I Can Already Guess Your Answer: Predicting Respondent Reactions during Dyadic Negotiation

Sunghyun Park, Stefan Scherer, Jonathan Gratch, *Senior Member, IEEE*, Peter J. Carnevale, and Louis-Philippe Morency

Abstract—Negotiation is a component deeply ingrained in our daily lives, and it can be challenging for a person to predict the respondent's reaction (acceptance or rejection) to a negotiation offer. In this work, we focus on finding acoustic and visual behavioral cues that are predictive of the respondent's immediate reactions using a face-to-face negotiation dataset, which consists of 42 dyadic interactions in a simulated negotiation setting. We show our results of exploring four different sources of information, namely nonverbal behavior of the proposer, that of the respondent, mutual behavior between the interactants related to behavioral symmetry and asymmetry, and past negotiation history between the interactants. Firstly, we show that considering other sources of information (other than the nonverbal behavior of the respondent) can also have comparable performance in predicting respondent reactions. Secondly, we show that automatically extracted mutual behavioral cues of symmetry and asymmetry are predictive partially due to their capturing information of the nature of the interaction itself, whether it is cooperative or competitive. Lastly, we identify audio-visual behavioral cues that are most predictive of the respondent's immediate reactions.

Index Terms—Human behavior analysis, negotiation, nonverbal behavior, prediction

1 Introduction

Negotiation is a complex and dynamic process in which two or more parties, often having non-identical preferences or agenda, attempt to reach agreement. Be it in our workplace or with our family or friends, negotiation comprises such a fundamental fabric of our everyday lives that we sometimes engage in the act without even being consciously aware of it. A real-time system that can automatically analyze human behavior in terms of negotiation and predict respondent reactions to negotiation offers has the potential to help people in their daily lives. For instance, such a system could function as a real-time decision support tool, especially in the online environment, to directly help a person during a negotiation process by providing various automatic analyses of the other person's behavior while teleconferencing. Computational analyses and models of behavior during negotiation could also be useful in training a person to be a better negotiator by applying them to create virtual characters for training and simulation.

- *S. Park, S. Scherer, and J. Gratch are with the Institute for Creative Technologies, University of Southern California, 12015 E Waterfront Dr., Los Angeles, CA 90094.*
  E-mail: {park, scherer, gratch}@ict.usc.edu.
- *P.J. Carnevale is with the Marshall School of Business, University of Southern California, Los Angeles, CA 90089-0808.*
  E-mail: peter.carnevale@marshall.usc.edu.
- *L.-P. Morency is with the Language Technology Institute, Carnegie Mellon University, Gates Hillman Center, 5000 Forbes Avenue, Pitts-burgh, PA 15213. E-mail: morency@cs.cmu.edu.*

Manuscript received 1 Dec. 2013; revised 20 Dec. 2014; accepted 9 Jan. 2015. Date of publication 27 Jan. 2015; date of current version 3 June 2015.
Recommended for acceptance by S. D'Mello, M. Pantic, and A. Nijholt.
For information on obtaining reprints of this article, please send e-mail to: reprints.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TAFFC.2015.2396079

Automatically predicting the respondent's reactions to offers during negotiation, that is whether the respondent will accept or reject an offer, is a challenging problem. Despite a long history of research on negotiation [38], much work is still needed in order to fully understand how people display various nonverbal behavioral cues in the context of negotiation. There has been very limited work that investigated nonverbal behavior and computational approaches, but recent progress in computer vision and audio signal processing technologies is enabling automatic extraction of various visual and acoustic behavioral cues without having to depend on costly and time-consuming manual annotations. In this work, we use automatic feature extractions for many low-level behavioral cues such as head displacements and rotations, but we also use manual annotations of several high-level behavioral cues such as head nods that cannot yet be reliably extracted automatically.

In this paper, we present a computational analysis of face-to-face dyadic negotiation sessions to investigate multiple behavioral factors predictive of respondent reactions to negotiation offers (see Fig. 1). For this challenging prediction problem, analyzing nonverbal behavior of the respondent would intuitively be first to consider, but we hypothesize that ample predictive information can reside in other sources as well. Specifically, nonverbal behavior of the proposer might hint at the status of an ongoing negotiation process, and past negotiation history between the two negotiators could shed light on their current relationship, making the respondent more likely to act in a reciprocal manner to a given negotiation offer. Additionally, we explore mutual behavior, which is defined as a set of nonverbal characteristics that occurs due to interactional influence in terms of behavioral symmetry and asymmetry between the two

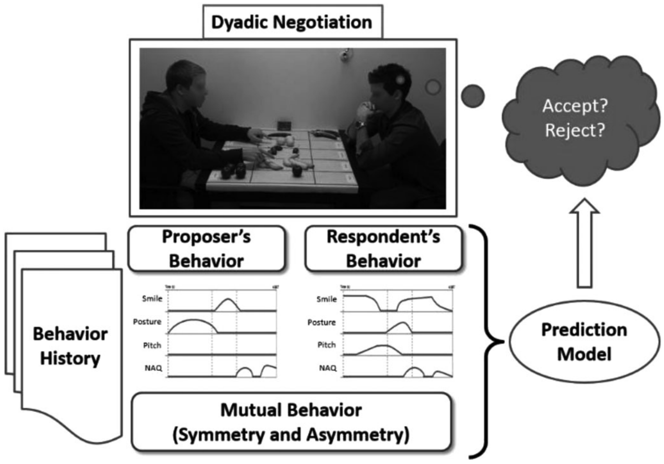

Fig. 1. Overview of our approach to predict respondent reactions (acceptance or rejection) to negotiation offers using predictive computational descriptors from various sources of information.

negotiators. We hypothesize that mutual behavior is important in the context of negotiation because people unconsciously engage in constant adaptation to others' behavior during face-to-face interaction. Then, the degree of behavioral matching or mismatching could show the overall atmosphere or rapport of the participants in the interaction.

With a face-to-face negotiation dataset consisting of 42 dyadic interactions, we present our experimental results to show that such nonverbal cues in various sources of information can be encoded as computational descriptors for a statistical model to automatically predict the respondent's immediate reaction to a negotiation offer. In particular, we examined the following four sources of information: nonverbal behavior of the proposer, that of the respondent, mutual behavior of symmetry and asymmetry between the two negotiators, and the past negotiation history. In addition to demonstrating that the nonverbal behavior of the respondent is not the only source of information useful for making the prediction, we also concentrate on showing mutual behavioral cues that can be extracted automatically to explore the possibility of building an automatic system for the prediction task.

The following section describes previous work on negotiation and computational prediction. In Section 3, we present theoretical background of nonverbal behavior in dyadic interactions and describe how they relate to our prediction problem in negotiation. Section 4 describes detail of our computational descriptors, and Section 5 describes our dataset and experiments. We present and discuss our results in Sections 6 and 7 and conclude in Section 8.

2 Related Work

Negotiation has long been and still is an active topic of research, and the reader can find a brief history of the psychological study of negotiation in [38]. For researchers endorsing a traditional cognitive view, negotiation is essentially a decision-making process, the people involved dispassionate negotiators, and the outcome a result of dynamics governed by rational strategies. There are also researchers who put more emphasis on affective aspects [18]. Some have tried to understand the general role of affect in different stages of negotiation [6] while others have investigated the influence of mood [4], [11], emotion [1], [43], and personality [5]. In addition, researchers focusing on social contexts further deepen our understanding of negotiation dynamics [24], [29].

The affective and social perspectives of negotiation give intuitions that nonverbal behavior can give clues to the ongoing state of a negotiation process. Although negotiation research abounds in literature, there has still been limited work investigating nonverbal behavior in the context of negotiation, let alone computational models. Probably a research problem that is most analogous to our line of research was explored in [15] and [32] in which the authors simulated an employment negotiation and interview scenarios and found that certain nonverbal behavioral features were predictive of the overall outcome in the end. In [15], Curhan and Pentland mainly explored behavioral features in speech and showed that four speech features, including activity, conversational engagement, prosodic emphasis, and vocal mirroring during the first 5 minutes of interaction predicted 30 percent of the variance in individual negotiation outcomes on the terms of employment. One noteworthy finding of this research lay in its research focus on thin slices, the idea that observing only a narrow window of behavior is highly predictive of subsequent evaluations. Additionally, all of the speech features were extracted and encoded automatically. In [32], Nguyen et al. explored features from both speech and visual behavior, not only looking at the behavior of the interviewees but also that of the interviewers, and found that ridge regression explained about 36 percent of the variance in predicting hirability scores. This work used a combination of automatically and manually coded features.

Whereas previous pieces of work were mainly focused on predicting overall negotiation outcomes in the end, our focus is on making immediate predictions of respondent reactions (acceptances or rejections) to individual proposals made during negotiation. In our previous work on this research problem [35], we explored various predictive cues that were manually annotated in three different sources of information, including nonverbal behavior of the proposer, that of the respondent, and negotiation history between the two negotiators. The results served as a proof-of-concept to show that such prediction is possible with a reasonable accuracy, and in a following work [36], we focused on finding multimodal cues from another information source of mutual behavior related to behavioral symmetry and asymmetry, while putting emphasis on automatic feature analysis and extraction. In this paper, we consider all four sources of information together and show a deeper analysis that expands on our previous works.

3 Nonverbal Factors in Face-to-Face Negotiation

In this section, we introduce relevant research on nonverbal behavior in dyadic interaction and highlight four potential sources where predictive cues can be extracted for our research problem: nonverbal behavior of the proposer, that of the respondent, mutual behavior between the two negotiators, and past negotiation history (see Fig. 1).

3.1 Proposer's and Respondent's Behavior

In a business negotiation setting, Niemeier [33] investigated various nonverbal communication channels including proxemics, body postures, gestures, facial expressions, and paralanguage, arguing that they could hint at emotional attitudes of the negotiators. In a study of cooperativeness and competitiveness during negotiation, Johnson [26], [27] similarly found that cooperativeness is expressed through "warm" behavior including a soft tone of voice, smiles, interested facial expressions, direct eye contact, open gestures, close spatial distance, and an occasional soft touch, while competitiveness is expressed through "cold" behavior including tense postures, avoidance of eye contact, closed gestures, distant spatial distance, and avoidance of touching.

Head movements can also provide rich information. For instance, the proposer could show eagerness by nodding his head while staring at the respondent, giving more emotional burden to the respondent if not accepting the offer. Similarly, the respondent could shake his head while listening to the proposal or tilt his head in confusion. Another interesting nonverbal behavior in the context of negotiation is self-touching, which Ekman and Friesen [21] call a type of adaptors. According to Harrigan et al. [25], the overall consensus is that negative affect, such as anxiety or discomfort, triggers self-touching behavior.

3.2 Mutual Behavior

Extensive research shows that we have a tendency to match our behavior to our interactional partners in various ways [30], and it is described with many terms in the literature including behavior matching, imitation, mimicry, synchrony, or chameleon effect. The changes in our behavior often occur unconsciously and in many different channels of communication from facial expressions to speech patterns [13], [14], [30]. Such behavioral characteristics are a part of what we broadly refer to as mutual behavior in this paper, which is not only limited to behavioral symmetry but spans more to also include any nonverbal characteristics that occur due to interactional influence, including behavioral asymmetry.

Mutual behavior is important in the context of negotiation because much evidence exists that it is related to social rapport. In general, people simply seem to get along better when their behavior is well coordinated [9], and it is shown that displaying similar behavior helps with the smoothness of an interaction and also builds a feeling of liking or positivity between interactional partners [13], [14]. The phenomenon is so prevalent that even computer agents that mimic human interactional partners are seen with a more positive feeling than non-mimicking agents [3]. Moreover, studies [8], [23] show that observable nonverbal cues can be indicative of rapport, suggesting that it is possible to detect and gauge rapport between interactional partners, which in turn can be used to assess the status of a negotiation process.

More specifically, Bernieri et al. [8] studied observable nonverbal cues that were indicative of rapport in two different contexts of adversarial and cooperative settings, and the list of behavior include gestures, posture shifts, proximity, back-channel responses, eye contact, and forward lean. Tickle-Degnen and Rosenthal [42], who describes rapport in terms of three components of mutual attentiveness, positivity and coordination, also studied several nonverbal cues associated with rapport that included a similar set of behavior.

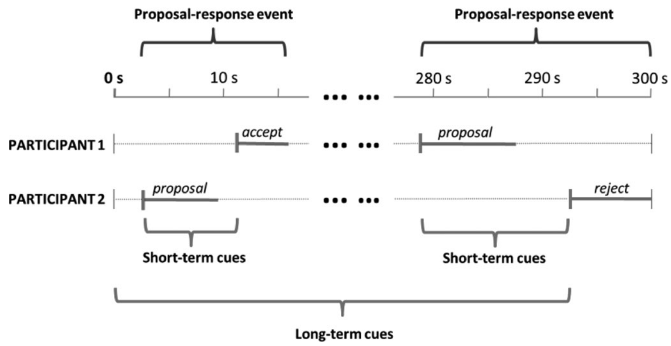

Fig. 2. An illustration of the proposal-response events and different time windows where computational descriptors were extracted.

Mutual behavior that hints at rapport can also reside at the speech level. People are known to imitate various acoustic characteristics of interactional partners including accents, pauses, speech rate, and tone of voice [30]. Some researchers focus more on the smoothness of turn taking, which is usually measured with simultaneous speech, mutual silence, and interruption [9]. Many researchers also investigate synchronization or accommodation in prosody and various vocal qualities to try capturing the interpersonal dynamics in social interaction [17], [39].

3.3 History

The history information can be thought of as capturing the ongoing relationship between the negotiators. For instance, in the absence of other contexts, if the respondent has mostly rejected the proposer's offers in the past, it would mean something quite different from the opposite case. Moreover, reciprocity can be a good predictor of negotiation outcomes in mixed-motive settings [20].

4 Computational Descriptors

In creating computational descriptors for predicting the respondent's reactions in a dyadic session, we identified the following four different sources of information in which predictive cues could reside: nonverbal behavior of the proposer, that of the respondent, mutual behavior between the two negotiators, and their past negotiation history.

Another factor we considered in creating our computation descriptors was time dependency (see Fig. 2). Since negotiation is an ongoing process in which participants constantly adapt themselves to each other, we noted that assessing both short-term and long-term cues could provide a deeper understanding of the current state of negotiation on which to base prediction of future actions. For this purpose, we defined a *proposal-response event* as a time window when the proposer made an utterance with a clear negotiation offer followed by the respondent's clear verbal utterance of acceptance or rejection. In each *proposal-response event*, short-term cues were explored only within the time boundary from start of the proposal until start of the response. For long-term cues, cumulative history of cues were explored from start of the interaction until start of the response. We note here that no information was used from the response part in a *proposal-response event* even when there was an overlap between the proposal and the response.

- *Long-term cues*. These descriptors are designed to model social engagement and rapport that are created over a longer period of time. For example, a continuous mutual gaze is often correlated with high rapport, which in turn can be correlated with successful collaboration.
- *Short-term cues*. These descriptors are designed to model recent momentum in negotiation. For example, the negotiation momentum could change rapidly because of cheating or mockeries, and the short-term descriptors are designed to quickly adapt.

4.1 Proposer's Behavior

For each *proposal-response event*, the following nonverbal behavioral cues displayed by the proposer were explored as potential short-term cues:

- *Head nod*. A vertical downward (or repeated upward and downward) movement of the head.
- *Head shake*. A repeated horizontal left and right movement of the head.
- *Head tilt*. A rotation of the head to the left or to the right (rotation around the z-axis with a frontal view of the face in 3D coordinates).
- *Gaze*. Gaze direction toward the other party, the table, or somewhere else.
- *Smile*. Presence of smiling.
- *Self-touch*. Touching his/her own body with his/her hands (e.g. touching the face with the hand). Only the upper portion of the body was visible in the videos.

The proposer's behavioral cues were manually annotated within the time window of each proposal-response event and were encoded as binary descriptors (except for the gaze, which had three different states) at the event level. For example, the proposer's smile descriptor depended on whether the proposer portrayed a smile or not from start of the proposal until start of the response in each proposal-response event. In summary, from this source of information, a total of six computational descriptors were encoded as short-term cues.

4.2 Respondent's Behavior

In creating computational descriptors of the respondent's behavior in each proposal-response event, we used the same set of behavioral cues and followed the same approach as described for creating the descriptors of the proposer's behavior. In addition to the event-level binary descriptors of the respondent's behavior, we also added another descriptor called binary response time that encoded the respondent's behavior in terms of his/her response time to the proposal.

- *Binary response time*. For each proposal-response event, the response time was computed as the time when the respondent started uttering acceptance or rejection minus the time when the proposer finished uttering his/her proposal. After taking the means of the response times for all accepted and for all rejected cases, the midpoint of the two means was found and used as a threshold, which was 1.37 seconds in our experiments.[1] Using this threshold, the response time in each proposal-response event was converted into a binary descriptor.

In summary, a total of seven computational descriptors were encoded as short-term cues from this information source.

4.3 Mutual Behavior

In creating computational descriptors of mutual behavior, we considered the following three main aspects: behavioral symmetry/asymmetry, automatic extraction, and multi-modality.

Although past research principally focused on symmetric mutual behavior, such as social rapport and behavior matching (see Section 3.2), we note that much information can also reside in asymmetric mutual behavior, such as opposite postures, in the context of our negotiation problem. For behavioral symmetry, we considered the similarity of behavioral patterns of the two negotiators. And for behavioral asymmetry, we considered behavioral patterns of one negotiator that contrasted with those of the other negotiator.

- *Behavioral symmetry*. This behavioral characteristic describes similarity and synchrony in the negotiators' behavior. For example, mutual gaze or reciprocal smile can show a general feeling of rapport and connection. We expect to see these more often in cooperative settings.
- *Behavioral asymmetry*. This behavioral characteristic describes unilateral behavior or behavioral patterns that contrast between the negotiators. For example, if only one of the two negotiators is smiling or if they show opposite body postures, these are possible signs of disengagement and competition.

Such behavioral characteristics of symmetry and asymmetry were captured with the following three computational descriptors that were derived from each type of behavioral cues (see Fig. 3). For instance, a visual signal such as smile or an acoustic signal such as pitch was extracted for both negotiators in each dyad (more detail about specific behavioral cues is in the following sections). Then symmetric and asymmetric characteristics were summarized as follows:

- *Correlation*. Pearson's correlation coefficient was computed for each behavioral cue between the two negotiators in a dyad. The higher the correlation, the more symmetric the behavior in the specific behavioral dimension. The correlation coefficient of –1 would mean perfect asymmetry.
- *Difference in the means*. For each negotiator in a dyad, the mean value was computed for each behavioral cue, and the absolute difference between the two mean values was computed. A higher difference

1. We note that recomputing this threshold yielded a very similar value with the mean of 1.50 seconds and the standard deviation of 0.22 seconds when using only training and validation folds in our 12 experiments from 3 randomly balanced sets x 4-fold cross validation. See Section 5.3 for experiment detail.

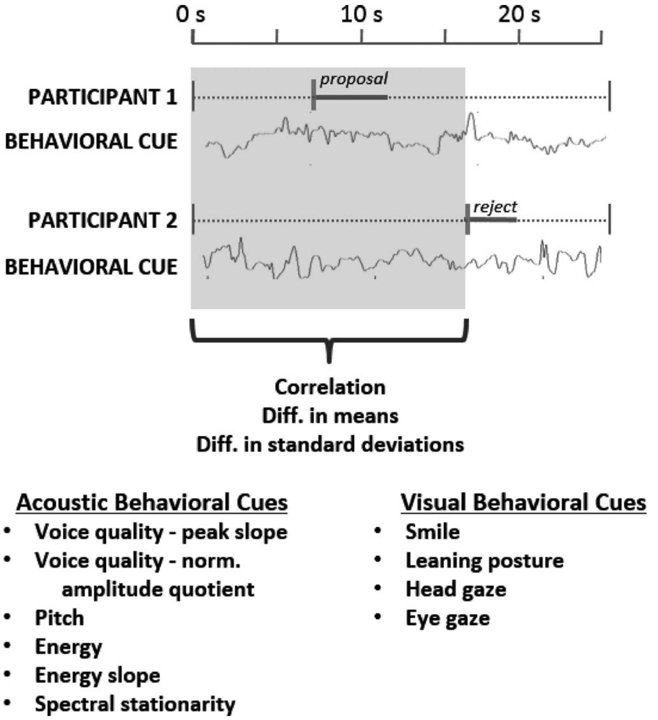

Fig. 3. An illustration of how audio-visual mutual behavior of symmetry and asymmetry is encoded as three computational descriptors for each type acoustic and visual behavioral cues and also as long-term cues. For short-term cues, only behavior from the beginning of the proposal until the beginning of the corresponding response is considered.

value signifies more asymmetry between the two negotiators' behavior.

- *Difference in the standard deviations*. As in computing the difference in the means, the same approach was taken to compute the difference with respect to the standard deviation values.

We concentrated only on nonverbal behavior that could be automatically extracted and that mutually occurred between the proposer and the respondent. That is, in extracting automatic mutual behavioral descriptors, we derived each by considering jointly nonverbal behavior of both the proposer and the respondent together, and none of these descriptors were derived from nonverbal behavior of just one party in the interaction.

Finally, we explored such symmetric and asymmetric mutual behavior descriptors in two different modalities of acoustic and visual channels.

4.3.1 Acoustic Mutual Behavior

Using publicly available software for speech analysis called Covarep [16], the following acoustic descriptors were extracted at 100 Hz for each participant per proposal-response event. The descriptors were extracted only within the long-term time windows since the amount of time was often too short to compute meaningful descriptors within the short-term time windows:

- *Voice quality—peak slope*. Used to indicate breathiness or tenseness of the voice. Values closer to zero are considered as more tense [28], [39].
- *Voice quality—normalized amplitude quotient (NAQ)*. Another feature for the tenseness of the voice [39].
- *Pitch (f0)*. The base frequency of the speech signal. It is the frequency the vocal folds are vibrating during voiced speech segments. We utilized the method introduced in [19] in this work.
- *Energy*. Used to indicate the loudness and intensity of the voice.
- *Energy slope*. Extracted as the absolute value of the first derivative of the energy. High slope values indicate stronger changes in the energy and low values higher monotonicity of the energy.
- *Spectral stationarity*. A measure that captures the fluctuations and changes in the voice signal. High values indicate a stable vocal tract and little change in the speech (e.g. during hesitation or sustained elongated vowels) indicating higher monotonicity [22], [41].

For each proposal-response event, the acoustic descriptors extracted for each participant within the long-term time windows were processed with a linear filter, specifically using a time-aligned moving average (sliding window) technique [44] with a time window of 10 seconds. This step was taken since unlike visual signals, acoustic signals usually do not have overlapping regions to compute meaningful mutual behavior descriptors. Then, symmetric and asymmetric mutual behavior descriptors were computed for each acoustic behavioral cue by taking the correlation and the difference in the means and in the standard deviation values. In summary, a total of 18 computational descriptors were encoded as long-term mutual behavior descriptors from this source of information (three types of mutual behavior descriptors of correlation and difference in the means and in the standard deviations multiplied by six types of acoustic behavioral cues).

4.3.2 Visual Mutual Behavior

In order to automatically extract visual mutual behavior descriptors, we used commercial software [34] that detects a person's face from frame to frame in a video and outputs various low-level and high-level facial features. Below is a list of visual descriptors that were extracted as potential predictive cues for each participant per negotiation session. Each visual descriptor listed below was smoothed with a linear filter, and each descriptor, except for smile, was converted into a binary descriptor at each frame using an empirically determined threshold point.

- *Smile*. Used to indicate if the person is displaying positive affect with a smile. The smile intensity value ranges on a scale from 0, which means no smile, up to 100.
- *Leaning posture*. Used to indicate if the person is showing a forward or a backward lean (posture), approximated with face length and face size. The face length and face size values were z-normalized and the threshold points of 0, 0.25, 0.5, 0.75 and 1.0 were used to convert them to binary values at the frame level. With the five different thresholded versions of the descriptor, prediction performance was measured with each of them used in a single-feature predictor. The threshold that performed best was with the threshold point of 0.75, and it was used for all subsequent experiments.

- *Head gaze*. Used to indicate if the face is directed downward (toward the table). From the raw face direction signals in upward / downward rotational degrees, the threshold points of –5, –10, –15 and –20 were used to convert them to binary values at the frame level since the videos were recorded from a lower position at an angle. Based on the prediction performance as an individual descriptor (similarly as how the threshold point was determined in the leaning posture), the threshold point of –5 was eventually used for all the experiments.
- *Eye gaze*. Used to indicate if the gaze is directed downward (toward the table). The same approach was taken as head gaze for converting to binary values at the frame level and the same threshold point of –5 was eventually used.

For each proposal-response event, the visual descriptors above were extracted from two different time windows: within the short-term time window (from the start of the proposal until the start of the response) and within the long-term time window (from the start of the interaction until the start of the response) as shown in Fig. 2. Then, for each time window, symmetric and asymmetric mutual behavior descriptors were computed with Pearson's correlation coefficient for each descriptor between the two participants in each dyadic session. The difference in the mean values and the difference in the standard deviation values were also computed. In summary, a total of 24 computational descriptors were encoded as short-term and long-term mutual behavior descriptors from this source of information (four types of visual descriptors multiplied by three types of mutual behavior descriptors of correlation and differences multiplied by two types of short-term and long-term windows).

We note that unlike audio signals, visual signals tend to happen more simultaneously. For instance, when two people smile or have eye contact, even if the behavior is not perfectly synchronized, there tends to be an overlapping period when both interactants display the behavior at the same time, which we tried to capture with the correlation coefficients. We further note that it would be interesting to take into account time delays using a similar time-aligned moving average technique that we used for acoustic descriptors, which we leave for future work. Additionally, time delay in behavior is not relevant for the descriptors of the difference in the means and in the standard deviations because they are already summary statistics over a time period.

4.4 Negotiation History

To capture useful predictive cues from negotiation history, we explored the following descriptors from the long-term time windows:

- *Net negotiation history*. The total net response history of the respondent at the time of the proposal-response event (+1 and –1 for each previous acceptance and rejection respectively).
- *Last negotiation history*. The result of the proposal-response event (+1 for acceptance and –1 for rejection) immediately prior to the current one.
- *Response time history*. The mean of all the previous response times of the respondent at the time of the proposal-response event. This descriptor could help better understand the binary response time descriptor by providing the general response time characteristic / habit of each negotiator.

In summary, a total of three computational descriptors were encoded from this source of information.

5 Experiments

We designed and performed our experiments to address the following primary hypothesis to investigate the degree of benefit that can be gained with more sources of information where we seek potential predictive cues of respondent reactions:

Hypothesis 1 (H1). *For predicting respondent reactions during dyadic negotiation, other sources of information (proposer's nonverbal behavior, mutual behavior, and negotiation history) can yield comparable prediction performance to looking at nonverbal behavior of the respondent, and combining all sources together yields higher performance than using a single source of information.*

In addition to the primary research question, we also tested a secondary hypothesis regarding mutual behavior. If H1 is true and mutual behavior descriptors are predictive of respondent reactions during negotiation, we suspected that it may be due to the descriptors capturing the very nature of the interaction itself, whether it is cooperative or competitive.

Hypothesis 2 (H2). *Computational descriptors of mutual behavior that are predictive of respondent reactions are also useful for determining whether the negotiation interaction is cooperative or competitive.*

5.1 Dyadic Negotiation Dataset

A dataset of dyadic negotiation sessions was collected in order to understand how people negotiate with various incentive scenarios. In total, 84 undergraduate business major students (40 males and 44 females) participated in 42 dyadic negotiation sessions, of which one dyad was discarded because the participants deviated from the experimental procedure. Each dyadic session involved same-sex participants to control for the influence of gender. In addition, negotiators in each dyad were instructed to adopt only one of three motivational orientations that derived from the monetary incentive associated with the negotiation task: cooperative (maximize joint outcomes), individualistic (maximize own outcomes), and competitive (maximize own outcomes relative to the other's outcomes). Out of 42 sessions, 13 were cooperative, 13 individualistic, and 16 competitive. Negotiators in each dyad received the same motivational instruction and were aware that the other was so instructed. A total of three cameras were placed unobtrusively to record a near-frontal view of each negotiator, as well as an overall side view of the interaction.

In each session, two participants sat face-to-face across each other at opposite ends of a table, on which several types of plastic fruits or vegetables were placed. The participants were randomly assigned to represent one of two

TABLE 1
Proposal-Response Events Distribution in Our Face-to-Face Dyadic Negotiation Dataset

Total # of sessions		41 dyads
Total # of samples (proposal-response events)		253
Accept samples across sessions	- mean	4.63
	- standard dev.	2.38
Reject samples across sessions	- mean	1.54
	- standard dev.	2.37
Total samples across sessions	- mean	6.17
	- standard dev.	2.97

different restaurants, which had different pay-off matrices associated with the items on the table. Each participant knew only the pay-off matrix of his/her assigned restaurant, and the participants had 12 minutes to negotiate on how to distribute the items on the table. As an incentive, each participant could receive up to $50 depending on the final points earned for his/her restaurant (see Fig. 1).

5.2 Annotations

For each negotiation session, all the events of proposal-response pairs were identified by two coders with each coder annotating half of the dataset, and the inter-coder reliability on four randomly selected sessions (about 10 percent of the dataset) measured with Krippendorff's alpha was at 0.67 (following the approach of measuring Time-Slice Krippendorff's alpha described in [37] with seconds as the time-slice granularity). A proposal is defined as an utterance made with a clear offer related to negotiating the items on the table, and if it is followed by a clear verbal utterance of acceptance or rejection, we grouped the start of the proposal until the end of the matching response as a proposal-response event. A total of 253 proposal-response events were identified, out of which 190 were accepted proposals and 63 were rejected proposals (see Table 1). For each proposal-response event, a subset of nonverbal behavior (see Section 4.1) of the proposer and the respondent were annotated. For the purpose of extracting acoustic descriptors, speaker diarization was also performed with annotations, but we note that this step could have been done automatically with close-talk microphones equipped for both participants. All annotations were performed using ELAN software [10].

5.3 Prediction Model and Methodology

For the prediction models, support vector machine (SVM) classifiers with a radial basis kernel were trained and tested [12]. In all of our prediction experiments, four-fold cross-validation was performed with hold-out testing and also hold-out validation to find the optimal parameters (*gamma* and *C*) using a grid-search technique. An exhaustive feature selection looking at all possible combinations of features (computational descriptors) was performed in each of the four sources of information. For making predictions with combined sources of information, the same feature selection approach was performed after combining the features at the feature-level (early-fusion) using the best subset of features that was automatically determined in each source of information.

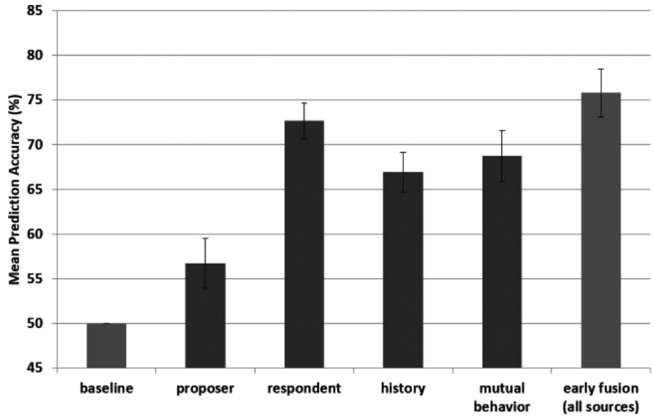

Fig. 4. The mean accuracies for predicting respondent reactions to negotiation offers using computational descriptors from each source of information. The right-most red bar shows the early fusion (feature-level) performance of combining all sources together.

In order to make balanced sample sets for predictor training and testing, all of the 63 samples of the rejected proposal-response events were combined with 63 randomly selected samples of the accepted events (making the baseline prediction at 50 percent), and three such randomly balanced sets were created. Each randomly balanced set was again randomly separated into four folds with almost an equal number of acceptance and rejection samples. All the prediction results were averaged over 12 test results (3 randomly balanced sets $\times$ 4-fold cross-validation). It should be noted that none of the folds contained samples from the same negotiation session for better generalizability. In other words, the four folds were created such that they were all session-independent to one another.

To test for our second hypothesis, we also investigated to what extent the prediction accuracies were due to the mutual behavior descriptors' capturing the different conditions of the negotiation sessions, specifically between the cooperative and competitive conditions. Using the same final descriptor set determined for the mutual behavior group predictor, another classifier was trained and tested in order to classify each negotiation session between the cooperative and competitive conditions. The samples were also randomly balanced with 13 cooperative sessions and 13 competitive sessions (making the baseline classification at 50 percent), and similar feature selection technique and 13-fold cross validation was performed. In each cross validation experiment, one hold-out fold was used for testing, eight folds for training, and four folds for validating the model parameters.

6 Results

All of the experimental results have the baseline prediction rate of 50 percent since all the samples were trained and tested using randomly balanced sets.

6.1 Predicting Respondent Reactions (H1)

For predicting respondent reactions to negotiation offers, combining the computational descriptors from all 4 sources of information at the feature level (early fusion) yielded the mean prediction accuracy of 75.8 percent, outperforming the prediction performance of using any single source of information alone (see Fig. 4). The prediction accuracy was

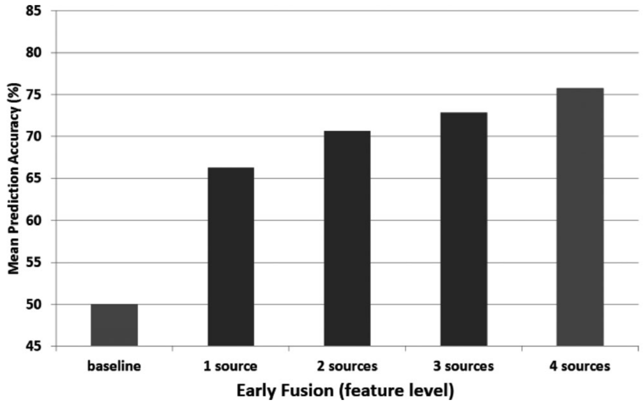

Fig. 5. The mean prediction accuracies for combining multiple sources of information together. The graph shows the results of early fusion at the feature level.

at 56.8 percent when using descriptors from only the proposer's behavior, 72.7 percent when using those from only the respondent's behavior, 66.9 percent when using those from only the history information, and 68.8 percent when using those from only the mutual behavior. Compared with the prediction accuracies of the early-fusion predictor ($M = 75.8$, $SD = 9.7$), the predictor using the proposer's behavior only ($M = 56.8$, $SD = 10.1$) showed the Cohen's d effect size value of 1.92 ($d = 1.92$) suggesting a large effect, the predictor using the respondent's behavior only ($M = 72.7$, $SD = 7.2$) showed $d = 0.36$ suggesting a small effect, the predictor using the history information only ($M = 66.9$, $SD = 8.1$) showed $d = 1.00$ suggesting a large effect, and the predictor using the mutual behavior only ($M = 68.8$, $SD = 10.4$) showed $d = 0.70$ suggesting a medium effect.

6.2 Benefit of More Sources of Information

As shown in Fig. 5, for the early fusion results of combining the computational descriptors at the feature level, the predictors on average performed at 66.3 percent when using only one source of information (mean of four different predictors, one predictor for each source of information), 70.7 percent when using two sources (mean of six different predictors possible by choosing two out of four sources), 72.9 percent when using three sources (mean of four different predictors possible by choosing three out of four sources), and 75.8 percent when using all four sources together. Moreover, the prediction performance of using all possible combinations of sources is summarized in Table 2.

6.3 Mutual Behavior and Classification of Cooperative versus Competitive Interactions (H2)

Using the same final descriptors selected for the information source of mutual behavior, four descriptors were relevant for the interaction condition classification because they were from the long-term time windows, making them sensible to be computed for the entire length of each negotiation session. The descriptors were correlation of smile, correlation of head gaze, difference in means for eye gaze, and difference in means for pitch.

TABLE 2
Prediction Performance of Early-Fusion with Different Combinations of Information Sources

Early fusion sources (● signifies inclusion)				Accuracy (%)
Proposer	Respondent	History	Mutual Behavior	
●	●	●	●	75.8
●	●	●		71.1
●	●		●	74.7
●		●	●	69.5
	●	●	●	76.0
●	●			69.8
●		●		66.9
●			●	65.1
	●	●		73.7
	●		●	74.5
		●	●	74.0
●				56.8
	●			72.7
		●		66.9
			●	68.8

Using the four descriptors, the best classification rate that could be achieved for classifying the negotiation sessions between cooperative or competitive condition was at 65.4 percent.

6.4 Top Performing Individual Descriptors

Table 3 summarizes the performance of top computational descriptors from all sources of information. The prediction accuracies in the table show the performance when a single-descriptor predictor was trained and tested using each individual descriptor alone. In other words, the prediction accuracies mean how much discriminative power each computational descriptor showed when it was considered alone in the prediction problem. From the proposer's nonverbal behavior, head tilt information was the only descriptor that performed above 55 percent at 56.8 percent. From the respondent's nonverbal behavior, three descriptors of binary response time, head nod, and head shake performed above 55 percent when used individually with the prediction accuracies of 63.8, 60.4, and 59.6 percent respectively. From the mutual behavior descriptors, the difference in the means for head gaze and eye gaze performed at 59.6 and 57.3 percent respectively, and the correlation in head gaze and smile both performed at 59.1 percent. Lastly, from the history information, last negotiation history and net negotiation history performed at 66.9 and 58.6 percent respectively.

7 Discussion

Overall, our results showed partial support for our two hypotheses outlined in Section 5. In this section, we elaborate our results with discussions in light of our hypotheses and in the same subtopics as in the results section.

7.1 Limitations

We first note that the samples in our experiments were randomly forced-balanced to have a majority baseline or chance-level prediction at 50 percent. Also, the threshold points for visual mutual descriptors (Section 4.3: Mutual Behavior) were determined using all the training and testing samples. We note that our main focus of this study was

TABLE 3
Top Descriptors According to Their Prediction Performance When Used Alone in a Single-Descriptor Predictor

Source	Descriptor	Accuracy (%)	Time dependency
Proposer	Head tilt	56.8	Short-term
Respondent	Binary response time	63.8	
	Head nod	60.4	
	Head shake	59.6	
Mutual Behavior (Symmetry / Asymmetry)	Leaning posture (correlation)	54.2	
	Head gaze (diff. in means)	59.6	Long-term
	Head gaze (correlation)	59.1	
	Smile (correlation)	59.1	
	Eye gaze (diff. in means)	57.3	
	Eye gaze (correlation)	54.9	
	NAQ (diff. in standard deviations)	54.2	
	Pitch (diff. in means)	54.2	
History	Last negotiation history	66.9	
	Net negotiation history	58.6	

to investigate whether different sources of information would also yield comparable prediction performances to that of looking at nonverbal behavior of the respondent and not on making best possible predictions, and the same thresholding advantage was given to all multimodal predictors that included mutual behavior.

7.2 Predicting Respondent Reactions (H1)

As shown in Fig. 4, other sources of information other than the respondent's nonverbal behavior also displayed comparable predictive power, especially history and mutual behavior sources showing slightly lower but still comparable performance, which confirmed the first half of our first hypothesis. However, when comparing the prediction accuracy results between the four-source predictor versus each of the single-source predictors, not all of them showed a large effect, not completely confirming our first hypothesis. It is not surprising that the nonverbal behavior of the respondent was the most predictive source of information when predicting the respondent reactions (see Fig. 4). After all, if one wishes to predict future actions of a person, it is only natural and intuitive to observe that specific person's behavior more than anything else. Mutual behavior between the negotiators also proved useful, probably due to the descriptors having captured the level of rapport between the negotiators and the overall atmosphere of cooperation or competition. Our previous work [36] specifically on mutual behavior showed similar results. The computational descriptors from the history information were also quite useful in the prediction, most likely by directly capturing whether the interaction was a cooperative or competitive condition. The nonverbal behavior of the proposer, although the least predictive of the four sources, still showed some predictive power.

7.3 Benefit of More Sources of Information

For the results of early fusion at the feature level, we could observe a slight trend that adding more sources of information leads to better prediction rates in general. In the graphs shown in Fig. 5, each bar shows the average prediction accuracy by using information from a certain number of sources in all possible combinations. On average, using information from all four sources of information yielded the best performance for the early fusion (75.8 percent) approach. These results are in line with our previous study [35] in which we considered a total of three sources of information without the mutual behavior and found that considering more sources of information generally helps with the prediction task. However, we note that this tendency is only in terms of general performance, and when we look at the performance of individual predictors with different combinations of sources, the tendency is not conclusive and not always true. For instance, the three-source predictor (using respondent's nonverbal behavior, history, and mutual behavior) actually outperformed the four-source predictor in our specific experiments.

7.4 Mutual Behavior and Classification of Cooperative versus Competitive Interactions (H2)

The condition classification performance of 65.4 percent is relatively lower compared to our early fusion prediction accuracy of 75.8 percent but still by far higher than the chance level at 50 percent. Considering that the engineering and selection of the computational descriptors were not completely focused on the purpose of the interaction condition classification, the classification performance suggests more meaning and implications. This result indicates that the descriptors that were useful for the reaction prediction were also helpful in determining the type of negotiation sessions, moderately confirming our second hypothesis (H2). The result also suggests that the performance of our respondent reaction predictors could be partially due to the mutual behavior descriptors' having captured the nature of the negotiation sessions (the overall atmosphere of cooperation or competitiveness).

7.5 Top Performing Individual Descriptors

As shown in Table 3, from the proposer's behavior, head tilt was slightly predictive of the respondent reactions, possibly because the behavior often showed lack of confidence from the proposer with the proposal, which was more likely to be rejected than accepted. From the respondent's nonverbal behavior, binary response time, head nod and head shake individually yielded the best prediction accuracy. We note that head nods and head shakes from the respondent's behavior individually only performed at about 60 percent prediction accuracy and were not determinant factors. Often in a dyadic session, the respondent gave head nods as a form of backchannel response to the proposer's speech, whose presence were somewhat related to the final respondent reaction to the offer but not to a great extent. From the history information, last negotiation history descriptor performed best at 66.9 percent, which was the best performance among all descriptors and all sources of information. This is most likely due to the descriptor having captured the degree of cooperation or competition in the interaction, as well as the tendency of acceptances or rejections to occur in closer temporal proximity (acceptances tend to happen in blocks and so are rejections). Also, there were several mutual behavior descriptors that performed very close to about 60 percent prediction accuracy even when used alone in a single-feature predictor.

8 Discussion

In this paper, we presented our experimental results showing that we could predict respondent reactions to negotiation offers (whether the respondent will accept or reject) with reasonable accuracy using computational descriptors from four different sources of information: nonverbal behavior of the proposer, nonverbal behavior of the respondent, mutual behavior between the negotiators, and the negotiation history. Furthermore, we presented the results with an early-fusion approach of fusing information at the feature level and showed our qualitative observation that adding more sources of information generally improves the prediction performance. Specifically, we moderately confirmed our first hypothesis that other sources of information other than the nonverbal behavior of the respondent can also be useful in our prediction problem. Furthermore, we moderately confirmed our second hypothesis that the computational descriptors from the mutual behavior are useful in the prediction problem due to their capturing the very nature of the interaction itself between the cooperative and competitive atmosphere. For future work, more behavioral cues and feature engineering can be explored, especially by taking into account time delays in behavior between the negotiators. We also plan to investigate how humans perform for the same prediction problem for comparison.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. IIS-1118018 and the U.S. Army Research, Development, and Engineering Command. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the Government. The authors would also like to thank the USC Annenberg Graduate Fellowship Program for supporting the first author's graduate studies. Sunghyun Park is the corresponding author.

References

[1] K. Allred, J. Mallozzi, F. Matsui, and C. Raia, "The influence of anger and compassion on negotiation performance," *Organ. Behav. Hum. Dec.*, vol. 70, no. 3, pp. 175–187, Jun. 1997.

[2] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis," *Psychol. Bull.*, vol. 111, no. 2, pp. 256–274, 1992.

[3] J. Bailenson and N. Yee, "Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments," *Psychol. Sci.*, vol. 16, no. 10, pp. 814–819, Oct. 2005.

[4] R. Baron, "Environmentally induced positive affect: Its impact on self-efficacy, task performance, negotiation, and conflict," *J. Appl. Soc. Psychol.*, vol. 20, no. 5, pp. 368–384, Mar. 1990.

[5] B. Barry and R. Friedman, "Bargainer characteristics in distributive and integrative negotiation," *J. Pers. Soc. Psychol.*, vol. 74, no. 2, pp. 345–359, 1998.

[6] B. Barry and R. Oliver, "Affect in dyadic negotiation: A model and propositions," *Organ. Behav. Hum. Dec.*, vol. 67, no. 2, pp. 127–143, Aug. 1996.

[7] M. Bazerman, J. Curhan, D. Moore, and K. Valley, "*Negotiation*," *Annu. Rev. Psychol.*, vol. 51, no. 1, pp. 279–314, 2000.

[8] F. Bernieri, J. Gillis, J. Davis, and J. Grahe, "Dyad rapport and the accuracy of its judgment across situations: A lens model analysis," *J. Pers. Soc. Psychol.*, vol. 71, no. 1, pp. 110–129, Jul. 1996.

[9] F. Bernieri and R. Rosenthal, "Interpersonal coordination: Behavior matching and interactional synchrony," in *Fundamentals of Nonverbal Behavior*, R. Feldman and B. Rime, Eds., New York, NY, USA: Cambridge Univ. Press, pp. 401–432, 1991.

[10] H. Brugman and A. Russel, "Annotating multimedia/multimodal resources with ELAN," in *Proc. 4th Int. Conf. Lang uage Res. Eval.*, 2004, pp. 2065–2068.

[11] P. Carnevale and A. Isen, "The influence of positive affect and visual access on the discovery of integrative solutions in bilateral negotiation," *Organ. Behav. Hum. Dec.*, vol. 37, no. 1, pp. 1–13, Feb. 1986.

[12] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, Apr. 2011.

[13] T. Chartrand and J. Bargh, "The chameleon effect: The perception behavior link and social interaction," *J. Pers. Soc. Psychol.*, vol. 76, no. 6, pp. 893–910, 1999.

[14] T. Chartrand, W. Maddux, and J. Lakin, "Beyond the perception behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry," in *The New Unconscious*, R. Hassin, J. Uleman, and J. Bargh, Eds., New York, NY, USA: Oxford Univ. Press, 2005, pp. 334–361.

[15] J. Curhan and A. Pentland, "Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes," *J. Appl. Soc. Psychol.*, vol. 92, no. 3, pp. 802–811, 2007.

[16] G. Degottex, J. Kane, T. Drugman, T. Raitio, and S. Scherer, "COVAREP—A collaborative voice analysis repository for speech technologies," in *Proc. IEEE 39th Int. Conf. Acoust., Speech, Signal Process.*, 2014, pp. 960–964.

[17] C. De Looze, S. Scherer, B. Vaughan, and N. Campbell, "Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction," *Speech Commun.*, vol. 58, pp. 11–14, 2014.

[18] D. Druckman and M. Olekalns, "Emotions in negotiation," *Group Decision Negotiation*, vol. 17, no. 1, pp. 1–11, Jan. 2008.

[19] T. Drugman and A. Abeer, "Joint robust voicing detection and pitch estimation based on residual harmonics," in *Proc. 12th Annual Conf. Int. Speech Commun. Assoc.*, 2011 pp. 28–31.

[20] M. Dufwenberg and G. Kirchsteiger, "A theory of sequential reciprocity," *Games Eco. Beh.*, vol. 47, no. 2, pp. 268–298, 2004.

[21] P. Ekman and W. Friesen, "The repertoire of nonverbal behavior," *Semiotica*, vol. 1, pp. 49–98, 1969.

[22] S. Finkelstein, S. Scherer, A. Ogan, L.-P. Morency, and J. Cassell, "Investigating the influence of virtual peers as dialect models on students' prosodic inventory," in *Proc. Workshop Child, Comput. Interaction*, 2012.

[23] J. Grahe and F. Bernieri, "The importance of nonverbal cues in judging rapport," *J. Nonverbal Behav.*, vol. 23, no. 4, pp. 253–269, Dec. 1999.

[24] L. Greenhalgh and D. Chapman, "Negotiator relationships: Construct measurement and demonstration of their impact on the process and outcomes of negotiation," *Group Decis. Negot.*, vol. 7, no. 6, pp. 465–489, Nov. 1998.

[25] J. Harrigan, J. Kues, and J. Weber, "Impressions of hand movements: Self-touching and gestures," *Perceptual Motor Skills*, vol. 63, no. 2, pp. 503–516, 1986.

[26] D. Johnson, "Effects of warmth of interaction, accuracy of understanding, and the proposal of compromises on listener's behavior," *J. Couns. Psycol.*, vol. 18, no. 3, pp. 207–216, 1971.

[27] D. Johnson, K. McCarty, and T. Allen, "Congruent and contradictory verbal and nonverbal communications of cooperativeness and competitiveness in negotiations," *Commun. Res.*, vol. 3, pp. 275–292, 1976.

[28] J. Kane and C. Gobl, "Identifying regions of non-modal phonation using features of the wavelet transform," in *Proc. 12th Annu. Conf. Int. Speech Commun. Assoc.*, 2011 pp. 177–180.

[29] R. Kramer and D. Messick, *Negotiation as a Social Process*. Thousand Oaks, CA, USA: Sage, 1995.

[30] M. Louwerse, R. Dale, E. Bard, and P. Jeuniaux, "Behavior matching in multimodal communication is synchronized," *Cognitive Sci.*, vol. 36, no. 8, pp. 1404–1426, Nov./Dec. 2012.

[31] A. Mehrabian, *Nonverbal Communication*. Chicago, IL, USA: Aldine 2007.

[32] L. Nguyen, D. Frauendorfer, M. Mast, and D. Gatica-Perez, "Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior," *IEEE Trans. Multimedia*, vol. 16, no. 4, pp. 1018–1031, Jun. 2014.

[33] S. Niemeier, "Nonverbal expressions of emotions in a business negotiation," in *The Language of Emotions*, S. Niemeier and R. Dirven, Eds. Amsterdam, The Netherlands: John Benjamins Publishing Company, 1997, pp. 277–306.

[34] OKAO Vision [Online]. Available:http://www.omron.com/r_d/coretech/vision/okao.html

[35] S. Park, J. Gratch, and L.-P. Morency, "I already know your answer: Using nonverbal behaviors to predict immediate outcomes in a dyadic negotiation," in *Proc. 14th ACM Int. Conf. Multimodal Interaction* , Oct. 2012, pp. 19–22.

[36] S. Park, S. Scherer, J. Gratch, P. Carnevale, and L.-P. Morency, "Mutual behaviors during dyadic negotiation: Automatic prediction of respondent reactions," in *Proc. Int. Conf. Affective Comput. Intell. Interaction*, 2013, pp. 423–428.

[37] S. Park, P. Shoemark, and L.-P. Morency, "Toward crowdsourcing micro-level behavior annotations: The challenges of interface, training, and generalization," in *Proc. 19th Int. Conf. Intell. User Interfaces*, Feb. 2014, pp. 37–46.

[38] D. Pruitt, "A history of social conflict and negotiation research," in *Handbook of the History of Social Psychology*, A. Kruglanski and W. Stroebe, Eds. New York, NY, USA: Psychology Press, 2012, pp. 431–452.

[39] S. Scherer, Z. Hammal, Y. Yang, and J. Cohn, "Dyadic behavior analysis in depression severity assessment interviews," in *Proc. 16th Int. Conf. Multimodal Interaction*, 2014, pp. 112–119.

[40] S. Scherer, J. Kane, C. Gobl, and F. Schwenker, "Investigating fuzzy-input fuzzy-output support vector machines for robust voice quality classification," *Comput. Speech Language*, vol. 27, no. 1, pp. 263–287, Jan. 2013.

[41] S. Scherer, N. Weibel, L.-P. Morency, and S. Oviatt, "Multimodal prediction of expertise and leadership in learning groups," in *Proc. 1st Int. Workshop Multimodal Learn. Anal.*, 2012.

[42] L. Tickle-Degnen and R. Rosenthal, "The nature of rapport and its nonverbal correlates," *Psychol. Inq.*, vol. 1, no. 4, pp. 285–293, 1990.

[43] G. Van Kleef, C. De Dreu, and A. Manstead, "The interpersonal effects of emotions in negotiations," *J. Pers. Soc. Psychol.*, vol. 87, no. 4, pp. 510–528, 2004.

[44] B. Vaughan, "Prosodic synchrony in cooperative task-based dialogues: A measure of agreement and disagreement," in *Proc. 12th Annu. Conf. Int. Speech Commun. Assoc.*, 2011, pp. 1865–1868.

Sunghyun Park received the bachelor's and master's degrees in computer science from the Georgia Institute of Technology. He is currently working toward the PhD degree supported by the Annenberg Graduate Fellowship in the Computer Science Department at the University of Southern California. His research interests include computational analysis of nonverbal social communication, affective computing, and machine learning.

Stefan Scherer received the degree of Drrernat from the faculty of Engineering and Computer Science at Ulm University in Germany with the grade summa cum laude (i.e., with distinction). He is a research assistant professor in the Department of Computer Science at the University of Southern California (USC) and a research associate at the USC Institute for Creative Technologies where he leads research projects funded by the US National Science Foundation (NSF) and ARL. His research aims to automatically identify, characterize, model, and synthesize individuals' multimodal nonverbal behavior within both human-machine as well as machine-mediated human-human interaction.

Jonathan Gratch received the PhD degree in Computer Science at the University of Illinois in Urban-Champaign in 1995. He is the director for Virtual Human Research at the University of Southern California's (USC) Institute for Creative Technologies, a research full professor of computer science and psychology at USC and the director in USC's Computational Emotion Group. His research focuses on computational models of human cognitive and social processes, especially emotion, and explores these models' role in shaping human-computer interactions in virtual environments. He studies the relationship between cognition and emotion, the cognitive processes underlying emotional responses, and the influence of emotion on decision making and physical behavior. He is the founding current editor-in-chief of the *IEEE's Transactions on Affective Computing* (3.5 impact factor in 2013), an associate editor of *Emotion Review* and the *Journal of Autonomous Agents and Multiagent Systems*, and the former president of the Association for the Advancement of Affective Computing (AAAC). He is the author of more than 200 technical articles. He is a AAAI fellow, received the SIGART Autonomous Agent's Award, a senior member of the IEEE, and a member of the International Society for Research on Emotion (ISRE).

Peter J. Carnevale is a professor of business, psychology and communication at the University of Southern California (USC) Marshall School of Business. He is an expert on negotiation, mediation, group problem solving, and creativity. His work is published in leading psychology and management journals, and he has served on numerous journal editorial boards. His book with Carsten de Dreu, *Methods of Negotiation Research* received an outstanding book award from the International Association for Conflict Management. He was elected the chair of the Conflict Management Division of the Academy of Management as well as President of IACM. His current research investigates the role of emotion expression by artificial humans (e.g., "avatars") in negotiation, and is currently funded by the National Science Foundation program on Social-Computational Systems (SoCS).

Louis-Philippe Morency received the master's and PhD degrees from MIT Computer Science and Artificial Intelligence Laboratory. He is an assistant professor in the Language Technologies Institute, Carnegie Mellon University, where he leads the Multimodal Communication and Machine Learning Laboratory (MultiComp Lab). His research interests include computational study of nonverbal social communication, a multidisciplinary research topic that overlays the fields of multmodal interaction, computer vision, machine learning, social psychology, and artificial intelligence. He was selected in 2008 by *IEEE Intelligent Systems* as one of the 10 to Watch for the future of AI research. He received six best paper awards in multiple ACM- and IEEE-sponsored conferences for his work on context-based gesture recognition, multimodal probabilistic fusion, and computational modeling of human communication dynamics.