Delft University of Technology

The Likeability-Success Tradeoff
Results of the 2nd Annual Human-Agent Automated Negotiating Agents Competition

Mell, Johnathan; Gratch, Jonathan; Aydogan, R.; Baarslag, Tim; Jonker, Catholijn M.

DOI
10.1109/ACII.2019.8925437

Publication date
2019

Document Version
Final published version

Published in
2019 8th International Conference on Affective Computing and Intelligent Interaction, ACII 2019

Citation (APA)
Mell, J., Gratch, J., Aydogan, R., Baarslag, T., & Jonker, C. M. (2019). The Likeability-Success Tradeoff: Results of the 2nd Annual Human-Agent Automated Negotiating Agents Competition. In *2019 8th International Conference on Affective Computing and Intelligent Interaction, ACII 2019* Article 8925437 (2019 8th International Conference on Affective Computing and Intelligent Interaction, ACII 2019). IEEE. https://doi.org/10.1109/ACII.2019.8925437

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' – Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)

The Likeability-Success Tradeoff: Results of the 2nd Annual Human-Agent Automated Negotiating Agents Competition

Johnathan Mell
USC Institute for Creative Technologies
Los Angeles, USA
mell@ict.usc.edu

Jonathan Gratch
USC Institute for Creative Technologies
Los Angeles, USA
gratch@ict.usc.edu

Reyhan Aydoğan
Özyeğin University
Istanbul, Turkey
reyhan.aydogan@ozyegin.edu.tr

Tim Baarslag
Centrum Wiskunde & Informatica | Utrecht Univ.
Amsterdam | Utrecht, The Netherlands
T.Baarslag@cwi.nl

Catholijn M. Jonker
Delft Univ. of Technology | Leiden University
Delft | Leiden, The Netherlands
C.M.Jonker@tudelft.nl

***Abstract*—We present the results of the 2nd Annual Human-Agent League of the Automated Negotiating Agent Competition. Building on the success of the previous year's results, a new challenge was issued that focused exploring the *likeability-success tradeoff* in negotiations. By examining a series of repeated negotiations, actions may affect the relationship between automated negotiating agents and their human competitors over time. The results presented herein support a more complex view of human-agent negotiation and capture of integrative potential (win-win solutions). We show that, although likeability is generally seen as a tradeoff to winning, agents are able to remain well-liked while winning if integrative potential is not discovered in a given negotiation. The results indicate that the top-performing agent in this competition took advantage of this loophole by engaging in favor exchange across negotiations (cross-game logrolling). These exploratory results provide information about the effects of different submitted "black-box" agents in human-agent negotiation and provide a state-of-the-art benchmark for human-agent design.**

Keywords*—human agent interaction; negotiation; empirical results in HCI

I. INTRODUCTION

Understanding how humans negotiate has been a key question in business and psychological literatures for many years—it is a complex social task [7][16][17]. But as humans increasingly rely on automated agents to interact with the world around them, we must increasingly design agents that are capable of interacting with humans in these sorts of social tasks. However, negotiation involves a delicate balance between two entwined goals: success and likeability. Succeed too much and too often, and negotiating partners will be hard to find or hostile. Bend over backwards to ensure likeability in negotiation, and one may find oneself on the losing end of too many deals. In this year's Automated Negotiating Agent Competition (ANAC) [6], we explore these dual goals, and begin to disentangle their complicated relationship through high-level analysis of a series of designer-submitted agents. While ANAC has been a recurrent, successful competition for 9 years (2019 has marked the 10th annual ANAC), it has been focused primarily on agent-agent negotiation. Human-agent negotiation is fundamentally different than agent-agent negotiation, and the Human-Agent Track of ANAC was added in 2017 to promote research into this promising area.

But while the results of ANAC 2017 indicated that indeed, the initial agents were capable of rising to these challenges and negotiating on a reasonable level with humans, the negotiating space was fundamentally limited. Negotiations among humans are rarely single interactions—rather, interactions often recur, and the same negotiating partners will begin to be part of an ongoing negotiating relationship. This brings with it a new set of social problems for agents to tackle: rapport-building [14], favor-exchange [11], and reputation effects [19]. With these social problems in tow, negotiation with humans takes on new subtleties—one cannot pursue shortsighted strategies that damage relationships too early without paying the price in future interactions.

In this work, we take an important step to widespread analysis of likeability and success, within the more realistic and far broader world of *repeated negotiation.* The Human-Agent League for ANAC 2018 created a challenge to negotiate with humans over a series of three repeated negotiations. The competition received 10 submissions from a variety of worldwide organizations, and these agents employed a variety of strategies to succeed against humans. Agents were forced to consider the likeability-success tradeoff, and needed to find robust strategies that allowed them to succeed in a broader domain.

In analysis of the competition data, we discover that the previously reported tradeoff between agent success and agent likeability [10] is more complex than at first assumed. Critical to negotiation, and to this important relationship, is the idea of "integrative potential". While there exist many competing notions of integrative potential[1], we look at

---

[1] Integrative potential refers generally to the ability for negotiators to "grow the pie" and increase the total points available to *either* party. One can measure the "amount of integrative potential discovered' by looking at several metrics, including opposition metrics that relate to the Kalai-Smorodinsky Solution, or to the Nash Product [2]. To parallel previous work on ANAC, we define this quantity as the "joint points", or the total value of the agreed solution as discovered by the agent and the human.

978-1-7281-3888-6/19/$31.00 ©2019 IEEE

"joint points", or the total amount of utility generated by the final agreement, when discussing this phenomenon. Specifically, in these results, we find that that when integrative potential is not discovered, agents remain likeable even when winning. We also present the strategy of the winning agent of the competition, which benefited from this observation by exchanging favors across games, thus maintaining distributive solutions within negotiations.

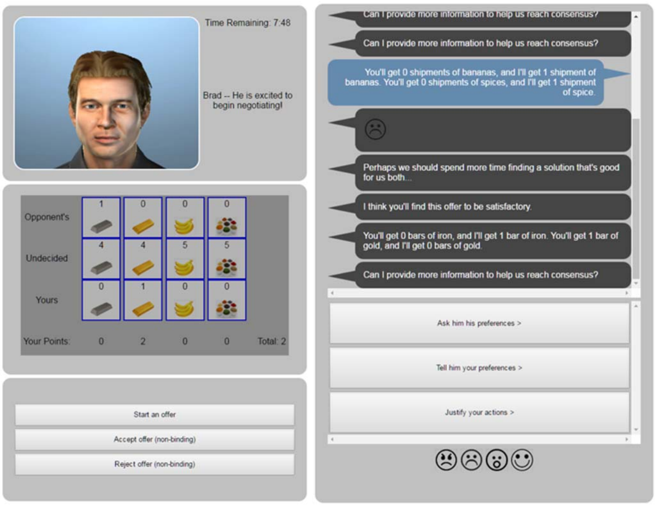
II. COMPETITION DESIGN

A. *IAGO Negotiation Platform*

The IAGO Negotiation platform was proposed and designed by Mell et al. and was selected to be used for the Human-Agent League of ANAC [13] in 2017. Results from this first human-agent competition have been previously published [10]. IAGO provides a front-facing GUI for the human-participants (see Fig. 1). In particular, 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. These channels are in addition to the traditional methods supported by agent-agent negotiation platforms, such as exchanging offers. IAGO also allows offers to be sent that do not involve all the items in the negotiation ("partial offers").

These features of IAGO mean that it provides a platform to address the basic features that intelligent negotiating agents require. It provides information that allows for robust user modeling and allows multiple channels for communicating in different ways. IAGO provides information that agents require to reason about their own preferences and allows them to pursue a number of more complex strategies that require specific features (such as partial offers).


Time Remaining: 7:48

Brad -- He is excited to begin negotiating!

Opponent's	1	0	0	0
Undecided	4	4	5	5
Yours	0	1	0	0
Your Points:	0	2	0	0	Total: 2

Can I provide more information to help us reach consensus?

Can I provide more information to help us reach consensus?

You'll get 0 shipments of bananas, and I'll get 1 shipment of bananas. You'll get 0 shipments of spices, and I'll get 1 shipment of spice.

☹

Perhaps we should spend more time finding a solution that's good for us both...

I think you'll find this offer to be satisfactory.

You'll get 0 bars of iron, and I'll get 1 bar of iron. You'll get 1 bar of gold, and I'll get 0 bars of gold.

Can I provide more information to help us reach consensus?

Ask him his preferences >
Tell him your preferences >
Justify your actions >
Start an offer
Accept offer (non-binding)
Reject offer (non-binding)

Fig. 1. IAGO Research Platform (Client View)

B. *Human-Agent Competition Design*

1) *General Information*

Much like the First Annual Competition, this second competition featured an array of participant-submitted agents competing against humans in multi-issue negotiation. However, in contrast to a single, 10-minute, multi-issue negotiation, participants engaged in three, 7-minute multi-issue negotiations. While this by no means encapsulates the entirety of the human-agent negotiation space, it does allow us to narrow in on a set of questions about the dynamics of likeability and success over multiple negotiations, rather than in a single snapshot.[2]

Participants were also asked a series of questions, ranging from demographic information to reviews of the agent behavior. Some of these questions were repeated, such that responses were received between the three negotiations (as well as at the beginning and end of the exercise). The submitted agents were judged according to the total number of points they earned, and prizes were awarded to the best agents. All agent designers submitted their own agents, which had to implement their own strategies with regards to opponent modeling, emotional understanding, and natural language response (see below). We were able to discern these strategies on a high-level, based on manual code inspection, process variable measurement, and designer-submitted descriptions.

Results from the 2017 competition had indicated that there was a tradeoff between scoring well and being well-liked—this competition was designed to craft a measure of score that took into account likeability. If such a tradeoff truly exists, then pursuing short-sighted strategies that increase points in the first negotiation but come at a severe cost to likeability may result in fewer points overall. Therefore, no prizes were considered for likeability alone, as agent score was thought to include the "useful" component of that measure. Likeability was still measured through self-report and was determined by user-submitted responses to Likert-scale questions after the end of each negotiation.

Agent designers were provided with a set of guidelines that restricted the domain of the negotiation within moderate bounds, but they were not given the details of the task itself, which was determined secretly prior to agent submission. These guidelines stipulated that the total for each side would be the same if that side got every item. Furthermore, an additional restriction was added to the competition this year, such that each negotiation would have the same total utility (although the *specific* issues could change between negotiations). In this way, the agent designers were given some idea of the scope of the negotiation and would be confident that no single negotiation would decide the overall outcome.

Designers were also provided with a limited set of natural language utterances that the humans could use in the negotiation. These utterances were preselected based on a modified set used in the previous year's competition. They were communicated to the agent designers in advance, and human users were able to send them using a

---

[2] Indeed, other topics considered for future competitions include the effectiveness of these human-like agents against their peer agents, mixed human-human/human-agent competitions, and partner choice. We welcome papers and future work that examine these topics.

GUI button interface. This allowed designers to craft agents that could respond to human messages without having to design/implement NLP modules for their agents (which was not the focus of this year's competition). Human players could also send pre-coded messages that contained information about their preferences, in addition to using the emotional and offer channels. Agents were unrestricted in the types of messages they could send back to players. Finally, agent developers were provided with the source code for a baseline agent ("Pinocchio") which was provided with the IAGO platform.

The competition's three negotiations had 4 issues, with a varying amount of levels to each. The issues had 3, 2, 6, and 3 items, but this initial ordering varied for each of the negotiations (e.g., the second negotiation could have been 6, 2, 3, 3). The task was partially integrative, with both sides gaining the most points by receiving the 6-item issue, but differing on their preferences for the two 3-item issues. Both sides also included a BATNA, which gave both players a minimum number of points should they fail to reach agreement. These three negotiations precisely mirrored the structure of ANAC 2017, in order to facilitate further scrutiny of the winning-likeability tradeoff.

2) *Participant Information*

Competition subject participants were selected from the MTurk subject pool. Subjects were adults and asserted that they were permanent residents of the US (verified with IP address). Restriction to the US was chosen in order to reduce cross-cultural variance. Each submitted competition agent was tested against ~25 participants. Participants were not re-used or matched against more than one agent. Subjects were asked a set of verification questions/attention checks to ensure they comprehended and were engaged in the negotiation. Additional participants were run to ensure n > 18 for each agent.[3]

All participants were presented with a tutorial of the system before use. Participants were paid regardless of their success in the negotiation. However, they were also awarded "lottery tickets" based on their performance. These lottery tickets then entered them into a prize drawing for one of several $10 MTurk credits, incentivizing good performance during the negotiation. This design allowed the competition to follow best practices for subject recruitment and handling, in line with other research [1].

III. AGENT DESIGN

As with human-human negotiation, there are many effective tactics that can lead to success in agent negotiation. These strategies have grown only more complex as work has examined *repeated negotiations*, where previous actions have impact on future interactions. This need to manage both human logic and human social perceptions of an agent led to agents that exploited numerous channels in their design.

Among the agents that were submitted to this competition, for example, there are several that use emotion in an attempt to influence their opponent. This strategy (particularly the use of negative emotions to gain concession) has been well-documented both in human-human and human-agent contexts [18]. Agent Keni, for example, tried to adjust its strategy based on the number of human negative emotions received. Other agents attempted rapport with the human-participant through positive emotion, with the hope that it would lead to greater value. The top-rated likeability agent, Glinda, used this strategy.

Within repeated negations, the idea of favor exchange (or "logrolling") has been explored Since agents are evaluated on their ability to win over several games, losing a single game may be a viable strategy for building likeability and thus winning more in the long run. Indeed, Agent Equalist used this strategy to score the highest of any agent in the competition.

All the agents also model their opponent to an extent, although agents do differ in how they accomplish this. The baseline agent (Agent Pinocchio) merely used information about preferences that had been explicitly given by the human user and then took a highly optimistic view of the remaining logical opponent models. Other agents, such as Quirinal, Keni, or Morty, all use some measure of the opponent's offer history to create better opponent models.

We provide brief accounts of the submitted agents below, and summarize their provenance in TABLE I. These accounts are based on designer-submitted descriptive documents, manual code examination, and examination of process variables and online testing. While we present these accounts in an attempt to outline the broad strokes of this year's agent designs, our analysis remains somewhat "black-box", and relies only on the process measures that are recorded within IAGO.

1) *Pinocchio (Baseline)*

The baseline agent was provided to all participants of the competition. Pinocchio followed a straightforward strategy that has been discussed in detail in previous work, as a part of the IAGO toolkit [12], but it follows a strategy most similar to the "Conceder" as described in [8]. Its behavior did not change over time, nor did it attempt to benefit from prior knowledge.

TABLE I. AGENT PROVENANCE

Agent	Institution	Authors
Pinocchio (Baseline)	Univ. of Southern California, USA	J. Mell
Athena	Southwest University, China	X. Li, Z. Dou, J. Li
Cyh	Tianjin University, China	Y. Cui, J. Hao
Equalist	Bar Ilan University, Israel	G. Yadgar, D. Juravski, N. Tshuva, C. Rozenshtein, K. Babay
Glinda	Bar Ilan University, Israel	D. Nisim, N. Yakar, I. Nimni, N. David
Keni	Bar Ilan University, Israel	I. Achituve, E. Orbach, K. Gilad, N. Shectman
Morty	Southwest University, China	L. Yuan, S. Chen
Quirinal	Bar Ilan University, Israel	P. Roit, A. Zanbar, H. Besser, A. Maymon, A. Fux
Smoc	Bar Ilan University, Israel	O. Drein
Emma	University of Strathclyde, UK	S. Daronnat
XDS	Tianjin University, China	D. Xie, J. Hao

---

[3] Participants were roughly gender-balanced, and followed common demographic trends in MTurk users.

2) *Athena*

Athena focuses on developing a dynamic natural language interaction with the human user, accounting for human emotions. In the face of aggressive human partners, Athena will refuse to concede, but with more passive partners, Athena will be cooperative. Athena also concedes more and "forgives" bad behavior in earlier rounds, in an attempt to teach its partner over time. This strategy mirrors on a high-level the "Tit-for-tat" agents in [8].

3) *Equalist*

Equalist explicitly attempts to exchange favors across negotiations (cross-game logrolling). Specifically, it attempts to offer a very positive deal for the human in the first negotiation, in exchange for a very positive deal for it in the second negotiation. However, in the third negotiation, Equalist maintains a tough perspective, in order to pull ahead of its opponent. Equalist therefore alternates between "Conceder" and "Hardheaded" agents, per the terminology in [8].

4) *Glinda*

Glinda is based primarily off of "Elphaba", an agent submitted to ANAC 2018. However, Glinda focuses entirely on trying to improve likeability through friendliness, and is a "Conceder" agent.

5) *Keni*

Keni bases its opponent model on both preference statements by the user as well the frequency of issues within offers. Keni attempts to take issues that it believes to be of little worth to the player, and attempts to change its natural language statements based on expressed emotion. It is conciliatory, and best fits the "Conceder" description.

6) *Morty*

Morty attempts to glean a behavioral model of its partner by examining the issues the partner used in his offers. Furthermore, Morty attempts to "stall" the first negotiation in order to gain more information. In later negotiations, it adjusts its concession rate based on this first negotiation, thus fitting the "Tit-for-tat" approach.

7) *Quirinal*

Quirinal incorporates human offers as well as human preference statements to try to determine an accurate ordering of issues for its opponent. Quirinal also generally tries to find mutually beneficial deals, but will occasionally explore the space with non-optimal deals. Quirinal's margin for acceptance widens as the negotiation goes on, and is a "Conceder" agent.

8) *Emma*

Emma takes a two-phase approach to negotiation. In the first half of a negotiation, Emma provides counter-offers based on the total number of items offered (regardless of their actual value). In the second half, Emma instead focuses on the value of items. Since exact utilities of opponent items are often unknown early on, this may be a helpful approach. This strategy varies from others, and may fall into the "Random" agent category, per [8].

9) *XDS*

The XDS agent seeks out integrative potential in all rounds and tries to come to a "win-win" solution. XDS also purports to be affected by human emotion and tries to account for human lying. XDS is a "Conceder".

IV. RESULTS

A. *Method*

For the purposes of the competition, all agents' scores were compared one-to-one. Dunnett's 2-sided test confirmed any significant differences for one-way contrasts against the baseline agent, Pinocchio. Significant differences between submitted agents were determined with post-hoc analysis, using Bonferroni correction. Likeability was determined by a series of self-reported 7-point Likert questions after negotiation:

- How satisfied were you with the final agreement?
- How much do you like your opponent?
- Would you negotiate with this opponent again?

Likeability was previously used in the ANAC 2017 results and found to have high reliability.

B. *Likeability*

Likeability varied substantially across the submitted agents. We examined the total likability by summing the three individual likeability scores for each negotiation. Quirinal had the lowest overall likeability (mean of 12.2 points) while Glinda had the highest overall (mean of 17.7 points). The average was 14.4 points. Pinocchio, the baseline agent, had a likeability of 15.8 points. The mean likeability scores for each agent are shown in Fig. 3.

C. *Agent Score*

Agent score took into account the total agent points summed over all three negotiations. It considered agents points earned only (and did not consider the human score). However, Equalist was the top-scoring agent both under that methodology and by examining "point lead" (human score subtracted from agent score). Equalist scored 89.72 points on average, putting it well-above the baseline agent Pinocchio (which scored 72.82) and the average score (mean 76.96). The lowest scoring agent was Emma, with an average of 68.35 points. These scores can be seen in Fig. 4. Equalist was significantly higher than its closest competitor (XDS), under Bonferroni correction ($p < 0.05$).

We also break out the individual results in Fig. 5. In general, the average scores for each negotiation across all agents were similar: negotiation 1 mean was 24.8, negotiation 2 mean was 26.7, negotiation 3 mean was 25.9. However, among individual agents, there were significant differences. Emma saw a large drop in performance between N1 and N2 (25.4 to 20.5) while Equalist showed a staggering increase (24.5 to 36.1).

D. *Cross-Game Logrolling*

Due to the large differences between agent score across games in certain agents, we decided to examine the differences as a structural feature of the interaction. We examine the maximum point spread between negotiations and performed regression analysis to examine correlations to "winning" (as measured by the agent point lead). We found a significant, positive correlations such that agents that have larger differences between their scores in different negotiations tend to score better overall ($t = 3.211$, $N=240$, $p = 0.002$). This effect is largely driven by the Equalist agent, which had a strategy that was designed to maximize these differences by exchanging favors and "logrolling" across multiple negotiations.

E. *Winning and Likeability*

Previous work has indicated that there is a tradeoff between winning and likeability. This effect is not surprising from a human perspective; losing is generally not pleasant and may negatively influence perceptions of a partner. Furthermore, negative perceptions may lead to tougher bargaining in the future ("tit for tat" strategies). However, since repeated negotiations by necessity require balancing winning across multiple interactions in order to maximize payout, we examined this relationship in detail.

We performed regression analysis to examine the relationship between "winning" (as measured by the point lead) and likeability. We found a highly significant effect such that increased point leads led to reduced overall likeability ($t = -2.692$, $N=240$, $p < 0.01$). Since increased points are tied to discovery of "integrative potential" within games (or "win-win" scenarios), we also examined the effects of joint points discovered. We likewise found a significant, positive correlation between liking the agent and discovering joint points ($t = 4.016$, $N=240$, $p < 0.001$). Entering all three of these variables into a regression analysis found that both main effects held, and there was an interaction between joint points and likeability such that in negotiations where joint value was discovered, winning was correlated with lower likeability, but where joint value was *not* discovered, winning was correlated with higher likeability (see Fig. 2). The main effect of liking on point lead controlling for joint points was still negative and significant ($t = -2.110$, $N=240$, $p = 0.036$). Similarly, we found the main effect of joint points on point lead controlling for likeability was also negative and significant ($t = -2.513$, $N=250$, $p = 0.013$). Finally, the interaction term was significant ($t = -2.385$, $N=240$, $p = 0.018$).

V. DISCUSSION & CONCLUSIONS

The results of this competition confirm the complexity of strategic tradeoffs in repeated negotiations. We isolate agent likeability (as reported by their human partners) as being particularly significant in understanding human-agent outcomes over time. Admittedly, these exploratory results do not put to rest questions over the *causal* relationship between likeability and scoring well (i.e., is not being well-liked an indicator that you are winning, or does winning cause unhappiness and a lack of likability?). Such is the cost of competition-based analyses, wherein the exact strategies of the submitted agents must needs remain somewhat "black-box". But these results do indicate that there is a tradeoff, and that the tradeoff is also dependent on the joint points earned by both parties. Since all of the negotiations in this study were at least partly integrative, we can look at joint points as a proxy for integrative potential discovered.

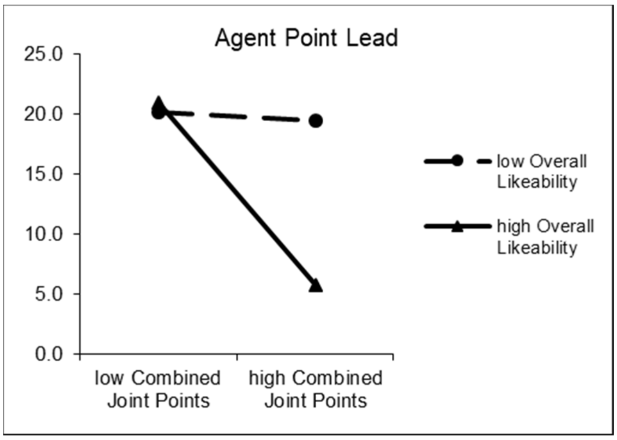


Fig. 2. Winning-Likeability Tradeoff, Interaction with Joint Points

Per Fig. 2, we can acknowledge several expected results: when integrative potential is discovered (high joint points), agents are well-liked when they have a small lead, and not well-liked when they are far ahead. But, unusually, when the outcome is distributive (there are low joint points), there is no correlation between winning and likeability. There are several potential explanations for this effect that warrant further follow-up beyond these early speculations. First, it is possible that people who fail to discover integrative potential are simply less aware that they are losing and are therefore less likely to blame their partner via a dip in likability. Alternatively, it may be that people are simply less concerned with fairness when integrative potential has not been grown. In short, people are more likely to be angered if the partners "grow the pie" but then the pie is stolen, than if the pie fails to grow at all—they are more accepting of a lopsided/unfair solution in this case. There has been some prior work that indicates this latter view may hold some merit [15].

Still, this result seems to create an unfortunate strategic catch-22: since discovering integrative potential tends to lead to more points, but claiming that value leads to low likeability, there may be no way to do well over time. However, the Equalist agent finds a loophole to this problem. By engaging in cross-game logrolling, Equalist manages to seek largely distributive but unfair solutions in each *individual* game. This doesn't appear to come at a hit to likeability, and Equalist is able to exchange favors across games to ensure it still comes out ahead (particularly in Negotiation 2). Equalist does end up with more than its fair share, but this may be indicative of human behavior with regards to favors: they are able to exchange favors but have trouble keeping track of the exact magnitude owed. That way, Equalist does *somewhat* poorly in Negotiation 1 (when it gives a favor) but does *extremely* well in Negotiation 2 (when the favor comes due). Previous work does indicate that people are perfectly capable of understanding reciprocity, but we posit that they may, in human-agent scenarios, fall back on heuristics that do not fully capture the complexities of favor exchange. Indeed, such subtleties are often the subject of advanced negotiation training courses. Of course, we emphasize that Equalist does *not* succeed against every user; systems and studies examining individual differences remain highly valuable.

In general, these results further expand the picture of human-agent negotiating behavior, and provide more insight into the tradeoffs between winning and likeability. Several of the submitted agents include improvements to the state-of-the-art for human agent negotiation, and we hope to continue to push the boundaries of socially-aware agents through the human-agent league of ANAC.

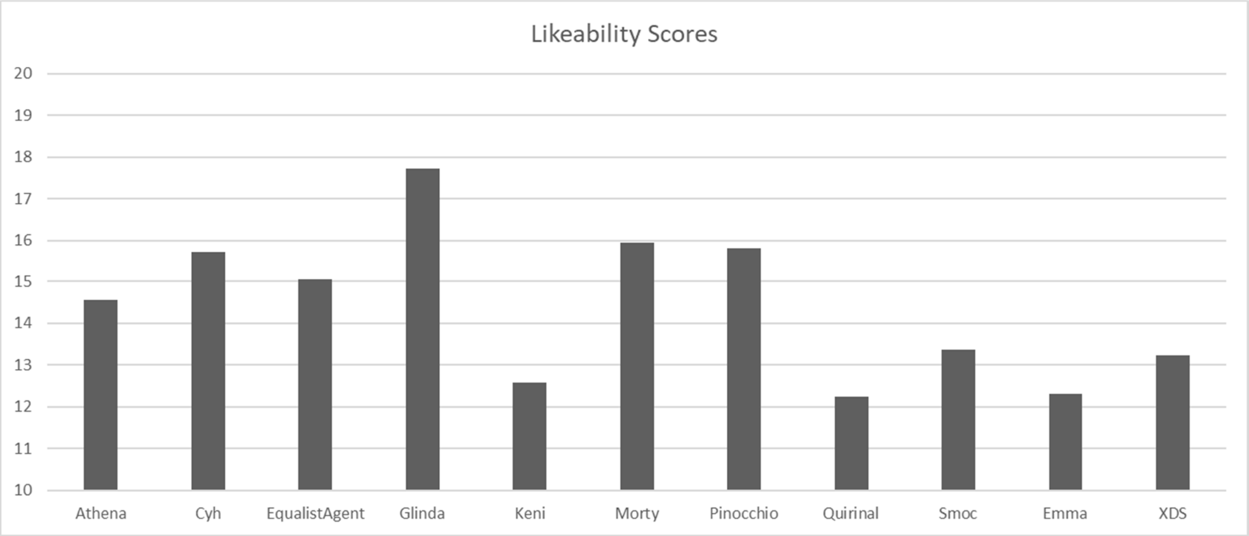


Fig. 3. Total Likeability Score (Summed, three Negotiations, 7-Point Likert)

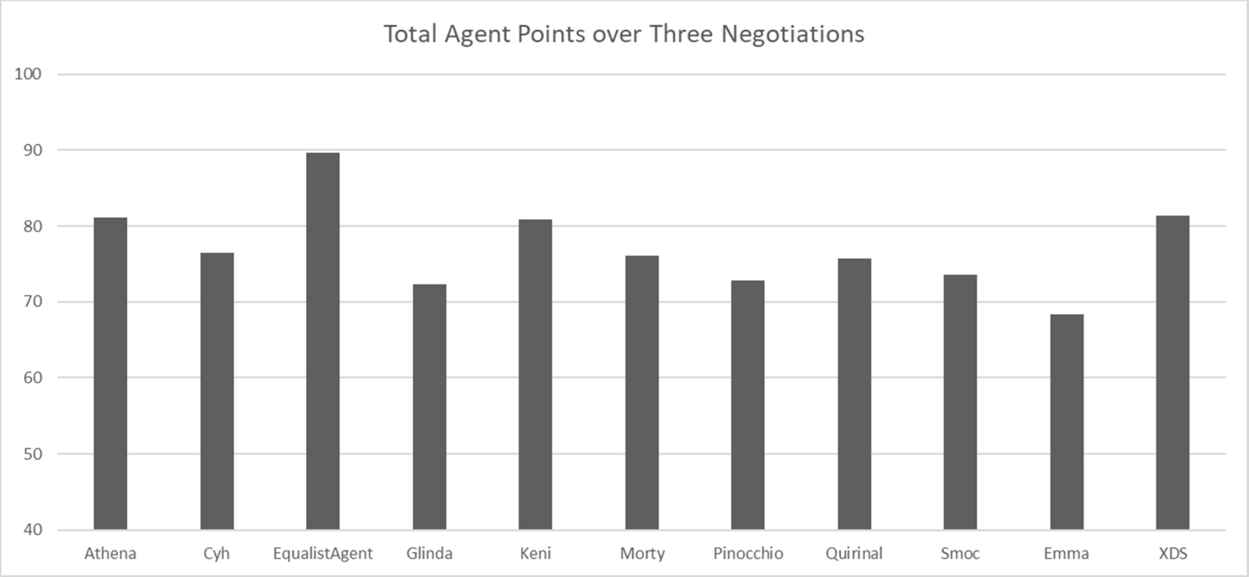


Fig. 4. Total Agent Score (Summed three Negotiations)

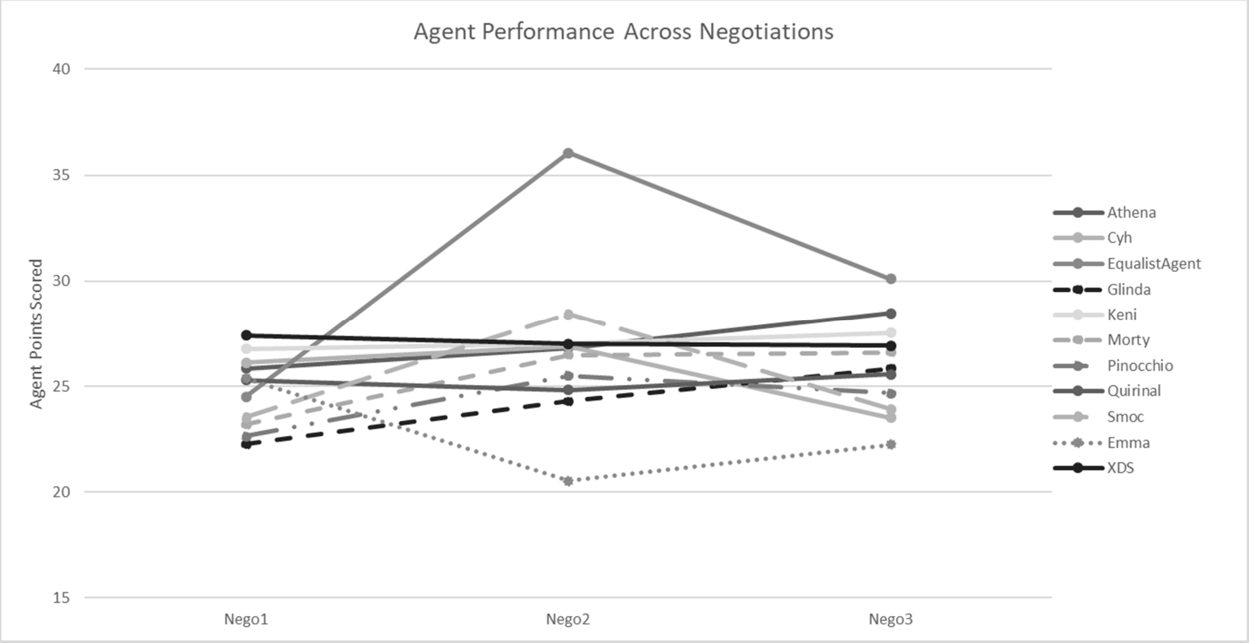


Fig. 5. All Agent Performance across All Three Negotiations Individually

ACKNOWLEDGMENT

We wish to acknowledge the support of the ANAC Board in allowing use of their anonymized data, as well as the competition participants, without whom this would not be possible. This work is part of the Veni research program with project number 639.021.751, which is financed by The Dutch Research Council (NWO). This work is supported by the U.S. Air Force Office of Scientific Research, and the opinions expressed herein should not be inferred to be endorsed by the U.S. Government.

REFERENCES

[1] Kathleen M. Alto, Keiko M. McCullough, and Ronald F. Levant, 2018. “Who is on Craigslist? A novel approach to participant recruitment for masculinities scholarship”. *Psychology of Men & Masculinity*, 19(2), p.319.

[2] Baarslag, T., 2014. What to bid and when to stop (Doctoral dissertation, Delft University of Technology).

[3] Baarslag, T., Kaisers, M., Gerding, E., Jonker, C. M., & Gratch, J. (2017). When will negotiation agents be able to represent us? The challenges and opportunities for autonomous negotiators. 26th International Joint Conference on Artificial Intelligence. Melbourne, Australia.

[4] Joost Broekens, Maaike Harbers, Willem-Paul Brinkman, Catholijn M. Jonker, Karel Van den Bosch, and John-Jules Meyer, 2012, September. Virtual reality negotiation training increases negotiation knowledge and skill. In *International Conference on Intelligent Virtual Agents* (pp. 218-230). Springer, Berlin, Heidelberg.

[5] Jonathan Gratch, David DeVault, Gale M. Lucas, and Stacy Marsella, 2015, August. Negotiation as a challenge problem for virtual humans. In *International Conference on Intelligent Virtual Agents* (pp. 201-215). Springer, Cham.

[6] Catholijn M. Jonker, Reyhan Aydogan, Tim Baarslag, Katsuhide Fujita, Takayuki Ito, and Koen V. Hindriks, 2017. Automated Negotiating Agents Competition (ANAC). In *AAAI* (pp. 5070-5072).

[7] Kelley, H. H. (1966). A classroom study of the dilemmas in interpersonal negotiations. Strategic interaction and conflict, 49, 73.

[8] Koeman, V.J., Hindriks, K.V., Gratch, J. and Jonker, C.M., 2019, May. Recognising and Explaining Bidding Strategies in Negotiation Support Systems. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems* (pp. 2063-2065). International Foundation for Autonomous Agents and Multiagent Systems.

[9] Raz Lin, Yinon Oshrat, and Sarit Kraus, 2009, May. Investigating the benefits of automated negotiations in enhancing people's negotiation skills. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1* (pp. 345-352). International Foundation for Autonomous Agents and Multiagent Systems.

[10] Mell, J., Gratch, J., Baarslag, T., Aydoğran, R. and Jonker, C.M., 2018, November. 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.

[11] Mell, J., Lucas, G. and Gratch, J., 2015, May. An effective conversation tactic for creating value over repeated negotiations. In Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (pp. 1567-1576). International Foundation for Autonomous Agents and Multiagent Systems.

[12] Johnathan Mell, and Jonathan Gratch, 2017, May. Grumpy & Pinocchio: Answering Human-Agent Negotiation Questions through Realistic Agent Design. *In Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems* (pp. 401-409). International Foundation for Autonomous Agents and Multiagent Systems.

[13] Johnathan Mell, and Jonathan Gratch, 2016, May. IAGO: interactive arbitration guide online. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems* (pp. 1510-1512). International Foundation for Autonomous Agents and Multiagent Systems. IAGO

[14] Nadler, J., 2003. Rapport in negotiation and conflict resolution. Marq. L. Rev., 87, p.875.

[15] Zahra Nazari, Gale M. Lucas, and Jonathan Gratch, 2017, August. Fixed-pie Lie in Action. In *International Conference on Intelligent Virtual Agents* (pp. 287-300). Springer, Cham.

[16] Patton, B. (2005). Negotiation. The Handbook of Dispute Resolution, Jossey-Bass, San Francisco, 279-303.

[17] Raiffa, H. (1982). The art and science of negotiation. Harvard University Press.

[18] Gerben A. Van Kleef, Carsten KW De Dreu, and Antony SR Manstead, 2004. The interpersonal effects of anger and happiness in negotiations. *Journal of personality and social psychology*, 86(1), p.57.

[19] Zacharia, G. and Maes, P., 2000. Trust management through reputation mechanisms. Applied Artificial Intelligence, 14(9), pp.881-907.