Modeling Side Participants and Bystanders: the Importance of Being a Laugh Track

Jina Lee and Stacy Marsella

University of Southern California
Institute for Creative Technologies
12015 Waterfront Drive, Playa Vista, CA 90094, USA
{jlee, marsella}@ict.usc.edu

Abstract. Research in virtual agents has largely ignored the role and behavior of side participants and especially bystanders. Our view is that the behavior of these other participants is critical in multi-party interactions, especially in interactive drama. In this paper, we provide an analysis of nonverbal behaviors associated with these roles. We first review studies of interpersonal relationships and nonverbal behavior. From this review, we construct an analysis framework based on characters' interpersonal relationships, conversational roles, and communicative acts. We then assess this framework by analyzing improv sessions of an old west scenario involving 4 characters. Informed by this analysis, we implemented a general model for participant and bystander behavior.

Keywords: Virtual Agents, Embodied Conversational Agents, Nonverbal Behaviors, Multi-modal Communication

1 Introduction

Imagine an old cowboy western. The good sheriff is at the bar having a whisky. The evil gunslinger, sworn enemy of the sheriff, enters the bar. All the people initially gaze at the gunslinger, some start to move away wanting to avoid trouble. Others avert their gaze, pretend to do something else, hoping not to be noticed by the gunslinger who runs the town. Such scenes make it clear that much of the drama of a performance is in the reactions provided by the living backdrop of other performers. The performance arts have long acknowledged the important role that audience response plays. In Greek drama, the chorus served in part to provide the context of an ideal audience response for the actual audience. In modern day TV and films, laugh tracks are added to stimulate audience responses. The fact that our responses are mediated by others' responses has also influenced theatrical actors, trained to react to the main action [5] and film editing's use of reaction shots. A variety of psychological theories, such as social comparison theory, social referencing and emotional contagion, similarly argue that the social milieu influences the individual.

With few notable exceptions, research in virtual agents has largely ignored the role of conversation participants other than the speaker and the addressee,

perhaps because many virtual agent applications are limited to dyadic interactions. In this paper, we discuss how to analyze and model the behaviors of not only the core conversation participants, but also the side participants and bystanders that can influence the human observer's reaction to the interaction. We focus on modeling them without requiring a complex cognitive system that forms the goals of the agents.

The approach we take is based on psychological research on interpersonal relational dimensions that inform the pattern of people's interactions. Interpersonal circumplex theories [6, 9] argue for two fundamental dimensions, *affiliation* (hostility-friendliness) and *control* (dominance-submissiveness), to explain how different kinds of action elicit predictable responses from others [6]. Interaction partners along the affiliation dimension will elicit similar behaviors (e.g. friendly behaviors evoke friendly behaviors) and those along the control dimension will elicit complementary behaviors (e.g. dominant behaviors evoke submissive behaviors). These dimensions are common to a range of psychological theories and are often used in work of nonverbal behaviors [1, 2].

In this paper, we present an analysis framework based on the agents' interpersonal relationships, communicative acts, and conversation roles. We then analyze a set of improv sessions of an old west gunslinger scenario, which includes dramatic behaviors that convey rich interpersonal relationships and emotional reactions, as a testbed to construct mappings to various nonverbal behaviors. This mapping provides us with a model for the behaviors of the participants in the interaction.

2 Analysis Framework

Inspired by interpersonal circumplex theories and techniques in theatrical performances, we first developed an analysis framework to study the behaviors exhibited in a set of improv sessions. Below we provide more details about how we define the analysis framework.

Conversation Roles: We define four conversation roles: speaker, addressee, side participant, and bystander. Speaker and addressee are the core participants of the conversation whereas side participants are the "un-addressed recipients" of the speech at the moment [3][4]. Bystanders are openly present in the environment but do not participate in the conversation.

Interpersonal Relationships: The relationship between characters is described in terms of dominance and friendliness, following theories of interpersonal circumplex [6, 9]. In the gunslinger scenario, there are four characters: Rio, Harmony, Utah, and Ranger. Rio is the dominant and hostile character. Harmony is submissive to Rio and 'acts' friendly to him when in truth she dislikes Rio. She is neither dominant nor submissive to Utah and Ranger, but is particularly friendly to Ranger. Similar to Harmony, Utah is submissive to Rio and 'acts' neutral to him in terms of friendliness, but in truth he dislikes Rio. Ranger has neutral relationships with all the characters in terms of dominance and friendliness. Table 1 specifies the interpersonal relationships of the characters.

Table 1. Interpersonal relationship between Gunslinger characters in terms of dominance and affiliation. The symbol in parenthesis represents the masked relationship a character hides from the other character. (D: dominant, S: submissive, H: hostile, F: friendly, N: neutral)

	Rio	Harmony	Utah	Ranger
Rio	-	D/H	D/H	D/H
Harmony	S/F(H)	-	N/N	N/F
Utah	S/N(H)	N/N	-	N/N
Ranger	N/N	N/N	N/N	-

Communicative Acts: Communicative acts are broadly constructed here to include not only the dialogue acts of character utterances but also events that take place which may engender emotional responses from the characters. For example, Rio's entrance into the saloon or even a mere mention of his name may cause strong fear within Utah and Harmony. The following lists the communicative acts defined in the Gunslinger scenario: (S) asks-question-to (A); (S) confirms (A); (S) disconfirms (A); (S) requests (A); (S) accepts-request-of (A); (S) declines-request-of (A); (S) suggests (A); (E)/(C) threatens (C); (E) removes-threat-from (C). (S) and (A) indicate speaker and addressee, (C) indicates character and (E) indicates event. Although we can define a much richer set of communicative acts, here we simplify them to find a commonality and see how the characters' conversation roles or relationships may influence the behaviors given the same communicative act.

3 Analysis and the Behavior Model

To assess the analysis framework described in the previous section, we used it to analyze the Gunslinger improv sessions and mapped each factor of the framework to various nonverbal behaviors exhibited in the videos. We first describe the Gunslinger improv sessions, then the result of the analysis, from which we construct the behavior model.

Gunslinger Improv Sessions: The setting of the Gunslinger Improv sessions is a saloon somewhere in western USA, circa the 1800s. The characters in Gunslinger are extreme stereotypes drawn from the mythology of the Old West: the friendly bartender Utah, the psychotic gunslinger Rio, the fille de joie Harmony and the lawman Ranger. The script begins with Utah and Harmony talking about Rio, who runs the town. The Ranger then enters the bar looking to arrest Rio, unaware of how bad Rio is. Rio stops in the bar on his way to get some smokes in order to tell Harmony that they are leaving town. As Harmony rejects the idea, he shoots up the bar, a nonverbal way of emphasizing that he is the one in control. Upon seeing Ranger's badge, Rio threatens to kill him if he is still in town when he gets back from buying tobacco. Rio exits, leaving Utah, Harmony and Ranger to plot his demise. Upon Rio's return the gunfight ensues.

A troop of 8 actors were recruited and videotaped performing the improvisations based on the script (see Fig. 1 for the improv setting). The actors were broken up into two groups of four, playing the roles of Rio, Harmony, Utah and the Ranger. The actors in the same group performed once following the script

Fig. 1. Improv of Gunslinger scenario

Fig. 2. The Gunslinger set

and twice improvising. Each session lasted for about 7-10 minutes. These improv sessions were undertaken to inform the design of a mixed-reality interactive virtual human entertainment experience (see Fig. 2 for the set design) developed at the University of Southern California.

Results of the Analysis and the Behavior Model: Using the analysis framework, we constructed a corpus mapping the factors of the framework to various nonverbal behaviors. Here we focused on gaze, posture shifts, and the dynamics of physical distance between characters (e.g. approach vs. move away). Table 2 shows this mapping exhibited in the video. The following summarizes the results of the video analysis and modifications made to the analysis framework.

First of all, we made several modifications to the list of communicative acts to capture important reactions exhibited by the actors. Communicative act *(S) accepts-request-of (A)* was generalized to *(S) informs (A)* because the analysis showed no differences in the characters' behaviors. On the other hand, we created new communicative acts to handle the cases when characters showed emotional reactions (i.e. (S) expresses-to (A)) and to differentiate different levels of threats exerted by Rio. Finally, we did not observe any cases of communicative act *(S) suggests (A)*.

Foreshadowing behaviors were displayed mainly at the beginning of the Gunslinger scenario. Rio, being dominant and hostile, imposes a large threat to others and tensions are built up or released depending on his actions. For instance, when Rio enters the saloon, his presence engenders fear within other characters, which causes avoidance behaviors such as gazing away or stepping back. When Harmony refuses Rio's order to pack and leave the town with him, Utah shifts his gaze nervously between Rio and Harmony as if to expect something bad to happen. These foreshadowing behaviors informs the audience that Rio is associated with danger and threat.

As expected, the interpersonal relationship was found to affect the behaviors of characters, even when they were not one of the core conversation participants. For example, Harmony showed completely different attitudes toward Ranger and toward Rio (flirtatious vs. submissive). Utah also exhibited different behaviors as a bystander. When Harmony speaks to Ranger, Utah holds a more relaxed posture, whereas when she speaks to Rio, he crouches his posture, puts his head down, and avoids gaze or quickly glances between Harmony and Rio.

The main difference between listener and bystander behaviors was in the gaze. Addressees mainly looked directly at the speaker (with possibly different

Table 2. Mapping from communicative act, conversation role, and interpersonal relationship to nonverbal behaviors (S: speaker, A: addressee, SP: side participant, B: bystander, D: dominant, S: submissive, H: hostile, F: friendly, N: neutral).

Case	Communicative Act	Conv. Role	Interp. Rel.	Nonverbal Behaviors
1-1	**(S) asks-question-to (A)**	S	N/N	Look at (A)
	(Harmony asks Utah or Ranger)	S	N/F	Look at (A), lean forward, smile
		A	N/N	Look at (S)
		SP	N/N	Look between (S) and (A)
		B	N/N	Quick glances between (S) and (A) occasionally
1-2	**(S) asks-question-to (A)**	S	D/H	Look directly at (A), erect posture
	(Rio asks Utah /Harmony or Ranger)	A	S/N(F), S/F(H)	Look at (S), crouched posture, head down, may step back
	(Rio asks Ranger)	A	N/N	Look at (S)
		B	N/N	Look between (S) and (A) (attention drawn)
		B	S/N(F), S/F(H)	Gaze aversion, occasionally glances between (S) and (A), crouched posture, head down
1-3	**(S) asks-question-to (A)**	S	N/N	Look at (A)
	(Ranger asks Rio)	A	D/H	Look at (S), erect posture
		B	S/N(F), S/F(H)	Gaze aversion, occasionally glances between (S) and (A), may step back, crouched posture, head down
2	**(S) confirms (A)**	S	N/N	Look at (A), head nod(s)
	(Ranger confirms Utah/Harmony)	S	N/F	Look at (A), relaxed posture, smile
	(Harmony confirms Ranger)	A	N/N	Look at (S), may nod after (S)'s speech
		B	N/N	Quick glances between (S) and (A) occasionally
3	**(S) disconfirms (A)**	S	N/N	Look at (A), head shake(s)
	(Ranger disconfirms Utah/Harmony)	A	N/N	Look at (S)
		B	N/N	Quick glances between (S) and (A) occasionally
4-1	**(S) informs (A)**	S	N/N	Look at (A)
	(Harmony, Utah, and Ranger talk to each other)	S	N/F	Look at (A), lean forward, smile (being flirtatious)
		A	N/N	Look at (S)
		B	N/N	Glances between (S) and (A) occasionally, relaxed posture
4-2	**(S) informs (A)**	S	S/F(H)	Look at (A), crouched posture, uneasy smile (masking fear)
	(Harmony or Utah talks to Rio)	S	S/N(H)	Look at (A), crouched posture
		A	D/H	Look at (S), erect posture
		SP	N/N	Look between (S) and (A)
		B	S/N(H), S/F(H)	Crouched posture, gaze aversion, head down, quick glances between (S) and (A)
4-4	**(S) informs (A)**	S	D/H	Look at (S), erect posture
	(Rio talks to Harmony or Utah)	A	S/F(H), S/N(H)	Look at (S), crouched posture
		SP	N/N	Look between (S) and (A)
		B	S/F(H), S/N(H)	Crouched posture, gaze aversion, head down
5-1	**(S) informs-negative-to (A)**	S	N/N	Look at (A)
	(Utah tells Harmony Rio is coming)	A	N/N	Look at (S)
	(Harmony describes Rio to Ranger)	B	N/N	Look between (S) and (A)
5-2	**(S) informs-negative-to (A)**	S	D/H	Glares at (A), erect posture
	(Rio threatens to kill Ranger)	A	N/N	Look at (S), erect posture
		B	S/N(F), S/F(H)	Look between (S) and (A), crouched posture, step back, distressed expression
6-1	**(S) expresses-to (A)**	S	N/N	Look at (A), lean forward, brow frowned, head shakes, disgusted face (nose wrinkle, squinted eyes)
	(Harmony tells Utah not to joke)			
	(Utah tells Ranger about Rio)	A	N/N	Look at (S)
		B		*(not present in the scenario)*
6-2	**(S) expresses-to (A)**	S	N/F	Eyes open, brow raise, lean forward
	(Harmony is surprised that Ranger haven't heard about Rio)	A	N/N	Look at (S)
		B		*(not present in the scenario)*
7-1	**(S) requests (A)**	S	N/F	Inner brow raise, lean forward (almost begging)
	(Harmony asks Ranger to help get rid of Rio)	A	N/N	Look at (A)
		B	N/N	Look between (S) and (A) (attention drawn)
7-2	**(S) requests (A)**	S	D/H	Straight gaze at (A), pound foot on ground, lean forward (as if attacking)
	(Rio tells Ranger to be quiet)	A	N/N	Look at (A)
		B	S/N(F), S/F(H)	Look between (S) and (A), crouched posture, may step back

7-3	**(S) requests (A)** *(Rio orders Harmony to pack)*	S A SP B	D/H S/F(H) N/N S/N(H)	Straight gaze at (A) Look at (A), distressed expression Look between (S) and (A) Look between (S) and (A), crouched posture, may step back
7-4	**(S) requests (A)** *(Rio orders Utah to countdown for gunfight)*	S A SP B	D/H S/N(H) D/N S/F(H)	Straight gaze at Ranger Look at (A), crouched posture, distressed expression Look at (A) Look around characters, crouched posture, step back, distressed expression
7-5	**(S) requests (ALL)** *(Rio tells everyone to be prepared when he comes back)*	S A A	D/H S/F(H)S/N(H) N/N	Look around Look at (A), crouched posture, distressed expression Look at (S), neutral posture
8	**(S) declines-request-of (A)** *(Harmony refuses to leave with Rio)*	S A SP B	S/F(H) D/H N/N N/N	Lean forward, brows raised, crouched posture (as if begging) Look at (A), erect posture Look between (S) and (A) Look between (S) and (A), crouched posture, distressed expression
9	**(E) threatens-1st (ALL)** *(Rio enters the saloon)*	Utah, Harmony Ranger	S/N(F), S/F(H) N/N	Gaze aversion, step back, crouched posture, head down Look towards Rio
10	**(E) threatens-2nd (ALL)** *(Rio enters the saloon the second time)*	Utah, Harmony Ranger	S/N(F), S/F(H) D/N	Look at Rio, step back, crouched posture, head down Look straight at Rio, erected posture, body oriented to Rio
11	**(C) threatens (ALL)** *(Rio shoots gun)*	Utah, Harmony Ranger	S/N(F), S/F(H), N/N	Startled, duck down
12	**(E) heightens-threat** *(Countdown to gunfight)*	Rio, Ranger Utah Harmony	D/H, D/N S/N(H) S/F(H)	Erect posture, stare at each other Look between Rio and Ranger, crouched posture Eyes wide open, look between Rio and Ranger, hand to face (panic)
13	**(E) removes-threat-from (ALL)** *(Rio leaves the saloon)*	Utah, Harmony Ranger	S/N(F), S/F(H), N/N	Relaxed posture, mutual gaze among characters
14	**threat-removed-permanently** *(Rio is killed)*	Utah, Harmony Ranger	S/N(F), S/F(H) D/N	Mutual gaze among characters, relaxed posture, eyebrow raise, smile Mutual gaze among characters, relaxed posture, eyes open

postures) but bystanders displayed more gaze aversion or quick glances between the speaker and the addressee using only the eyes (i.e. gaze without revealing gaze). Only when the bystander felt strong fear or surprise did they make more obvious gaze movements between the speaker and the addressee.

Implementation: The mapping shown in Table 2 has been constructed as a set of rules within the Nonverbal Behavior Generator (NVBG) [7], the behavior planner of our virtual human system. For speakers, the communicative act and conversation role further modify the existing nonverbal behavior rules, especially with regards to posture and facial expressions. For non-speakers, new rules were added to generate listener or bystander behaviors. During the system initialization step, the NVBG receives a message specifying the interpersonal relationships of each character and registers this information. Upon receiving an input message including the speaker utterance and communicative acts from the dialogue manager [8], NVBG then detects the agent's conversation role and checks the interpersonal relationship with the speaker or addressee and selects the corresponding nonverbal reaction from the rules.

4 Conclusions and Future Work

In this paper, we presented an analysis framework and a model of side participant and bystander behaviors along with speaker and addressee behaviors based on interpersonal circumplex theories, techniques in theatrical performance, and our own analysis of improvised acting. The model could be used to improve the capability of not only the core conversation characters but also the background characters by generating appropriate reactions that reveal relational factors of the characters. These behaviors can lead to more dramatic impact on human participants and observers.

In the future, we plan to extend our model to include a wider range of interpersonal relationships and communicative acts. The work presented here is based on a limited data and we wish to collect a larger corpus to cover more diverse interpersonal relationships and communicative acts. We also plan to evaluate the model with human participants. Of particular interest is in how the side participant and bystander behaviors impact the user experience including whether they reveal the agents' relational factors, improve the perception of agents, and increase the user's engagement level.

Acknowledgments This work was sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM). The content does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

References

1. Argyle, M.: Bodily communication. Internat. Universities Press, New York (1975)
2. Burgoon, J.K.: Nonverbal Communication: The Unspoken Dialogue. Harper-Collins College Div (1988)
3. Clark, H.: Using Language. Cambridge University Press (1996)
4. Goffman, E.: Footing. Semiotica 25(1/2), 1–29 (1979)
5. Johnson, W.L.: Dramatic expression in opera, and its implications for conversational agents. In: Workshop on Embodied Conversational Characters as Individuals, AAMAS 2003 (2003)
6. Leary, T.: Interpersonal diagnosis of personality; a functional theory and methodology for personality evaluation. Ronald Press, New York (1957)
7. Lee, J., Marsella, S.: Nonverbal behavior generator for embodied conversational agents. In: Proc. of 6th Int. Conf. on Intelligent Virtual Agents. pp. 243–255 (2006)
8. Leuski, A., Traum, D.: A statistical approach for text processing in virtual humans. In: Proc. of 26th Army Science Conference. Orlando, Florida, USA (Dec 2008)
9. Wiggins, J.S.: A psychological taxonomy of trait-descriptive terms: The interpersonal domain. Journal of Personality and Social Pscyhology 37, 395–412 (1979)