Posture, Relationship, and Discourse Structure

Models of Nonverbal Behavior for Long-Term Interaction

Daniel Schulman and Timothy Bickmore

College of Computer and Information Science, Northeastern University
360 Huntington Ave – WVH 202, Boston MA 02115
{schulman,bickmore}@ccs.neu.edu

Abstract. We present an empirical investigation of nonverbal behavior in long-term interaction spanning multiple conversations, in the context of a developing interpersonal relationship. Based on a longitudinal video corpus of human-human counseling conversation, we develop a model of the occurrence of posture shifts which incorporates changes that occur both within a single conversation and over multiple conversations. Implications for the design and implementation of virtual agents are discussed, with a particular focus on agents designed for long-term interaction.

Keywords: relational agent, embodied conversational agent, nonverbal behavior, relationship, posture, discourse structure

1 Introduction

Embodied Conversational Agents (ECAs) are increasingly applied to tasks which require or benefit from multiple conversations with each user, potentially over a long period of time; Examples include education, counseling, and social engagement. Maintaining long-term user engagement in such applications may be challenging [5].

Typically, ECAs are designed to accurately reproduce human verbal and nonverbal behavior, within the limits of their design. The conversational behavior of human dyads has been shown to change over time as their relationship evolves (e.g. [7]). Simulating these changes in an ECA may produce more realistic and engaging ECAs for long-term interaction.

This study is a step toward that goal. We focus on a single nonverbal behavior — posture shifts — which are both a part of the standard repertoire of many ECAs, and may also be an indicator of the interpersonal relationship of a dyad. Based on an examination of a longitudinal corpus of human-human interaction, which contains multiple conversations between the same dyads, we construct a model of the occurrence of posture shifts in conversation, including changes over time, both across and within conversations.

2 Background and Related Work

Several decades of research, dating at least to work by Scheflen [18], has focused on examining postural alignment or mirroring as an indicator of rapport. To the extent that a dyad is likely to build stronger rapport over multiple conversations, this predicts increasing postural alignment over time. However, empirical tests have been mixed [13, 14], Bernieri reports that movement synchrony (i.e. similarity in timing) may be an indicator of rapport while behavior matching (i.e. taking the same position at the same time) is not [3]. Tickle-Degnen and Gavett suggest that postural alignment may have a positive association with rapport only in later interactions, rather than initial interactions [20].

There is also a well-studied association between posture shifts and discourse structure. Many authors have noted that posture shifts tend to occur at topic boundaries (e.g. [12]). Cassell and Nakano et al. give empirical evidence of this phenomenon, based on an examination of direction-giving dialogues [8].

Little prior work examines any possible association between posture shifts and aspects of interpersonal relationship other than rapport. To our knowledge, no prior empirical work examines simultaneously an association between posture shifts and discourse structure and an association between posture and interpersonal relationship. Here, as an initial step, we focus on behavior changes that occur when a dyad has repeated interaction and a longer history together, rather than examining many other possible aspects of interpersonal relationship.

3 The Exercise Counseling Corpus

We collected a longitudinal corpus of dyadic interaction, containing multiple conversations between each dyad. This approach allows us to examine any changes over time in detail, and separate them from other differences between individuals. We chose to study behavior change counseling for exercise promotion, an area in which conversational agents have been applied (e.g., [4]), in which the counselor-client relationship has an effect on outcomes [15], and nonverbal behavior is associated with the development of this relationship [20].

We recruited clients (N=6) who stated they did not currently exercise regularly. Each client was asked to have up to 6 weekly conversations with a counselor. The same counselor conducted all conversations, and in each conversation attempted to encourage the client to increase his or her daily physical activity. All conversations were held in the same room, with both client and counselor seated in office chairs, and were videotaped from three angles. The participants were informed the conversations would be taped and examined, but were not told what behaviors were of interest.

The final corpus contains 32 conversations (mean duration 15.6 minutes), comprising approximately 8.3 hours of recorded video, and approximately 100,000 words of spoken dialogue.

4 Methods

We separately coded the exercise counseling corpus to identify occurrences of posture shifts and of topic shifts. Coding was performed by the primary author.

To check reliability, three conversations were randomly selected for each coding task and analyzed by a second coder.

4.1 Coding of Topic Shifts

Topic shifts were coded using transcripts of the corpus produced for a previously-reported study [19]. Video was not viewed in order to avoid confounding topic shifts with visible posture shifts or other nonverbal indicators of discourse structure. The transcripts were segmented based on the occurrence of silence, and topic shifts were coded as occurring at the beginning of the segment that introduced a new topic.

Following Grosz and Sidner [10], we defined a topic as a shared conversational goal to which the participants were mutually committed. A topic shift was marked whenever the coder believed a participant was attempting to introduce such a shared goal (whether or not the attempt was successful). The agreement rate between coders was 96.4% (Cohen's $\kappa = 0.68$).

4.2 Coding of Posture Shifts

Posture shifts were coded using muted video, in order to avoid confounding posture shifts with audible topic shifts. A posture shift was defined as a gross movement of the body, including the trunk, the legs and lower body, or both. Movements that appeared to be caused by the performance of a communicative gesture (e.g. a large hand gesture) were excluded, as were repetitive motions lasting more than a couple seconds (e.g. repeatedly rocking back and forth in the chair). Initially, chair rotation was coded as posture shifts, but preliminary examination revealed that these movements were very difficult to code reliably, and they were excluded.

Coders were asked to judge the start and end times of each posture shift, and several additional features, including movement direction, co-occurrence of grooming behavior (e.g. brushing hair or adjusting clothes), and an estimated energy level. Energy was judged on a linear scale ranging from 1 (the smallest perceptible shift) to 10 (the most energetic possible shift without leaving the chair).

An initial examination showed that reliability was very poor on low-energy shifts, and consequently all shifts with an energy of less than 4 were discarded. Aside from this, features of posture shifts besides the time of occurrence were not used in the present study. To compute inter-rater reliability, the corpus was divided into 1-second intervals, and each interval was considered to have been marked as a shift if the majority of it was covered by any coded posture shift. Cohen's κ was 0.58.

5 Results

31 conversations were coded for both posture and topic shifts; one conversation had large portions of unintelligible speech, and could not be coded for topic shifts.

Table 1. A mixed-effect logistic regression model predicting the probability of occurrence of posture shifts

Parameter	Est.	SE	p
Intercept	-4.17	0.21	.000***
Sessions (# of previous conversations)	0.16	0.05	.002**
Speaker (0=client, 1=counselor)	0.15	0.17	.376
Minutes (from the start of conversation)	-0.03	0.01	.027*
$\text{Tshift}_{\text{self}}$ (a topic shift occurs within 2 seconds)	0.94	0.31	.002**
$\text{Tshift}_{\text{other}}$ (a topic shift occurs within 2 seconds)	0.47	0.29	.103
Sessions × Speaker	-0.11	0.04	.010**
Sessions × Minutes	-0.02	0.00	.000***
Sessions × $\text{Tshift}_{\text{self}}$	-0.09	0.07	.225
Sessions × $\text{Tshift}_{\text{other}}$	-0.13	0.08	.110
Speaker × Minutes	-0.04	0.01	.013*
Speaker × $\text{Tshift}_{\text{self}}$	-0.52	0.24	.031*
Speaker × $\text{Tshift}_{\text{other}}$	0.22	0.27	.425
Minutes × $\text{Tshift}_{\text{self}}$	-0.02	0.02	.470
Minutes × $\text{Tshift}_{\text{other}}$	-0.03	0.03	.226
$\text{Tshift}_{\text{self}}$ × $\text{Tshift}_{\text{other}}$	0.16	0.63	.794

[a] Random intercepts on dyads (SD=0.26) and conversations (SD=0.33).
[b] Coefficients indicate change in the log-odds of a posture shift occurring within one second.
[c] All p values are derived from Z tests.

A total of 803 posture shifts were identified in the remaining conversations. The rate of posture shifts varied widely across conversations, ranging from 0.035 to 4.92 per minute (median 0.71).

The start time of each posture shift was aligned to the nearest second. We then modeled the occurrence of a posture shift as a binary outcome, with one observation per second. A logistic mixed-effect regression model [16] was used. This model generalizes logistic regression to account for observations that are non-independent due to being grouped or nested — in this case, within conversations and dyads — by adding "random effects" which model the group-level variance.

In order to model change over time, we included the number of previous sessions, and the time since the start of the conversation as predictors. To control for varying discourse structure, we included the co-occurrence of topic shifts, by both a speaker and their conversation partner, as predictors. To allow for variability among different speakers, dyads, and conversations, we include the speaker (counselor or client) as a predictor, along with random effects on dyads and individual conversations. Finally, we included two-way interactions among all predictors; a model comparison by Akaike Information Criterion [6] strongly preferred this more complex model (ΔAIC = 9.8).

The final model (Table 1) was fit using R 2.12.1 [17] and the lme4 [2] package. Posture shifts are significantly more likely to occur at topic shifts (the coefficient "$\text{Tshift}_{\text{self}}$" in Table 1); this replicates results by Cassell et al. [8]. We see no

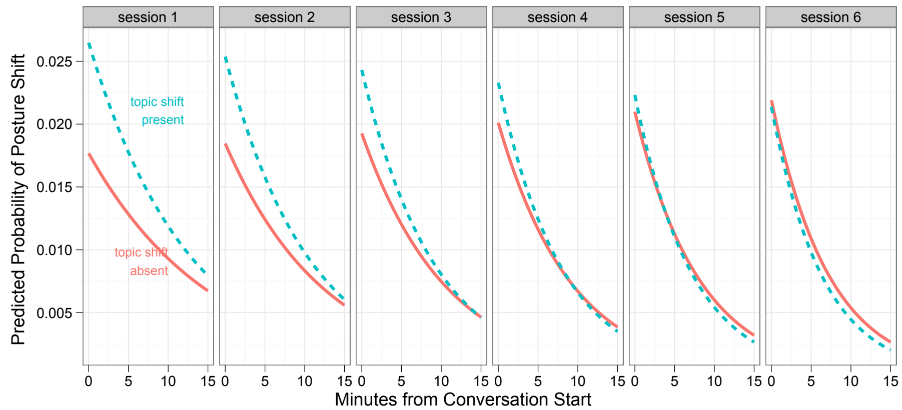

Fig. 1. Predicted probability of a posture shift occurring within a 1-second interval

significant effect for topic shifts (“$\text{Tshift}_{\text{other}}$”) introduced by the conversation partner rather than a participant, although there is a trend in the same direction.

There are significant changes over time, both within and across conversations. Figure 1 shows the predicted probability of a posture shift at different points over six weekly sessions, averaged over dyads. Each column shows a single session. Posture shifts occur much more frequently in the beginning of a conversation, as indicated by the steeply downward sloping lines in each column (and the coefficient “Minutes” in Table 1). There is an interaction with the number of previous sessions: the rate of decrease is greater in later conversations (“Sessions $\times$ Minutes”).

6 Discussion

We show evidence of changes in the occurrence of posture shifts over time, both within conversations and across multiple conversations. Based on prior work on the occurrence of posture shifts [8], nonverbal behavior generation frameworks (e.g., BEAT [9]) have implemented stochastic, rule-based approaches to the generation of posture shifts, where the probability of generation of a posture shift is determined by discourse structure. These results suggest that such models could be easily modified to produce more realistic behavior by decreasing the probability of a posture shift over time. Using the regression coefficients estimated here, the log-odds of a posture shift[1] would change by $(0.16s - 0.03m - 0.02sm)$, where s is the number of previous sessions, and m is minutes from the start of the conversation.

We do not yet have clear evidence for a mechanism behind these effects, but instead offer some conjectures, based partially on subjective examination of the

[1] the odds o of a probability p are $o = \frac{p}{1-p}$; the log-odds are $\log o = \log \frac{p}{1-p}$

corpus. The early portion of many conversations includes posture shifts that appeared to be part of a process of "settling in", with most shifts leaving the participant in a more relaxed body posture. A relaxed body posture is an indicator of nonverbal immediacy [1] (i.e., intimacy, warmth, or closeness). Increasingly rapid decreases in the rate of posture shifts in later conversations may indicate that, as a stronger interpersonal relationship develops over time, participants will more quickly and easily adopt a body posture that indicates high immediacy. Future work may investigate these conjectures through more detailed coding of posture shifts and by examining other indicators of immediacy for similar patterns of change.

Limitations of this study include a small number of participants (and a single counselor). The corpus is also limited to a single task, in a single setting, and validation of this model in an ECA is necessary. To address some of these limitations, we plan a longitudinal evaluation study, in which participants have multiple conversations with an ECA designed according to the model developed here.

Associations between nonverbal behavior and other aspects of interpersonal relationship are also of interest. We have collected longitudinal assessments of the strength of the counselor-client therapeutic alliance from participants in the exercise counseling corpus, using the Working Alliance Inventory [11], as well as other assessments of interpersonal relationship. Future work will incorporate this information into more complete models.

Finally, we hope that future work in this area will begin to develop more detailed and complete models of behavior, enabling the development of more lifelike, engaging, and efficacious virtual agents.

Acknowledgements. Thanks to Jenna Zaffini for her work in collecting the corpus, Brandon Gier and Connor Westfall for assistance in coding, and to the other members of the Relational Agents Group for much help and useful discussion. This work was supported by NSF CAREER IIS-0545932.

References

1. Andersen, P.A.: Nonverbal immediacy in interpersonal communication. In: Siegman, A.W., Feldstein, S. (eds.) Multichannel Integrations of Nonverbal Behavior, pp. 1–36. Lawrence Erlbaum, Hillsdale, NJ (1985)
2. Bates, D., Maechler, M., Bolker, B.: lme4: Linear mixed-effects models using S4 classes (2011), http://CRAN.R-project.org/package=lme4, R package version 0.999375-39
3. Bernieri, F.J.: Coordinated movement and rapport in teacher-student interactions. Journal of Nonverbal Behavior 12(2), 120–138 (Jun 1988)
4. Bickmore, T.: Relational Agents: Effecting Change through Human-Computer Relationships. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA (2003)
5. Bickmore, T., Schulman, D., Yin, L.: Maintaining engagement in long-term interventions with relational agents. Applied Artificial Intelligence 24(6), 648–666 (Jul 2010)

6. Bozdogan, H.: Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika 52(3), 345–370 (Sep 1987)
7. Cassell, J., Gill, A.J., Tepper, P.A.: Coordination in conversation and rapport. In: Workshop on Embodied Language Processing. pp. 41–50. Association for Computational Linguistics (Jun 2007)
8. Cassell, J., Nakano, Y.I., Bickmore, T.W., Sidner, C.L., Rich, C.: Non-verbal cues for discourse structure. In: ACL '01: Proceedings of the 39th Annual Meeting on Association for Computational Linguistics. pp. 114–123. Association for Computational Linguistics, Morristown, NJ, USA (2001)
9. Cassell, J., Vilhjálmsson, H.H., Bickmore, T.: BEAT: the Behavior Expression Animation Toolkit. In: SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques. pp. 477–486. ACM, New York, NY, USA (2001)
10. Grosz, B.J., Sidner, C.L.: Attention, intentions, and the structure of discourse. Computational Linguistics 12(3), 175–204 (1986)
11. Hatcher, R.L., Gillaspy, A.J.: Development and validation of a revised short version of the Working Alliance Inventory. Psychotherapy Research 16(1), 12–25 (Jan 2006)
12. Kendon, A.: Some relationships between body motion and speech. In: Seigman, A., Pope, B. (eds.) Studies in Dyadic Communication, pp. 177–216. Pergamon Press, Elmsford, NY (Jun 1972)
13. Lafrance, M.: Nonverbal synchrony and rapport: Analysis by the Cross-Lag panel technique. Social Psychology Quarterly 42(1), 66–70 (1979)
14. LaFrance, M., Ickes, W.: Posture mirroring and interactional involvement: Sex and sex typing effects. Journal of Nonverbal Behavior 5(3), 139–154–154 (Mar 1981)
15. Martin, D.J., Garske, J.P., Davis, M.K.: Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. Journal of consulting and clinical psychology 68(3), 438–450 (2000)
16. McCulloch, C.E., Neuhaus, J.M.: Generalized Linear Mixed Models. John Wiley & Sons, Ltd (2005)
17. R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2011), http://www.R-project.org/, ISBN 3-900051-07-0
18. Scheflen, A.E.: The significance of posture in communication systems. Psychiatry 27, 316–331 (Nov 1964)
19. Schulman, D., Bickmore, T.: Modeling behavioral manifestations of coordination and rapport over multiple conversations. In: Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., Safonova, A. (eds.) Intelligent Virtual Agents, Lecture Notes in Computer Science, vol. 6356, chap. 14, pp. 132–138. Springer Berlin / Heidelberg, Berlin, Heidelberg (2010)
20. Tickle-Degnen, L., Gavett, E.: Changes in nonverbal behavior during the development of therapeutic relationships. In: Philippot, P., Feldman, R.S., Coats, E.J. (eds.) Nonverbal behavior in clinical settings, chap. 4, pp. 75–110. Oxford University Press, New York, NY, USA (2003)