch, J. (2016) "IAGO: Interactive Arbitration Guide Online", In *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems.* International Foundation for Autonomous Agents and Multiagent Systems.

Mell, J., Gratch, J. (2017) "Grumpy & Pinocchio: Answering Human-Agent Negotiation Questions through Realistic Agent Design", *Proceedings of the 2017 International Conference on Autonomous Agents and Multiagent Systems International Foundation for Autonomous Agents and Multiagent Systems.*

Mell, J., Gratch, J., Baarslag, T., Aydoğan, R., Jonker, C. (2018) "Results of the First Annual Human-Agent League of the Automated Negotiating Agents Competition", in *Proceedings of the 2018 International Conference on Intelligent Virtual Agents.*