Towards Emotion-Aware Agents for Negotiation Dialogues

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY-SA 4.0

SUBMISSION DATE / POSTED DATE

26-02-2022 / 03-03-2022

CITATION

Chawla, Kushal; Clever, Rene; Ramirez, Jaysa; Lucas, Gale M.; Gratch, Jonathan (2022): Towards Emotion-Aware Agents for Negotiation Dialogues. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.19242939.v1>

DOI

[10.36227/techrxiv.19242939.v1](https://doi.org/10.36227/techrxiv.19242939.v1)

Towards Emotion-Aware Agents for Negotiation Dialogues

Kushal Chawla, Rene Clever, Jaysa Ramirez, Gale M. Lucas,
and Jonathan Gratch, *Senior Member, IEEE*

(Invited Paper)

Abstract—Negotiation is a complex social interaction that encapsulates emotional encounters in human decision-making. Virtual agents that can negotiate with humans are useful in pedagogy and conversational AI. To advance the development of such agents, we explore the prediction of two important subjective goals in a negotiation – outcome satisfaction and partner perception. Specifically, we analyze the extent to which emotion attributes extracted from the negotiation help in the prediction, above and beyond the individual difference variables. We focus on a recent dataset in chat-based negotiations, grounded in a realistic camping scenario. We draw extensive qualitative and quantitative comparisons between three types of emotion variables – *emoticons*, as well as *lexical* and *contextual* variables, by leveraging affective lexicons and a state-of-the-art deep learning architecture. To further validate the findings, we analyze the prediction of these subjective negotiation goals after controlling for the objective performance of the participants. We also study the temporal effects, understanding the contribution of emotion expressed in the initial and latter parts of the conversation. Finally, we discuss our insights, which will be helpful for designing adaptive negotiation agents that interact through realistic communication interfaces.

Index Terms—Negotiation dialogues, Human-Agent Interaction, Emotion Models, Outcome Prediction, Regression

1 INTRODUCTION

NEGOTIATION is a core task for studying emotional feelings and expressions in human decision-making [1]. Being a mixed-motive task, it creates both interpersonal and intrapersonal conflicts for the negotiators. Motivational tensions often arise when negotiators pit aspirations for individual accomplishments against the demands of sustaining social connections. This leaves difficult decisions for the negotiators about working towards their own self-oriented outcomes or making sacrifices for others. Such situations can be fraught with emotional encounters. For instance, as a negotiator strives to get as much as possible for themselves, in doing so, they need their partner to go along as well. However, the partner's willingness to go along is essentially an emotional decision. A negotiator that tries to take too much can annoy their partner, and in turn, hurt their likeability in the eyes of their partners and also the partner's affective evaluation of the outcome (that is, their satisfaction). Instead, it is desirable for the negotiator to strive for maximum performance while ensuring that the partner is satisfied [2] and leaves with a positive perception of the partner [3], [4]. Therefore, predicting the partner's satisfaction and perception in advance can be crucial for an AI assistant that aims to negotiate with its users. These virtual negotiation agents find broad applications in advancing conversational AI such as the Google Duplex agent that books a haircut appointment over the phone [5]. Further,

such agents can also be useful in increasing access to social skills training [6].

A number of prior studies rely on user individual difference attributes to explain negotiation behaviors and outcomes. For instance, this includes studies based on gender [7] and personality traits such as Social Value Orientation and Machiavellianism [8]. This research is crucial from the perspective of practical negotiation agents that interact with users on social media platforms since research suggests that demographic and personality attributes can be inferred from past user interactions [9], [10], [11]. Thus, they can be used to adapt the behavior of the deployed automated systems. However, relying only on these variables misses out on all the other available information such as affective attributes in the negotiation itself, which may further help in predicting the outcomes. While one might expect that such affect variables would help merely because they are manifestations of the individual differences, it is instead possible that affective factors reflect the recent interaction with the partner and might predict in their own right. This leads us to our key research question in this work: *For predicting a negotiator's satisfaction and liking for their partners, is there value in leveraging affective attributes from the negotiation, above and beyond the individual difference variables?*

Most prior efforts in negotiations focus on a menu-driven communication between the human and the agent. Such interfaces allow button clicks for sharing offers or preferences, which helps to keep the design tractable. This concreteness, however, comes at a cost – it hinders the analysis of several realistic aspects in a negotiation such as persuasion and free-form emotion expression. Due to the inherent design restrictions, emotion expression merely reduces to the use of emoticons and predefined sentence

- K. Chawla, G. M. Lucas, and Jonathan Gratch are with University of Southern California, Los Angeles, USA.
E-mail: {chawla, lucas, gratch}@ict.usc.edu
- Rene Clever and Jaysa Ramirez were interns at Institute for Creative Technologies, University of Southern California, Los Angeles, USA.
E-mail: rene.clever1@gmail.com and j.ramirez7384@gmail.com

Manuscript received February 28, 2022.

templates [12]. Instead, recent efforts have explored a fundamentally different, chat-based interaction that allows human-agent communication in free-form natural language such as English [13], [14], having the potential to fill the gap between existing emotion research and the challenges in real-world negotiations.

To this end, we analyze a large-scale, linguistically rich dataset [14] that is grounded in a realistic campsite negotiation task. The associated metadata about the participants enables us to study individual differences and linguistic dialogue behaviors in the same setup. Going beyond the capabilities of restrictive menu-driven interfaces, we develop three degrees of emotion recognition techniques: *emoticons*, *lexical*, and *contextual* emotion by leveraging recent advancements in deep learning (Section 3). We draw extensive qualitative and quantitative comparisons between these different methods. Using correlational and step-wise regression analysis, we quantify the extent to which affective attributes help to account for the variance in the outcomes, beyond just the individual user attributes (Section 5). We further validate our results by controlling for the objective performance of the participants in the negotiation and study the temporal effects of emotion expression on the final prediction performance. We conclude by discussing how our analysis can guide the development of automated negotiation agents and briefly summarise the ethical considerations around such systems (Section 6).

2 RELATED WORK

How humans negotiate has been extensively studied in multiple disciplines: Game Theory for identifying optimal behaviors [15], Psychology for understanding human decision-making [16], and Computer Science for building automated negotiation agents [17], [18]. Demographics, personality, and affective behaviors have played a central role in improving our understanding around negotiations [7], [8] and similar decision-making social interactions [19], [20]. For instance, de Melo and colleagues studied human-agent prisoner's dilemma interaction and found that the display of guilt after exploitation can elicit more cooperative responses as compared to an agent that smiled [20].

Both satisfaction and the relationship with the negotiation partner are seen as crucial metrics of negotiation performance. Oliver and colleagues studied how profit expectations relate to satisfaction after the bargaining in buyer-seller interactions, finding that higher expectations had the effect of decreased satisfaction [2]. Maintaining a positive relationship with the partner is especially crucial in repeated interactions, where poor relations in earlier negotiations can adversely impact the results of future ones [4]. Relationship also manifests in the context of rapport building [21], favor exchange [22], and reputation effects [23].

However, most efforts in human-agent negotiations are based on highly structured menu-driven designs such as the IAGO negotiation platform [12]. In many cases, the negotiation is seen as simply an exchange of offers [24]. Towards more realistic modes of communication, limited efforts have studied chat-based dialogue systems, that are fundamentally different than menu-driven systems and

hence, must be studied separately. Initial datasets in this direction looked at game settings [25], [26], which allow free-form language use but the limited semantic context around the issues in such negotiations inhibits any rich personal conversations. He and colleagues focused on more realistic buyer-seller price negotiations for the product listings on Craigslist [13]. These datasets have fueled numerous efforts in developing negotiation dialogue agents [27], [28], [29], [30], [31], [32]. Unfortunately, none of these datasets captured the demographics or the personality variables of the users. Further, all of them (and subsequently the developed dialogue agents) looked only at the objective outcomes of the negotiation performance such as the points scored, the final agreed price, or the optimality of the negotiated deal. Our focus is on the recently released *CaSiNo* dataset that enables linguistically rich conversations in a constrained, tractable environment [14]. The dataset further contains rich metadata of the participants and subjective negotiation outcomes, that we use to develop our measures in this work (Section 3).

Finally, we note that researchers have also looked at face-to-face negotiations [33], including speech and embodied agents that can be natural extensions to our current analysis in chat-based negotiations.

3 DATASET AND METHODS

Our analysis is based on the recently released *CaSiNo* dataset [14]. The design is an instance of Multi-Issue Bargaining Task (MIBT), which is a popular model for studying negotiations computationally [12], [34]. MIBT consists of a negotiation over a fixed set of issues with a predefined priority order for the negotiators. This provides a concrete structure to the negotiation and thus, it helps to keep the conversations focused and the evaluation of the negotiation performance tractable. Nevertheless, it can lead to a diverse set of dialogues, depending on whether the participant preferences align or not, making it suitable for a variety of applications such as in social skills training [6] or building artificial assistants [24]. In addition, the *CaSiNo* task is grounded in a real-world campsite scenario that promotes contextually-rich personal conversations in free-form natural language. Unlike prior datasets in chat-based negotiation systems [13], [25], [26], this design enables the study of a number of realistic aspects in negotiations such as persuasion and emotion expression in a constrained, tractable environment.

3.1 Participants

A total of 846 unique participants were recruited on Amazon Mechanical Turk for collecting the *CaSiNo* dataset¹. The participant pool was restricted to the United States. They were paid for their time and also with a performance-based bonus. Among the participants, 472 identified themselves as *Female*, 372 were *Male*, and 2 belonged to *Other* category. 73.9% of the participants were *White American*, 1.5% were *Native or Islander*, 8.7% were *Asian American*, 8.6% were *Black or African American*, 4.1% were of *Hispanic or Latino* origin, and remaining 3.1% belonged to *Other* category.

1. This research was approved by the Institutional Review Board at the University of Southern California, under Study ID: UP-14-00684.

The screenshot displays the CaSiNo data collection interface. On the left, there is a table for negotiating items:

Value	Item	Packages you get	Packages they get
High	Food	2	1
Medium	Firewood	3	0
Low	Water	0	3

Below the table is a 'Submit Deal' button and a 'Walk Away (Ends Negotiation)' button. On the right, a chat interface shows a dialogue:

You: Well you guys did say you did not care much about large fires. What if you gave all the firewood in replace for the water and you can still keep 1 food?

Fellow Turker: So I would get 3 water and 1 food and youd get 3 firewood and 2 food?

You: Yea that seems like an alright trade to me 😊

Fellow Turker: Hmm... alright then 😊

At the bottom, there is a text input field 'Please enter here...' with a 'Send' button and a row of emoticon icons.

Fig. 1: A screenshot from the CaSiNo data collection interface. Taken from [14]. Please refer to the original dataset paper for more screenshots and sample negotiation dialogues.

3.2 Design and Procedure

The CaSiNo dataset contains a total of 1030 negotiation dialogues in English. Each data point is rich with the participant metadata that includes demographics and personality traits. These are collected before the negotiation begins. Further, each participant is asked to watch a negotiation training module. This video module covers some of the best practices in negotiations and also provides an overview of the data collection interface. In each dialogue, two participants take the role of campsite neighbors and negotiate for additional packages of camping essential items. Following the MIBT structure, there are three issues: *Food*, *Water*, and *Firewood*. Each issue has an available quantity of three, thereby resulting in four *levels* from 0 to 3. Before the negotiation begins, each participant is randomly assigned a priority order for the three issues, which is a permutation of $\{High, Medium, Low\}$. Once the priorities are assigned, the participants also come up with their own justifications for needing or not needing a particular issue such as needing more water for a planned hike or food for the kids. This exercise *prepares* the participants for their upcoming negotiation, and subsequently results in the collection of linguistically rich conversations.

The dialogues in CaSiNo contain 11.6 utterances with 22 tokens per utterance, on average. The participants were also allowed to use emoticons for four basic emotions: *Joy*, *Sadness*, *Anger*, and *Surprise*. We provide a screenshot from the chat interface in Figure 1. Once the negotiation is over, the researchers record the performance of the participants based on the number of items that they are able to negotiate for. Alongside, the participants are also asked to report their satisfaction with the outcome, and their perception of the negotiation partner. Analyzing these subjective measures of negotiation performance is the primary focus of this work. We refer the readers to the original CaSiNo paper [14] for more information on the data collection procedure and a sample dialogue from the dataset.

3.3 Measures

We cluster the variables into three categories: 1) *Individual differences*, which captures the demographics and personality of the participants, 2) *Affect Variables*, which comprises all the affective attributes extracted from the negotiation

utterances, and 3) *Negotiation Outcomes*, where we describe the two primary dependent variables in our analysis.

3.3.1 Individual Differences

These variables are based on the responses reported by the participants in a survey before their negotiation. The participants self-identified their demographic attributes while the personality traits are based on the standard tests from the psychology literature.

Demographics: There are two continuous variables: *Age*² and *Education*. We encoded *Education* as a continuous variable, leveraging the inherent order in the highest level of education. It takes a value from 0 to 8, with an increasing level of education.

Further, there are two discrete demographic variables: *Gender* and *Ethnicity*. For our regression analysis, we dummy-coded these variables based on the categories discussed earlier.

Personality: There are two available measures of individual personality differences: Social Value Orientation [35] and the Big-5 personality traits [36]. These personality attributes have been extensively studied in the context of negotiation research [37], [38]. The Social Value Orientation (SVO) is defined as the stable preferences for certain patterns of outcomes for oneself or others [39]. A participant can either be categorized as *Prosocial*, that tend to be cooperative in their interactions, or *Proself*, that tend to serve their individual interests. In the CaSiNo data collection study, SVO was computed using the Triple Dominance Measure [35]. 463 participants were classified as *Prosocial*, 364 as *Proself*, while 19 were *Unclassified*. The Big-5 test consists of five personality dimensions: *Extraversion*, *Agreeableness*, *Conscientiousness*, *Emotional Stability*, and *Openness to Experiences*. These dimensions were computed using the Ten-Item Personality Inventory [40]. Each dimension takes a value between 1 and 7.

3.3.2 Capturing Affect

Natural language based negotiations provide exciting avenues for Affective Computing research. Unlike menu-driven systems, where communication is highly restricted,

2. One participant reported the age of 3, which we believed to be in error and was removed from all our analysis that uses *Age*.

language allows free expression of emotion and other affective attributes. In order to capture these attributes in CaSiNo, we consider three different degrees of affect recognition techniques: emoticon counts, use of emotive vocabulary based on affect lexicons (specifically LIWC [41]), and utterance-level emotion expressions (based on a pretrained deep learning model called T5 [42]). We discuss these approaches below. We will later compare these methods through our regression analysis for predicting satisfaction and liking.

Emoticons: Emoticons provide a structured way for emotion expression. Due to this reason, emoticons have been heavily used in menu-driven negotiation research. For instance, the IAGO platform for human-agent negotiations allows human participants to use emoticons via button clicks [12]. For CaSiNo, the participants were allowed to use shorthands for typing emoticons directly in the chat interface for four basic emotions: *Joy*, *Sadness*, *Anger*, and *Surprise* (see Figure 1). For a given participant and the negotiation dialogue, we count the number of emoticons used by the participant for each of these types and use that as a continuous measure of emotion expression. Approximately 15% of the utterances in CaSiNo make use of one or more emoticons. Among those that use it, nearly 80% use *Joy* emoticon.

Based on a manual inspection of the utterances, we find that the participants tend to use *Joy* emoticon in a number of scenarios, including small talk at the beginning or end of the conversation when showing agreement, and in some cases, interestingly, also when rolling out a strict offer. *Sadness* tends to be used when showing disagreement, showing empathy, or emphasizing personal needs. Some cases where the participants expressed *Surprise* are when pointing out strange behavior from their partner or showing empathy towards a specific need of their partner. *Anger* is used in the cases of strong disagreement where for instance, the partner does not empathize with the personal requirements of the participants or when they receive a very unfair offer that largely benefits the negotiation partner.

LIWC: We now go beyond the capabilities of menu-driven systems by extracting emotion attributes from the text in the utterances of the participants. Specifically, we look at word-level expressions by leveraging affect lexicons from the literature. Specifically, we make use of the Linguistic Inquiry and Word Count lexicon, popularly referred to as LIWC in the literature [41]. LIWC consists of a number of word vocabularies corresponding to everyday use that reflect our feelings and personality. Given our focus on affective attributes, we use four LIWC dimensions in this work: *Positive Emotions*, *Sadness*, *Anger*, and *Anxiety*. To compute the corresponding scores for a participant in a given dialogue, we simply count the number of words in the utterances of the participant that are present in the corresponding LIWC lexicons.

As one might expect, we find a number of instances where the participants express emotions without explicitly making use of emoticons. For instance, an excerpt where the participant clearly expresses anger without using the *anger* emoticon: ‘Are you mad. with out water what we do...’. Hence, with LIWC lexicons, we are able to capture more emotive utterances than by simply looking at the emoticon

usage. Nearly 60% of the utterances make use of at least one emotive word from our four LIWC categories.

Since LIWC variables are based on word usage, they provide a highly interpretable way of extracting emotions directly from the text. However, merely relying on word-level traits can still miss out on phrase-level or utterance-level emotion expressions. Hence, we next discuss how we make use of a state-of-the-art deep learning model from Natural Language Processing (NLP) to identify emotion in CaSiNo dialogues.

Utterance-level emotion: Large deep learning models based on the highly parallelizable Transformer architecture have gained enormous popularity in the field of NLP in recent years [43]. These models have achieved huge success on a variety of tasks ranging from sentiment analysis to machine translation, fundamentally changing the way researchers approach these tasks. The primary way of training a model for a specific task consists of two steps: 1) *Pre-train* a large language model with freely available text data on the web in an unsupervised or a self-supervised manner, and then 2) *Fine-tune* the model on the task-specific dataset based on supervised learning.

Leveraging these recent advancements, we make use of the Text-to-Text Transfer Transformer (T5) model [42]. T5 is an encoder-decoder architecture that is pretrained on a variety of NLP tasks by converting all the tasks in the same text-to-text format. No matter the task, the input to the model is a text sequence along with a task identifier prefix, and the target output is also a text sequence. Classification tasks can also be converted in this way by simply using a target output sequence that consists of a single word corresponding to the ground-truth label. This approach swiftly unifies a number of tasks, paving the way for parameter sharing throughout the model architecture. T5 achieves state-of-the-art performance on a number of NLP tasks such as classification, text summarization, and question answering.

For our purpose, we first pre-process the utterances in CaSiNo by removing all the emoticons. We then use the publicly available T5 model³ for making predictions. This model is *fine-tuned* on an emotion recognition dataset based on social interactions on Twitter [44]. Given an input text sequence, the model classifies it into one of the six classes: *Joy*, *Love*, *Sadness*, *Fear*, *Anger*, and *Surprise*. The model performs well on the test set of the emotion data, achieving an accuracy of 93% and a macro F1 score of 90%.

The output prediction of the model is based on the confidence scores that it generates for all the emotion labels, such that the label with the highest confidence is chosen as the prediction. In our case, since we are primarily interested in extracting the emotion scores for the utterances, we do not rely on the final predictions of the model. Instead, we directly use these confidence scores (probability values between 0 and 1) as the corresponding scores for each label. Given a negotiation dialogue and a participant, we simply sum up these confidence scores for all the participant’s utterances and use these as the continuous measures for the six emotion dimensions. We refer to these dimensions (and interchangeably, this method) as *T5-Emotion* in this paper.

3. <https://huggingface.co/mrm8488/t5-base-finetuned-emotion>

All the three methods described above capture the emotions expressed in CaSiNo dataset in fundamentally very different ways. To gain further insight into these methods, in Section 4, we will analyze the relationship between them through a combination of quantitative and qualitative techniques. First, we complete our description for measures by elaborating on the subjective negotiation outcomes.

3.3.3 Negotiation Outcomes

We focus on two outcome variables: 1) Satisfaction (How satisfied are you with the negotiation outcome?), and 2) Liking (How much do you like your opponent?). Both variables were self-reported by the participants after the negotiation, using a 5-point Likert scale. We coded the responses from 1 to 5 to create the continuous measures. We present the means and standard deviations for all the continuous variables in Table 4.

4 ANALYSIS OF EMOTION VARIABLES

Before presenting our results for predicting participant satisfaction and liking, we first validate *T5-Emotion* measures by understanding how they relate to *Emoticon* and *LIWC*.

4.1 Correlation among emotion dimensions

Table 1 presents the correlations of *T5-Emotion* measures among each other and with *Emoticons* and *LIWC* dimensions. For this analysis, we only focus on the dimensions that appear in all three models: *Joy*, *Sadness*, and *Anger*. For *LIWC*, we incorporate the *Positive* dimension. In general, we find significant positive correlations of the emotion dimensions with their own counterparts across the three models, which aligns well with our expectations and validates the model predictions on the CaSiNo dataset. We observe that although significant, most correlation values are weak to moderate, indicating the fundamental differences in what is captured by these three models. We also note that *Anger* shows significant positive correlations with *Sadness* as well. Since these measures are based on the confidence scores of the model, this suggests that the model might get confused between these two negative emotion dimensions.

4.2 Lexical correlates of emotion

To gain more insights, we perform additional qualitative analysis by generating lexical correlates for the emotion measures. We assigned an emotion label to each utterance for all three methods. For *Emoticons* and *LIWC*, this was based on the label that is reflected in the majority in the utterance. For *T5-Emotion* dimensions, we simply used the final predictions from the model. We next compute the log-odds ratio, informative Dirichlet prior [45] of all the tokens for each emotion dimension relative to all other dimensions. This allows us to study the associations between the tokens and each emotion dimension. We present the top 5 words in each category in Table 2 along with scaled Z-scores. We observe a number of highly significant associations across all categories that align with our intuition such as words like ‘awesome’, ‘great’, ‘thank’ for *Joy* and *Positive Emotions* categories. We found these words to be common in the beginning and end of the negotiation where participants

engage in small talk, resulting in such utterances being classified as *Joy* or *Positive*. Further, the associations are higher for *LIWC* and *T5-Emotion* which, as we observe, are able to recognize many other emotive utterances. Although significant, we observe that some words are less meaningful such as in the case of *Anger*. This suggests that some emotion dimensions may require more utterance-level context for better interpretation.

4.3 Sample Predictions

For a more contextual analysis, we explicitly observe the sample predictions for *T5-Emotion*. We still focus on the three dimensions that are the most common and are present in all three emotion methods. Table 3 presents a few of the most confident predictions by the model. To analyze the utterances that are specifically captured by the deep learning model, we only look at the utterances that are *undetected* by the other two methods. As the table depicts, the model is able to capture contextual emotion, beyond just the lexical or emoticon usage. Based on these observations, we expected the affective attributes extracted using the deep learning model to be better predictors than *Emoticons* and *LIWC* measures.

5 RESULTS

We now present our main results.

5.1 Correlation with outcomes

Table 4 summarizes the correlations of all the continuous measures defined in Section 3 with the negotiation outcomes. Overall, we observe a number of significant trends across individual differences and emotion variables. We find that satisfaction and partner liking both improve with age ($r=0.110$, $p < 0.01$). The outcomes are also positively correlated with Agreeableness and Emotional Stability ($ps < 0.01$). We observe that positive emotions tend to relate positively with both the outcomes, showing significant trends for liking in the case of *Emoticon* usage ($r=0.075$, $p < 0.01$) and *LIWC* ($r=0.061$, $p < 0.01$). We do not observe similar results for *T5-Emotion*. This may be due to the inherent bias that the deep learning model shows towards *Joy* label, reducing its precision. However, trends show significant negative correlations for both *Sadness* and *Anger* that are usually expressed when the negotiation is not going favorably. This is naturally associated with lower satisfaction of the participants and liking for their negotiation partners.

Among the discrete variables, we find that females are significantly more satisfied than males ($t=3.6$, $p < 0.001$) and significantly like their negotiation partners more ($t=4.1$, $p < 0.001$). We do not observe any significant associations with SVO and Ethnicity, based on t-test and one-way ANOVA respectively.

5.2 Regression Analysis

We now discuss the results of our regression analysis that answers whether emotion variables extracted from the negotiation dialogue collectively explain more variance in satisfaction and liking, above and beyond the individual

TABLE 1: Correlations of *T5-Emotion* variables (a) among each other, and (b) with corresponding *Emoticon* and *LIWC* variables. We only consider the dimensions that are common across all methods. * denotes $p < 0.05$. ** denotes $p < 0.01$.

(a)

	Joy	Sadness	Anger
Joy	1	-.288**	-.295**
Sadness	-.288**	1	.277**
Anger	-.295**	.277**	1

(b)

T5-Emotion	Emoticon Usage			LIWC Variables		
	Joy	Sadness	Anger	Positive Emotions	Sadness	Anger
Joy	.062**	-.043	-.040	.387**	-.125**	-.106**
Sadness	-.010	.213**	.086**	-.033	.550**	.033
Anger	-.052*	.113**	.111**	-.185**	.114**	.169**

TABLE 2: Top 5 words based on log-odds ratio for (a) *Emoticon* (b) *LIWC* and (c) *T5-Emotion* methods. * denotes $p < 0.05$. ** denotes $p < 0.01$. *** denotes $p < 0.001$.

(a)

Joy		Sadness		Anger		Surprise	
Word	Z-score	Word	Z-score	Word	Z-score	Word	Z-score
hope	2.89**	sorry	2.10*	animal	1.36	drive	1.43
buddy	2.72**	spilled	1.84	unacceptable	0.98	bargain	1.43
awesome	2.71**	n't	1.82	covid-19	0.84	awful	1.32
great	2.60**	reconsider	1.38	word	0.84	function	1.32
thank	2.53*	stopped	1.30	joking	0.84	snack	1.23

(b)

Positive Emotions		Sadness		Anger		Anxiety	
Word	Z-score	Word	Z-score	Word	Z-score	Word	Z-score
good	11.89***	sorry	9.93***	hate	2.08*	worried	7.18***
well	11.87***	low	6.16***	greedy	1.58	afraid	4.07***
okay	9.90***	alone	4.43***	critical	1.35	worry	3.5***
great	9.86***	sugar	2.44*	tricky	1.19	risk	3.22**
ok	9.68***	hear	2.43*	bother	1.06	confused	2.54*

(c)

Joy		Love		Sadness		Fear		Anger		Surprise	
Word	Z-score	Word	Z-score	Word	Z-score	Word	Z-score	Word	Z-score	Word	Z-score
you	8.62***	hot	6.4***	sorry	10.63***	worried	8.72***	cold	11.5***	funny	1.93
good	7.68***	generous	3.04**	unfortunately	4.68***	afraid	7.05***	no	5.76***	interesting	1.89
great	7.33***	hotter	1.89	dehydrated	4.03***	concerned	4.48***	need	4.77***	surprise	1.08
hello	6.84***	lovely	1.56	low	3.12***	dark	3.88***	thirsty	4.63***	ha	1.03
sounds	6.81***	liking	1.34	suffer	2.98**	scared	3.58***	hungry	4.36***	status	0.85

TABLE 3: High confidence sample predictions for *T5-Emotion* model that went *undetected* for both *Emoticon* and *LIWC* methods.

Prediction	Sample Utterances
Joy	1) I think that sounds reasonable to me. 2) I can make that deal work. 3) You sound very resourceful! Sounds like we have a deal then, yes?
Sadness	1) I feel like this deal just keeps getting worse for me. I won't take less than 2 waters. 2) That is unfortunate. How about I give you all my firewood for all your food? 3) Oh, that would make things difficult!
Anger	1) You are not getting everything, thats just selfish 2) I am about to walk, you are being so unfair 3) You gave me one on each, thats unfair man!

the following set of variables: 1) Individual differences, 2) Affect variables of the participant, and 3) Affect variables for the negotiation partner. The affect variables come from one of the emotion methods defined in Section 3, which also helps to compare them. As discussed earlier, these variables only encompass the information about the negotiation that would be entirely *visible* to an automatic negotiation agent, either explicitly from the negotiation dialogue, or implicitly by inferring the individual difference attributes from past online behaviors of the users on social media platforms. Hence, our insights from such an analysis can practically aid in designing sophisticated negotiation agents that incorporate user satisfaction and their perception of the agent itself in their modeling.

Table 5 and 6 summarize the results for satisfaction and liking respectively. Overall, we observe that the predictions are highly significant with just using the individual difference variables. However, incorporating the affective attributes of the participant and the negotiation partner explains significantly more variance.

difference variables. To achieve this, we perform regression with three steps where each subsequent step incorporates

TABLE 4: Statistics and correlations with outcome variables for all continuous measures. * denotes $p < 0.05$. ** denotes $p < 0.01$.

Variable	Mean	Std.	Correlations	
			Satisfaction	Liking
Individual Differences				
Age	36.97	10.81	.110**	.158**
Education	5.23	1.67	-.004	-.009
Extraversion	3.69	1.70	.055*	.060**
Agreeableness	5.27	1.29	.067**	.099**
Conscientiousness	5.60	1.26	.026	.046*
Emotional Stability	4.91	1.56	.062**	.087**
Openness to Experiences	5.04	1.31	.027	.042
Emoticons				
Joy	.77	1.24	.040	.075**
Sadness	.13	.42	-.151**	-.170**
Anger	.02	.18	-.074**	-.108**
Surprise	.07	.27	-.062**	-.059**
LIWC				
Positive Emotions	5.58	3.36	.033	.061**
Sadness	.22	.52	-.057**	-.091**
Anger	.05	.25	-.047*	-.043
Anxiety	.18	.52	-.040	-.014
T5-Emotion				
Joy	3.89	1.09	.010	.026
Love	.22	.21	-.010	-.030
Sadness	.35	.42	-.141**	-.172**
Fear	.43	.43	-.121**	-.117**
Anger	.81	.66	-.217**	-.295**
Surprise	.03	.08	-.011	-.016
Negotiation Outcomes				
Satisfaction	4.17	1.03	1	.702
Liking	4.11	1.12	.702	1

For satisfaction, the individual difference variables account for minimal yet significant variance ($F(14, 1997)=3.46$, $p < .001$, $R^2=.024$). Adding the participant's affect variables based on the *T5-Emotion* model in the second step, helps to account for a much higher variance ($F(20, 1991)=10.41$, $p < .001$, $R^2=.095$), such that this increase in the proportion is itself highly significant ($\Delta F(6, 1991)=26.02$, $p < .001$, $\Delta R^2=.071$). Yet, further variance is explained when the partner's affect variables are incorporated ($F(26, 1985)=10.88$, $p < .001$, $R^2=.125$; $\Delta F(6, 1985)=11.38$, $p < .001$, $\Delta R^2=.030$).

Similar trends can be observed for liking as well. Individual difference variables alone account for a significant proportion ($F(14, 1997)=6.05$, $p < .001$, $R^2=.041$) but adding the participant's *T5-Emotion* attributes further improves the prediction ($F(20, 1991)=18.16$, $p < .001$, $R^2=.154$) with a significant increase ($\Delta F(6, 1991)=44.58$, $p < .001$, $\Delta R^2=.114$). Finally, when the partner's attributes are also incorporated as predictors, this again shows significant improvements in the explained variance ($F(26, 1985)=19.07$, $p < .001$, $R^2=.200$; $\Delta F(6, 1985)=18.83$, $p < .001$, $\Delta R^2=.046$).

Lastly, we note that in line with our linguistic analysis, the ability for *T5-Emotion* to capture more contextual emotive utterances helps to account for more variance, as compared to *Emoticons* and *LIWC* independently. In our analyses, we also observed that even when *Emoticons* and *LIWC* are combined, just using *T5-Emotion* still achieves much better results. We further explore this comparison between emotion recognition techniques in Section 5.4, where we analyze these methods against each other.

5.3 Controlling for the Objective Performance

Apart from the *subjective performance* metrics of satisfaction and liking, the CaSiNo dataset also provides information on

TABLE 5: Regression results for predicting Satisfaction. ** denotes $p < 0.01$. *** denotes $p < 0.001$.

Variables	R ²	df	F	R ² Change	F Change
Affect Variables: Emoticons					
Individual Differences	.024	(14, 1997)	3.46***	—	—
+Participant Affect	.051	(18, 1993)	5.92***	.027	14.20***
+Partner Affect	.063	(22, 1989)	6.04***	.012	6.34***
Affect Variables: LIWC					
Individual Differences	.024	(14, 1997)	3.46***	—	—
+Participant Affect	.032	(18, 1993)	3.61***	.008	4.07**
+Partner Affect	.051	(22, 1989)	4.85***	.019	10.14***
Affect Variables: T5-Emotion					
Individual Differences	.024	(14, 1997)	3.46***	—	—
+Participant Affect	.095	(20, 1991)	10.41***	.071	26.02***
+Partner Affect	.125	(26, 1985)	10.88***	.030	11.38***

TABLE 6: Regression results for the prediction of Liking. *** denotes $p < 0.001$.

Variables	R ²	df	F	R ² Change	F Change
Affect Variables: Emoticons					
Individual Differences	.041	(14, 1997)	6.05***	—	—
+Participant Affect	.080	(18, 1993)	9.68***	.040	21.50***
+Partner Affect	.097	(22, 1989)	9.70***	.017	9.12***
Affect Variables: LIWC					
Individual Differences	.041	(14, 1997)	6.05***	—	—
+Participant Affect	.055	(18, 1993)	6.39***	.014	7.32***
+Partner Affect	.072	(22, 1989)	7.06***	.018	9.56***
Affect Variables: T5-Emotion					
Individual Differences	.041	(14, 1997)	6.05***	—	—
+Participant Affect	.154	(20, 1991)	18.16***	.114	44.58***
+Partner Affect	.200	(26, 1985)	19.07***	.046	18.83***

the *objective performance* of the participants in terms of their points scored at the end of the negotiation. Each item of *High*, *Medium*, and *Low* priority is worth 5, 4, and 3 points respectively. The total points scored by the participants can be computed based on how many items of each kind they are able to negotiate for.

One might argue that the subjective outcomes of satisfaction and liking could be explained by *objective performance* in the negotiation, and, moreover, that the predictive power of the emotion variables seen above (in Section 5.2) is merely a result of the relationship between emotion variables and the objective negotiation performance. That is, all of the emotion the participants express might just reflect how much they are getting in the negotiation, and, if that were the case, then the emotion variables would no longer predict the subjective outcomes when controlling for the points that the participant scored in the negotiation. Alternatively, we propose that emotion expression provides more signals (in terms of satisfaction and liking) than just how much the participants are getting in the agreed deal, and thus, the emotion variables should continue to predict significant variance above and beyond objective performance.

We assess this hypothesis through a modified regression, where we consider the points scored by the participant in the first step, followed by the same three steps used in Section 5.2. Tables 7 and 8 summarize the results. Although points scored itself explains significant variance in both satisfaction and liking metrics, we still observe similar significant trends for incorporating emotion variables, as before.

For satisfaction, the points scored themselves do explain relatively high variance ($F(1, 2010)=343.88$, $p < .001$, $R^2=.146$), but the results still show high significant improvements with the incorporation of individual differences

TABLE 7: Regression results for predicting Satisfaction after controlling for the points scored by the participants. ** denotes $p < 0.01$. *** denotes $p < 0.001$.

Variables	R ²	df	F	R ² Change	F Change
Affect Variables: Emoticons					
Points Scored	.146	(1, 2010)	343.88***	—	—
+Individual Differences	.172	(15, 1996)	27.70***	.026	4.52***
+Participant Affect	.191	(19, 1992)	24.76***	.019	11.51***
+Partner Affect	.199	(23, 1988)	21.52***	.008	5.19***
Affect Variables: LIWC					
Points Scored	.146	(1, 2010)	343.88***	—	—
+Individual Differences	.172	(15, 1996)	27.70***	.026	4.52***
+Participant Affect	.178	(19, 1992)	22.77***	.006	3.71**
+Partner Affect	.191	(23, 1988)	20.43***	.013	7.82***
Affect Variables: T5-Emotion					
Points Scored	.146	(1, 2010)	343.88***	—	—
+Individual Differences	.172	(15, 1996)	27.70***	.026	4.52***
+Participant Affect	.223	(21, 1990)	27.24***	.051	21.75***
+Partner Affect	.241	(27, 1984)	23.34***	.018	7.77***

TABLE 8: Regression results for the prediction of Liking after controlling for the points scored by the participants. *** denotes $p < 0.001$.

Variables	R ²	df	F	R ² Change	F Change
Affect Variables: Emoticons					
Points Scored	.077	(1, 2010)	167.38***	—	—
+Individual Differences	.121	(15, 1996)	18.39***	.045	7.23***
+Participant Affect	.152	(19, 1992)	18.82***	.031	18.10***
+Partner Affect	.165	(23, 1988)	17.14***	.013	7.88***
Affect Variables: LIWC					
Points Scored	.077	(1, 2010)	167.38***	—	—
+Individual Differences	.121	(15, 1996)	18.39***	.045	7.23***
+Participant Affect	.133	(19, 1992)	16.05***	.011	6.51***
+Partner Affect	.146	(23, 1988)	14.78***	.013	7.71***
Affect Variables: T5-Emotion					
Points Scored	.077	(1, 2010)	167.38***	—	—
+Individual Differences	.121	(15, 1996)	18.39***	.045	7.23***
+Participant Affect	.216	(21, 1990)	26.18***	.095	40.22***
+Partner Affect	.251	(27, 1984)	24.62***	.035	15.25***

and subsequent affect variables, achieving the best results with *T5-Emotion* ($F(27, 1984)=23.34$, $p < .001$, $R^2=.241$). In comparison, this is nearly twice the maximum variance explained by not incorporating the points scored into the prediction model, as reported in Section 5.2 ($F(26, 1985)=10.88$, $p < .001$, $R^2=.125$).

For liking, the points scored account for a significant but a much lower variance itself ($F(1, 2010)=167.38$, $p < .001$, $R^2=.077$). One explanation for this effect can be that the reported satisfaction of the participants is geared towards their own outcomes, while their liking is geared towards their relationship with their negotiation partners. While this shows that their overall perception towards their partners can improve with their own negotiation performance, the results can be hugely improved with the incorporation of affect variables, with *T5-Emotion* still proving to be superior ($F(27, 1984)=24.62$, $p < .001$, $R^2=.251$).

As hypothesized, these observations show that the emotion variables continue to predict significant variance above and beyond objective performance, and therefore, provide more signals (in terms of satisfaction and liking) than just how many points the participants are getting in the negotiation. This points to the value of extracting additional variables from the negotiation dialogue to better capture the subjective outcomes in a negotiation. In Section 6, we will briefly discuss a number of other variables suggested by the prior work, which can further improve this performance.

5.4 Comparing Emotion Recognition Methods

Our results indicate that emotion variables based on the deep learning model *T5* better explain the variance in satisfaction and liking compared to those based on other emotion recognition methods. However, till now, we have only compared the different types of emotion variables by looking at their predictive power in separate regression models. Accordingly, it is reasonable to expect that in addition to *T5-Emotion* variables explaining more variance than other types of emotion variables (based on emoticons and *LIWC*), *T5-Emotion* variables might also explain the same variance in outcomes (and more) that *Emoticon* and *LIWC* variables do. That is, the deep learning model might provide unique cues about the subjective outcomes but they also explain the variance that is explained by other methods. Alternatively, it is possible that *Emoticon* and *LIWC* variables might tap into different aspects of satisfaction and liking than *T5-Emotion*. In this case, incorporating *Emoticon* and *LIWC* variables should continue to predict significant variance above and beyond the *T5-Emotion* variables. To test this possibility, we now enter all three types of emotion variables as predictors in the same regression.

To this end, we design another regression with three steps: 1) Individual differences, 2) All *T5-Emotion* variables, and 3) All *Emoticon* and *LIWC* variables. For simplicity, here we combine the affect variables from the participant and the partner together in the same step.

The results are summarized in Table 9. First, we note that *T5-Emotion* variables themselves show huge improvements above only the individual difference variables, as already established by the analyses presented in the previous sections. Further, with the incorporation of *Emoticon* and *LIWC*, we find that the prediction model shows minimal but significant improvements above just using *T5-Emotion*, for both satisfaction ($F(42, 1969)=7.73$, $p < .001$, $R^2=.142$: $\Delta F(16, 1969)=2.42$, $p < .01$, $\Delta R^2=.017$) and liking ($F(42, 1969)=13.08$, $p < .001$, $R^2=.218$: $\Delta F(16, 1969)=2.87$, $p < .001$, $\Delta R^2=.018$). These results show that *Emoticon* and *LIWC* do provide some unique value to the model, which is not already captured by *T5-Emotion* variables, indicating that automated systems can benefit from leveraging these different methods of emotion recognition collectively, by extracting emotion along the emoticons, lexical, and contextual dimensions. However, given the small effect size, we would still encourage researchers to explore this further in the future, with potentially more sophisticated deep learning architectures and trained models.

5.5 Studying Temporal Effects

Finally, we conduct analysis to study the temporal effects of emotion expression on the prediction of outcomes. Our primary objective here is to guide the design of an automated negotiation agent. We aim to understand whether emotion expression at the start of the conversation is sufficient to build the prediction model or instead, whether a more continual learning approach is preferred. Since the number of turns in each dialogue are relatively small in CaSiNo, it is less meaningful to study the variation in emotion expression at a granular level. Hence, we divide the dialogues into two halves and compute the emotion variables on each half

TABLE 9: Regression results for predicting Satisfaction and Liking with emotion recognition techniques pitted against each other. We combine the affect variables corresponding to the participant and the partner in the same step. ** denotes $p < 0.01$. *** denotes $p < 0.001$.

Variables	R ²	df	F	R ² Change	F Change
Satisfaction					
Individual Differences	.024	(14, 1997)	3.46***	—	—
+T5-Emotion	.125	(26, 1985)	10.88***	.101	19.10***
+Emoticons+LIWC	.142	(42, 1969)	7.73***	.017	2.42**
Liking					
Individual Differences	.041	(14, 1997)	6.05***	—	—
+T5-Emotion	.200	(26, 1985)	19.07***	.159	32.90***
+Emoticons+LIWC	.218	(42, 1969)	13.08***	.018	2.87***

TABLE 10: Temporal analysis for predicting Satisfaction and Liking. We only consider *T5-Emotion* dimensions due to their superiority in other experiments. We combine the affect variables corresponding to the participant and the partner in the same step. *** denotes $p < 0.001$.

Variables	R ²	df	F	R ² Change	F Change
Satisfaction					
Individual Differences	.023	(14, 1995)	3.42***	—	—
+First Half Affect	.087	(26, 1983)	7.30***	.064	11.58***
+Second Half Affect	.131	(38, 1971)	7.83***	.044	8.28***
Liking					
Individual Differences	.040	(14, 1995)	6.00***	—	—
+First Half Affect	.139	(26, 1983)	12.36***	.099	19.03***
+Second Half Affect	.212	(38, 1971)	13.96***	.073	15.13***

separately. Further, we focus only on *T5-Emotion* variables since they do not rely on the usage of emoticons and specific lexical usage that can be limited when considering a portion of the dialogue. To answer our research question, we consider the individual differences, first half affect, and second half affect as three steps in our regression model.

We present the results in Table 10. As evident, for both satisfaction and liking, it is necessary to consider the affect attributes extracted from both the halves of the dialogues to match the overall performance as reported in previous sections. For satisfaction, while just incorporating the first half affect shows significant improvements over individual differences ($F(26, 1983)=7.30$, $p < .001$, $R^2=.087$), incorporating variables from the latter parts of the conversation shows further substantial gains ($F(38, 1971)=7.83$, $p < .001$, $R^2=.131$). We observed similar trends for liking prediction as well. Simply incorporating the affect expression in the first half shows significant benefits but fails to match the variance explained when the entire dialogue was used ($F(26, 1983)=12.36$, $p < .001$, $R^2=.139$), which is only matched after incorporating the affect expression from the second half as well ($F(38, 1971)=13.96$, $p < .001$, $R^2=.212$). These findings point towards the importance of measuring emotion variables continuously for an automated agent. The agent’s design should incorporate the entire information as it is received and update its prediction of the subjective outcomes in a negotiation.

6 DISCUSSION AND CONCLUSIONS

Our aim was to empirically investigate the extent to which the affect variables extracted from the negotiation itself are helpful for the prediction of two important metrics for negotiation agents – participant’s outcome satisfaction and

liking for the partner. We presented an extensive analysis based on a large-scale dataset of human-human negotiation dialogues, grounded in a realistic camping scenario. We devised three degrees of emotion dimensions, from leveraging emoticon usage similar to prior work in menu-driven systems, to going beyond and extracting emotion expression directly from the textual chat utterances. Our results show that such affect variables explain variance in these subjective outcomes, above and beyond the impact of individual difference variables that are available before the negotiation begins. We further find that the contextual emotion dimensions from the deep learning *T5-Emotion* model better account for the variance than emoticon or lexical level measures (Sections 4 and 5.4). We take this opportunity to encourage researchers to further expand the adoption of the recent advancements in deep learning architectures for future work in both the analysis and the design of automated agents.

One might expect that the affect variables would help merely because they are manifestations of the individual differences such as the social value orientation, agreeableness, or gender. If this was indeed the case, there would have been no need to incorporate such affective variables, but merely the individual difference variables would have sufficed. However, to the extent to which the affect variables help in predictions, above and beyond these individual differences, this suggests that there is utility in extracting these emotion dimensions explicitly for improved agent’s performance, being especially well-suited for more subjective outcomes in a negotiation.

We further note that the individual difference variables that we use are based on self-identification and standard personality tests from the psychology literature. This makes these variables more reliable than if they were to be implicitly inferred from past social behaviors of the users, as attempted in a number of previous works analyzing interactions on social media platforms [9], [10], [11]. If affective factors show significant improvements above the individual differences measured with minimal error, this further attests to their utility when the demographics and personality information is inferred, and thus, less reliable.

Since negotiations and other mixed-motive situations can be fraught with emotional decisions, our findings suggest that it would be useful for designers of agents that negotiate with humans to be *armed* with algorithms for understanding the unfolding emotions displayed by users in terms of their emoticon use and natural language. Our work suggests that demographic and personality attributes inferred from past social interactions (such as on social media) are not sufficient and that providing information to negotiating agents from affective channels tracked during the negotiation itself is important for developing agents that can predict or understand if their human counterpart is satisfied and likes them. Ultimately, outcomes like satisfaction and liking will be essential for such agents to cultivate if they, like human negotiators, hope to successfully negotiate with that same partner in future interactions [4].

We note that even after incorporating the emotion attributes along with the individual difference variables, the reported correlation values, although highly significant, are moderate. Explaining variance in such outcomes is challeng-

ing because real-world negotiations are complex and a myriad of factors can contribute. However, accordingly, models could explain more of the variance in negotiation outcomes by adding additional predictor variables. For example, as the results from Section 5.3 suggest, incorporating objective performance into the model explains more variance. Other possible predictors that could explain additional variance include prior expectations of the negotiators [2], selfish linguistic strategies such as undervaluing their partners [14], rapport building [21], favor exchange [22], and expression of empathy [46]. Prajod and colleagues manipulated the warmth expressed by the agent through multiple modalities such as facial expressions and eye gaze, and found that higher warmth is related to increased satisfaction, even with a similar objective performance [47]. Combining these variables along with the affective information considered in this work provides exciting avenues for future research in automated agents.

In order to promote ethical practices in Affective Computing research, we now briefly discuss the ethical considerations around the development of automated negotiation agents. There has been a significant amount of work on ethics in the negotiation literature. The acts of emotion manipulation, deception, bias, and misinterpretation have always been central concerns [48]. It is plausible that these behaviors also get reflected in the negotiation agents that are grounded in the insights gained from human-human negotiation datasets such as CaSiNo. To mitigate the impact of these ethical concerns when automated negotiation systems are deployed to the end-users, we lay out a few recommendations. We believe that transparency is the key. Being upfront and open about the identity, capabilities, and any known misbehaviors of the agent would prevent any misinterpretation and would set the right expectations for the end-users. Further, we note that our work encourages the agent to adapt its behavior by incorporating affective attributes along with user demography and personality. If the deployed agent is indeed adaptive towards these traits, we recommend a regular monitoring cycle for unexpected behaviors of the agent, preventing against any offensive or discriminative actions.

ACKNOWLEDGMENTS

This invited paper is an extension of our previously published work at the International Conference on Affective Computing and Intelligent Interaction (ACII 2021) [49]. We would like to thank the reviewers at ACII 2021 for their feedback that inspired the extensions in this paper. Further, we thank the attendees of ACII for helpful discussions. We are also grateful to the reviewers of IEEE Transactions on Affective Computing for their valuable time in reviewing our submission.

Our research was sponsored by the National Science Foundation and Army Research Office under Cooperative Agreement 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] J. Gratch, D. DeVault, G. M. Lucas, and S. Marsella, "Negotiation as a challenge problem for virtual humans," in *International Conference on Intelligent Virtual Agents*. Springer, 2015, pp. 201–215.
- [2] R. L. Oliver, P. S. Balakrishnan, and B. Barry, "Outcome satisfaction in negotiation: A test of expectancy disconfirmation," *Organizational Behavior and Human Decision Processes*, vol. 60, no. 2, pp. 252–275, 1994.
- [3] J. Mell, J. Gratch, R. Aydoğan, T. Baarslag, and C. M. Jonker, "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)*. IEEE, 2019, pp. 1–7.
- [4] R. Aydoğan, T. Baarslag, K. Fujita, J. Mell, J. Gratch, D. de Jonge, Y. Mohammad, S. Nakadai, S. Morinaga, H. Osawa et al., "Challenges and main results of the automated negotiating agents competition (anac) 2019," in *Multi-Agent Systems and Agreement Technologies*. Springer, 2020, pp. 366–381.
- [5] Y. Leviathan and Y. Matias, "Google duplex: An ai system for accomplishing real-world tasks over the phone," URL <https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html>, vol. 3, 2018.
- [6] E. Johnson, G. Lucas, P. Kim, and J. Gratch, "Intelligent tutoring system for negotiation skills training," in *International Conference on Artificial Intelligence in Education*. Springer, 2019, pp. 122–127.
- [7] A. F. Stuhlmacher and A. E. Walters, "Gender differences in negotiation outcome: A meta-analysis," *Personnel Psychology*, vol. 52, no. 3, pp. 653–677, 1999.
- [8] Y. Xu, P. Sequeira, and S. Marsella, "Towards modeling agent negotiators by analyzing human negotiation behavior," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2017, pp. 58–64.
- [9] Y. Dong, Y. Yang, J. Tang, Y. Yang, and N. V. Chawla, "Inferring user demographics and social strategies in mobile social networks," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 15–24.
- [10] A. Ortigosa, J. I. Quiroga, and R. M. Carro, "Inferring user personality in social networks: A case study in facebook," in *2011 11th International Conference on Intelligent Systems Design and Applications*. IEEE, 2011, pp. 563–568.
- [11] S. Adali and J. Golbeck, "Predicting personality with social behavior: a comparative study," *Social Network Analysis and Mining*, vol. 4, no. 1, pp. 1–20, 2014.
- [12] J. Mell and J. Gratch, "Iago: interactive arbitration guide online," in *AAMAS*, 2016, pp. 1510–1512.
- [13] H. He, D. Chen, A. Balakrishnan, and P. Liang, "Decoupling strategy and generation in negotiation dialogues," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 2333–2343.
- [14] K. Chawla, J. Ramirez, R. Clever, G. Lucas, J. May, and J. Gratch, "Casino: A corpus of campsite negotiation dialogues for automatic negotiation systems," *arXiv preprint arXiv:2103.15721*, 2021.
- [15] J. F. Nash Jr, "The bargaining problem," *Econometrica: Journal of the Econometric Society*, pp. 155–162, 1950.
- [16] P. J. Carnevale and D. G. Pruitt, "Negotiation and mediation," *Annual review of psychology*, vol. 43, no. 1, pp. 531–582, 1992.
- [17] C. Beam and A. Segev, "Automated negotiations: A survey of the state of the art," *Wirtschaftsinformatik*, vol. 39, no. 3, pp. 263–268, 1997.
- [18] T. Baarslag, M. J. Hendrikx, K. V. Hindriks, and C. M. Jonker, "A survey of opponent modeling techniques in automated negotiation," in *15th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2016*. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2016, pp. 575–576.
- [19] G. Stratou, R. Hoegen, G. Lucas, and J. Gratch, "Emotional signaling in a social dilemma: An automatic analysis," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2015, pp. 180–186.

- [20] C. M. de Melo, P. Carnevale, and J. Gratch, "The effect of expression of anger and happiness in computer agents on negotiations with humans," in *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, 2011, pp. 937–944.
- [21] J. Nadler, "Rapport in negotiation and conflict resolution," *Marq. L. Rev.*, vol. 87, p. 875, 2003.
- [22] J. Mell, G. M. Lucas, and J. Gratch, "An effective conversation tactic for creating value over repeated negotiations," in *AAMAS*, vol. 15, 2015, pp. 1567–1576.
- [23] G. Zacharia and P. Maes, "Trust management through reputation mechanisms," *Applied Artificial Intelligence*, vol. 14, no. 9, pp. 881–907, 2000.
- [24] T. Baarslag, K. Fujita, E. H. Gerding, K. Hindriks, T. Ito, N. R. Jennings, C. Jonker, S. Kraus, R. Lin, V. Robu *et al.*, "Evaluating practical negotiating agents: Results and analysis of the 2011 international competition," *Artificial Intelligence*, vol. 198, pp. 73–103, 2013.
- [25] N. Asher, J. Hunter, M. Morey, B. Farah, and S. Afantenos, "Discourse structure and dialogue acts in multiparty dialogue: the stac corpus," in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 2016, pp. 2721–2727.
- [26] M. Lewis, D. Yarats, Y. Dauphin, D. Parikh, and D. Batra, "Deal or no deal? end-to-end learning of negotiation dialogues," in *EMNLP*, 2017.
- [27] M. Cheng, W. Wei, and C.-J. Hsieh, "Evaluating and enhancing the robustness of dialogue systems: A case study on a negotiation agent," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 2019, pp. 3325–3335.
- [28] A. Parvaneh, E. Abbasnejad, Q. Wu, and J. Shi, "Show, price and negotiate: A hierarchical attention recurrent visual negotiator," *arXiv preprint arXiv:1905.03721*, 2019.
- [29] Y. Zhou, Y. Tsvetkov, A. W. Black, and Z. Yu, "Augmenting non-collaborative dialog systems with explicit semantic and strategic dialog history," in *International Conference on Learning Representations*, 2019.
- [30] Z. Zhang, L. Liao, X. Zhu, T.-S. Chua, Z. Liu, Y. Huang, and M. Huang, "Learning goal-oriented dialogue policy with opposite agent awareness," in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 2020, pp. 122–132.
- [31] R. Yang, J. Chen, and K. Narasimhan, "Improving dialog systems for negotiation with personality modeling," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2021, pp. 681–693.
- [32] R. Joshi, V. Balachandran, S. Vashishth, A. Black, and Y. Tsvetkov, "Dialograph: Incorporating interpretable strategy-graph networks into negotiation dialogues," in *International Conference on Learning Representations*, 2020.
- [33] D. DeVault, J. Mell, and J. Gratch, "Toward natural turn-taking in a virtual human negotiation agent," in *AAAI Spring Symposium*. Citeseer, 2015.
- [34] C. Fershtman, "The importance of the agenda in bargaining," *Games and Economic Behavior*, vol. 2, no. 3, pp. 224–238, 1990.
- [35] P. A. Van Lange, E. De Bruin, W. Otten, and J. A. Joireman, "Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence," *Journal of personality and social psychology*, vol. 73, no. 4, p. 733, 1997.
- [36] L. R. Goldberg, "An alternative" description of personality": the big-five factor structure," *Journal of personality and social psychology*, vol. 59, no. 6, p. 1216, 1990.
- [37] S. Bogaert, C. Boone, and C. Declerck, "Social value orientation and cooperation in social dilemmas: A review and conceptual model," *British Journal of Social Psychology*, vol. 47, no. 3, pp. 453–480, 2008.
- [38] R. G. Curtis, T. D. Windsor, and A. Soubet, "The relationship between big-5 personality traits and cognitive ability in older adults—a review," *Aging, Neuropsychology, and Cognition*, vol. 22, no. 1, pp. 42–71, 2015.
- [39] C. G. McClintock, "Social values: Their definition, measurement and development," *Journal of Research & Development in Education*, 1978.
- [40] M. G. Ehrhart, K. H. Ehrhart, S. C. Roesch, B. G. Chung-Herrera, K. Nadler, and K. Bradshaw, "Testing the latent factor structure and construct validity of the ten-item personality inventory," *Personality and Individual Differences*, vol. 47, no. 8, pp. 900–905, 2009.
- [41] J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic inquiry and word count: Liwc 2001," *Mahway: Lawrence Erlbaum Associates*, vol. 71, no. 2001, p. 2001, 2001.
- [42] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *Journal of Machine Learning Research*, vol. 21, pp. 1–67, 2020.
- [43] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in neural information processing systems*, 2017, pp. 5998–6008.
- [44] E. Saravia, H.-C. T. Liu, Y.-H. Huang, J. Wu, and Y.-S. Chen, "Carer: Contextualized affect representations for emotion recognition," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 3687–3697.
- [45] B. L. Monroe, M. P. Colaresi, and K. M. Quinn, "Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict," *Political Analysis*, vol. 16, no. 4, pp. 372–403, 2008.
- [46] O. M. Klimecki, "The role of empathy and compassion in conflict resolution," *Emotion Review*, vol. 11, no. 4, pp. 310–325, 2019.
- [47] P. Prajod, M. Al Owayyed, T. Rietveld, J.-J. van der Steeg, and J. Broekens, "The effect of virtual agent warmth on human-agent negotiation," in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, 2019, pp. 71–76.
- [48] R. J. Lewicki, B. Barry, and D. M. Saunders, *Essentials of negotiation*. McGraw-Hill, 2016.
- [49] K. Chawla, R. Clever, J. Ramirez, G. Lucas, and J. Gratch, "Towards emotion-aware agents for negotiation dialogues," in *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2021, pp. 1–8.

Kushal Chawla received his Bachelors in Computer Science and Engineering from the Indian Institute of Technology Guwahati, India. After graduating in 2017, he spent two years with Big Data Experience Labs at Adobe Research, Bangalore, India, where he focused on the applications of Natural Language Processing (NLP) in marketing, including affect modeling and text summarization. Currently, he is a Ph.D. student at the University of Southern California, Los Angeles, USA. He is broadly interested in the

overlap between Human-Computer Interaction and NLP, focusing on enabling human-machine negotiations through natural language. He is working towards the design of negotiation dialogue systems that exhibit appropriate emotional and strategic behavior for applications in conversational AI and pedagogy.

Rene Clever is a junior Java developer at Skillstorm. She received her Bachelor of Science in Computer Science and minor in Business Management at Lehman College. Before completing her undergraduate degree, she was a summer intern at the University of Southern California's Institute of Creative Technologies, where she contributed to the design of negotiation agents. She is also a part-time CS/Math tutor at Lehman College.

Jaysa Ramirez received her BA degree in Computer Science and Philosophy from Rollins College in Central Florida. Her undergraduate research was focused on Machine Ethics, an interdisciplinary field that seeks to understand whether or not machines can make ethical decisions autonomously – and if so, how they might be able to accomplish such a task. As part of the NSF REU program, she contributed to research done at the University of California's Institute for Creative Technologies concerning the automa-

tion of natural-language negotiations. She is currently a Media Services Technician at LAIKA Studios, the award-winning stop-motion animation studio that combines art and technology to produce films like *Coraline* (2009) and *Kubo and the Two Strings* (2016).

Gale M. Lucas received her Ph.D. degree in psychology from Northwestern University, Evanston, IL, USA, in 2010. She is a Research Assistant Professor with the Viterbi School of Engineering, University of Southern California (USC), Los Angeles, CA, USA, and works with the USC Institute for Creative Technologies, Los Angeles. She works in the areas of human-computer interaction, affective computing, and trust-in-automation. Her research focuses on rapport, disclosure, trust, persuasion, and nego-

tiation with virtual agents and social robots. Dr. Lucas is a member of Association for Computing Machinery.

Jonathan Gratch is a Research Full Professor of Computer Science. Psychology and Media Arts and Practice at the University of Southern California (USC) and Director for Virtual Human Research at USC's Institute for Creative Technologies. He completed his Ph.D. in Computer Science at the University of Illinois in Urbana-Champaign in 1995. Dr. Gratch's research focuses on computational models of human cognitive and social processes, especially emotion, and explores these models' role in advancing

psychological theory and in shaping human-machine interaction. He is the founding Editor-in-Chief (retired) of IEEE's Transactions on Affective Computing, founding Associate Editor of Affective Science, Associate Editor of Emotion Review, and the Journal of Autonomous Agents and Multiagent Systems, and former President of the Association for the Advancement of Affective Computing (AAAC). He is a Fellow of AAAI, AAAC, and the Cognitive Science Society, and an ACM SIGART Autonomous Agents Award recipient.