Modeling Rapport in Embodied Conversational Agents

Justine Cassell

Center for Technology and Social Behavior, Northwestern University

justine@northwestern.edu

What characterizes the language of people who have known one another for a long time? In the US one thinks of groups of friends, leaning in towards one another, smiling broadly, and finishing one another's sentences in their eagerness to contribute to the conversation. The details may differ from culture to culture, but the fact of differences between groups of friends and groups of strangers seems universal. And these styles of communication play an important role in identifying people who like and respect us, people we can trust, and people we wish to collaborate with, from the rest of the people in the world. In fact, those differences in communication style, and the feelings they engender, have important consequences for our behavior and actions. For example, we know that children learn better in interactions with teachers and peers that they like, and whom they feel like them. And people who are incapable of reading rapport language or rapport non-verbal behavior or of producing this kind of communicative style – such as people with autism spectrum disorder – not only may be more likely to suffer from depression and loneliness, but also have trouble learning and collaborating. However, we also know that rapport is not always desirable in interaction; in fact, different kinds of interpersonal affiliation behaviors can have different kinds of effects. For instance, the literature suggests that one type of rapport behavior leads survey respondents to answer in ways that they think will please the interviewer, even if these answers are dishonest, while another type of rapport behavior leads respondents to be as truthful as they can, regardless of the social desirability of their responses.

Given the importance of interpersonal rapport, it is surprising that so little attention has been paid to implementing notions of long-term rapport in the machines we communicate with. Advances in natural language and speech technology are ensuring that dialogue engines are increasingly emotionally expressive. For the most part, however, those same dialogue engines demonstrate amnesia, beginning every conversation with a user as if it is their first. Even embodied dialogue engines, or Embodied Conversational Agents (ECA), which are increasingly realistic and personality-rich, act like Drew Barrymore in the movie "50 First Dates" – always pleasant, but always meeting somebody for the first time.

Because of the importance of notions of rapport, and our belief that rapport can have important consequences for human-machine interaction in domains as diverse as intelligent tutoring systems, assistive technologies, GPS direction-giving, and online survey interviewing, my students and I are currently carrying out a number of different studies to better understand rapport, and to implement it in Embodied Conversational Agents. We are not the only researchers in the field of ECAs to examine this phenomenon. [7] have looked at how ECAs might change their personality and behaviors according to the likes and dislikes of a user, in order to become friends. [1] have reflected on the ways in which computers might need to act in order to maintain relationships with users over time. [5] has looked at the kinds of behaviors that lead users to say that they feel more rapport with an ECA in a first encounter. In our own previous work [2], we looked at the role of small-talk in conversation, and employed our results to implement ECAs that invited increased trust from their users.

Our work is characterized, however, by a reliance on models derived from study of human-human behavior. That is, we approach the issue by first carrying out a series of studies on rapport in human-human interaction, and using the results to model rapport in ECAs. Importantly, that series of studies attempts to parameterize rapport so that the resultant models can be used across a variety of situations, and so that the *functions* of rapport are tightly tied to the surface-level *behaviors* that the agent engages in with a user. Our models join short-term audience design (i.e. grounding phenomena) to long-term audience design (relationship and rapport). And, importantly, our analyses of the data we collect address both language and nonverbal behaviors, and both speech and dis-

course characteristics. Thus, in current work we are observing children as they learn with peers, and attempting to understand the role that dialect (as expressed in speech, and in syntax and morphology) and nonverbal behavior play in indexing feelings of "alikeness" among the children [6]. We are also observing dyads of friends and dyads of strangers as they give directions, and analyzing the relationship between dialogue acts and nonverbal behaviors in the interaction styles of these two groups [3]. We are studying professional survey interviewers, and comparing the behaviors of those who get the most honest responses to those who are more likely to obtain socially-desirable responses [4]. And finally, we are observing typically-developing children and children with autism spectrum disorder, as they play with other children, and comparing their ability to use pragmatic devices such as contingency to previous utterances, seamless turn-taking, reciprocity in conversation, and topic management [8].

In each instance, our studies are leading to implemented ECAs that we then test in the original community. We are therefore implementing virtual survey interviewers that can use rapport to elicit truthful responses, and virtual direction-giving agents that behave differently as they give directions over the lifetime of use. We are also implementing virtual peers that can engage in collaborative learning with children within different dialect communities, virtual peers that can scaffold the learning of rapport behaviors in children with autism spectrum disorder, and virtual peers that can be used to assess the social skills deficits of children with autism spectrum disorder so as to better plan their treatment. The goal of the research program is to better understand linguistic and social coordination devices from the utterance level to the relationship level – how they work in humans, how they can be modeled in virtual humans, and how virtual humans can be implemented to help humans have productive and satisfying relationships, with machines and with one another, over long periods of time

Acknowledgments

Thanks to all of the members of the ArticuLab research group at Northwestern University, and to the Autism Speaks Foundation, Charles Cannell Fund in Survey Methodology, Hiatt Fund, Motorola, the National Association of Autism Research, and the National Science Foundation, for generous funding.

References

[1] Bickmore, T. and Picard, R., (2005) "Establishing and Maintaining Long-Term Human-Computer Relationships," *ACM Transactions on Computer Human Interaction (ToCHI),* vol. 12, pp. 293-327.

[2] Cassell, J. and Bickmore, T., (2002) "Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents," *User Modeling and Adaptive Interfaces,* vol. 12, pp. 1-44.

[3] Cassell, J., et al., (2007) "Coordination in Conversation and Rapport," in *Proceedings of Workshop on Embodied Natural Language, Association for Computational Linguistics Annual Meeting*, Prague, CZ.

[4] Cassell, J. and Miller, P., (2007) "Is It Self-Administration if the Computer Gives You Encouraging Looks?," in *Envisioning the Survey Interview of the Future*, F. G. Conrad and M. F. Schober, Eds. Hoboken, N.J.: John Wiley & Sons, Inc., pp. 161-178.

[5] Gratch, J., et al., (2006) "Virtual Rapport," in *Proceedings of 5th International Conference on Interactive Virtual Agents (IVA)*, Marina del Rey, CA.

[6] Iacobelli, F. and Cassell, J., (2007) "Ethnic Identity and Engagement in Embodied Conversational Agents," in *Proceedings of 7th International Conference on Intelligent Virtual Agents*, Paris, France, pp. 57 -- 63.

[7] Stronks, B., et al., (2002) "Designing for friendship: Becoming friends with your ECA," in *Proceedings of Embodied conversational agents - let's specify and evaluate them!*, Bologna, Italy, pp. 91-97.

[8] Tartaro, A. and Cassell, J., (2007) "Peer Technology as an Intervention for Children with Autism," in *Towards Universal Usability: Designing Computer Interfaces for Diverse User Populations*, J. Lazar, Ed. Chichester, UK: John Wiley, pp. 231-262.