Towards Developing a Model to Handle Multiparty Conversations for Healthcare Agents

Reshmashree Bangalore Kantharaju
Institut des Systèmes Intelligents et de Robotique (ISIR), Sorbonne Université
Paris, France
bangalore_kantharaju@isir.upmc.fr

Catherine Pelachaud
Centre national de la recherche scientifique (CNRS) - Institut des Systèmes Intelligents et de Robotique (ISIR), Sorbonne Université, Paris, France
catherine.pelachaud@upmc.fr

ABSTRACT

Virtual agents are playing a key role in several application domains for learning, gaming, healthcare. Social virtual agents are being used in different healthcare applications to provide assistance, coaching, explanation. Most existing applications rely on dyadic interaction. However multiparty conversation where each agent can be designed with a specific role has been shown to increase the passing of messages. One of the most common aspect of human-agent interaction is turn-taking, where the participants take *turns* to speak and allow for a smooth exchange. Several nonverbal signals and lexical and semantic information is used to facilitate this. Several studies have provided models to handle turn-taking in conversations involving three or more participants. The Council of Coaches project aims to provide virtual coaching to users by making use of multiple healthcare agents with various domain expertise. To achieve this, we aim to develop a model that will be able to handle multi-conversation and adapt the turn strategy depending on the group size and the agent role dynamically. This paper presents an overview of the existing literature and the research questions we intend to address.

KEYWORDS

Turn-taking, Multi-agent system, Healthcare agents, Nonverbal signals

1 INTRODUCTION

Virtual humans are now becoming more affect-sensitive and have been able to incorporate social skills to build and maintain rapport [13]. The initial research on virtual agents focused on enabling the agents to be involved in dyadic conversations, making them human-like by displaying verbal and non-verbal signals. Now, researchers are focusing on making the agents context aware, engaging and be able to understand human conversational dynamics to handle multi-party conversations and adapt accordingly to provide rich human-computer interaction [15, 39, 40]. Virtual agents are used in several application domains for learning [34], coaching [1], and even in healthcare domain [9]. Embodied conversational coaches are being developed to be companions or coaches. Studies show that patients in general withhold information due to the stigma and fear of being perceived in negative light [11, 19]. Virtual humans can play a key role in providing a 'safe' environment, that can motivate the users for honest disclosure of important information [22].

The average human life expectancy has increased significantly over the past decades due to advancements in health treatment and care. Subsequently, older adults living under the effects of various chronic conditions has also increased substantially. Maintaining a healthy lifestyle can help in preventing chronic conditions and improve the overall quality of life. In this paper, we present the Council of Coaches project that aims to provide virtual coaching involving multiple autonomous agents to motivate and educate the user about healthy lifestyle habits through interactive group discussions. Use of a multi-agent system would allow the user to be more engaged in the discussion. Also, multiple agents can provide different perspectives on the same issue and facilitate the users approaching old age to understand better life style s/he should undertake and the medical treatment s/he should follow. We first present the overview of the project followed by the focus of research. In Section 2, we present existing works on virtual agents in healthcare domain and the various multiparty models that exist. In Section 3, we present the research motivations and approach for developing a model to handle multi-party conversations.

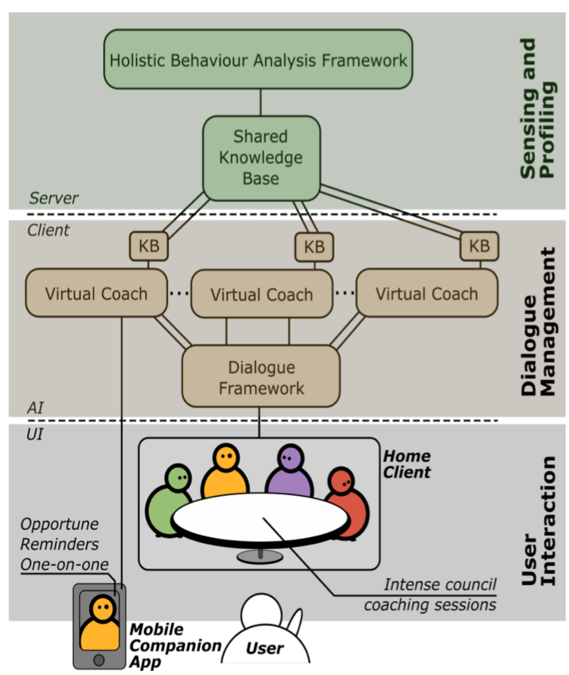

Figure 1: Overall concept overview for the Council of Coaches project.

1.1 Council of Coaches Project

The Council of Coaches is a European Horizon-2020 project that aims to develop a tool to provide virtual coaching for aged people to improve their physical, cognitive, mental and social health [24]. The council consists of a number of virtual embodied agents, each specialized in their own specific domain. They interact with each other and with the user to inform, motivate and discuss about issues related to health and well-being.

Six different expert coaches will be developed with knowledge in specific domains (see Figure 2). The core virtual coaches are Social-, Cognitive-, Physical-, and Mental Coaches and health specific coaches e. g., Diabetes and Chronic Pain coaches. The coaches will be designed to have different physical appearances, roles and use different persuasive strategies i. e., emotional or rational. The coaches will respond in a highly personalized manner based on the holistic behaviour modeling and analysis component which will detect short-term behaviour events and long term behaviour trends using sensors (on-body and home-environment). A multi-party Dialogue and Argumentation framework will provide the dialogue actions to be communicated by the virtual coaches based on [18]. It will introduce, analyze, and represent at a formal level, aspects of argumentation, dialogue, persuasion and explanation between (i) multiple virtual agents, and (ii) multiple and individual virtual agents and a user. The Council of Coaches will reside in the cloud environment and actively build and maintain a shared knowledge base in the form of user models, context models, interaction models. The coaches will be enabled to maintain their own private knowledge base, to improve the interactions.

The main contribution of this research work will be, to develop a multi-party conversation model that will facilitate turn-taking in a dyad or multi-party scenario that involves human user and virtual agents. The interaction will allow both reactive and proactive user participation and the speaking turns will be driven either by the users or the agents to ensure mixed initiative. The model will also be able to handle interruptions and manage dynamic group conversations in a multiparty scenario based on the role and domain of expertise. Further, we aim to maintain user's engagement by, gathering data on user's emotional state, providing information to the user appropriately, and ensuring that the user maintains involvement in the interactions with the coaches. This work will make use of and build upon the GRETA/VIB platform [25] for multimodal behaviour generation and for visualizing virtual coaches.

Figure 2: Embodied conversational coach archetypes.

2 RELATED WORK

In this section, we first present the virtual agents developed to assist users in the healthcare domain and then we provide an overview of the existing models developed to simulate and handle multiparty conversation.

2.1 Healthcare Agents

Several virtual agents have already been developed in the healthcare domain to aid patients in various aspects. In [9], a virtual human was developed for conducting interviews for healthcare support and it was shown that participants reported willingness to disclose, willingness to recommend and general satisfaction with the system [33]. In [38], Kristina, an ECA is developed to provide healthcare advice, assistance and act as social companions for the elderly. In [5] a virtual agent was used to explain health documents to patients. Evaluation study showed that, participants preferred agent over a human also reported that they felt more comfortable with the agent since they could ask for repetitions. ECAs are increasingly being considered and adopted for psychotherapeutic interventions [26] as well. A virtual agent was designed to improve communication skills among users with autism spectrum disorder [32]. A meditation coach was developed in [30], to aid the users through meditation to help users to relax. In [23] Monkaresi, developed a diabetes coach that helps individuals manage prescribed exercise, nutrition, monitoring blood glucose levels, and medication adherence. In [4], an agent was developed to help individuals with schizophrenia to manage their condition. A virtual agent that assists patients with chronic pain and depression was developed in [31] and patients reported significant improvements in depressive symptoms, social support, stress. In [21], virtual agent was able to perceive the user's facial expressions and text entries and deliver motivational intervention in an empathetic way. A relational agent in the role of exercise advisor was developed in [3] to be used by older adults in their homes on a daily basis over an extended period of time and the results showed an increase in physical activity. An agent was designed to provide social support for older adults and the results showed high levels of acceptance and satisfaction, by successfully providing a sense of companionship [37].

2.2 Multiparty Models

There are several models in the literature that are enabled to handle turn-taking and interruptions in dyadic conversations [8, 10, 20] and models that use different conversational settings e. g., presentation (interactive or non-interactive), multi-agent presentation or interaction with users [28, 29]. In this section, we describe the various multi-agent models that assign turns and handle interruptions in multi-party conversations.

Padilha and Carletta [12] presented a basic model for simulation of agents engaged in group conversation that involved turn-taking and the associated non-verbal behaviours. The agents were modeled to be independent and defined by a set of attributes i. e., *talkativeness*, likelihood of wanting to talk; *transparency*, likelihood of producing explicit positive and negative feedbacks, and turn- claiming signals; *confidence*, likelihood of interrupting, and continuing to speak during simultaneous talk; *interactivity*, the mean length of turn segments between TRPs; *verbosity*, likelihood of continuing

Table 1: An overview of the existing multiparty models

Model	Group size	Kind	Roles	States / Actions / Parameters	Factors	Handles	Non-Verbal Behaviours
[12]	>3	Agents	Speaker Addresse Side Participant	Talkativeness Transperency Confidence Interactivity Verbosity	Context	Pre-TRP cues Autonomous agents	Facial, gestures, posture, Gaze, Feedback
[36]	>3	>4 Agents	Speaker Addresse Side Participant Overhearer	Take Request Release Hold Assign	Discourse, Context	Multiple floors Make or break contact	Gaze, Gestures, Proxemics
[16]	6	Agents	Speaker Addresse Side Participant Overhearer	Hold Release Take	Context	Multiple floors Make or break contact Pre-TRP cues	Nods, Gestures, Posture Shifts, Gaze Feedback (Pos/Neg)
[17]	6	Agents	Speaker Addresse Side Participant Overhearer	Close to Speaker Away from Noise Proxemity to Agents Circular Formation	Context	Multiple floors Dynamic groups	Gestures, Postures, Gaze, Proxemics
[6]	>3	1 Agent 2-3 participants	Speaker Addresse Side Participant Overhearer Eavesdropper	Hold Release Take Null	Discourse, Context	Multiple floors Determines source of voice Prompts user to take turn	Gaze, Gestures, Facial
[27]	>3	>4 Agents	Speaker Addresse Side participant	Own-the-speech Compete-for-the-speech End-of-speech Interrupted Addressed-listener Unaddressed-listener Want-to-speak	Context	Interruptions Interpersonal attitudes	Gaze, Gestures, Proxemics

the turn after a TRP at which no one self-selected. The agents made probabilistic decision about the display of behaviour such as speaking and listening, feedback, head nods, gestures, posture shifts, and gaze. The algorithm generated a log of the decisions made by the agents and did not provide any visual representation.

In real life, groups are not always static and they may fragment into subgroups or merge into a larger group e. g., an agent can join an already existing conversation or two agents that are already a part of the conversation can start a new conversation of their own and form a different group. In [16] an algorithm that simulates behaviours and allows dynamic changes to conversational group structure is presented. The agents take turns to talk and make use of non-verbal cues to indicate end of turn. When an agent is speaking, the gaze behaviour of other agents is monitored and appropriate feedback is provided. The algorithm supports sharing of messages between the agents to communicate, but these messages do not result in any direct outcome. The decisions are made independently by an individual agent i. e., when a character receives a message it can react immediately or update the internal state and make decision at a scheduled time. One of the main limitations of this model is the lack of dynamic movement and positioning which was addressed in [17].

Appropriate positioning of the agents in groups with respect to other agents and the environment is important since it makes the simulation more engaging and realistic. Incorrect positioning of the agents can have a negative effect on the believability. [17] presents an algorithm for simulating the dynamic movements, orientation and positioning observable in multi-party group conversations and is a continuation of the work presented in [16]. The movement and positioning component allows agents to monitor "forces" that make it more desirable to move to one place or another while remaining engaged in conversation. The desire to move and the direction can be represented as a vectorial quantity which could be interpreted as social force. This force does not directly cause any movement but provides a motivation to move. Several forces can be active at any given point. The reasons for the agents to move can be (1) to come closer to the speaker, (2) avoid/ move away from background noise, (3) if an agent is too close than the comfortable distance, (4) if the speaker is occluded (move to make a circular formation).

In [35], Thorisson et. al, proposed the Ymir Turn Taking Model (YTTM) which is a computational agent-oriented model. The model has been implemented with up to 12 agents in a virtual world participating in real-time cooperative dialogue. Eight dialogue contexts each containing the various perception and action modules has been implemented. Each agent has an isolated context model and input from perceptions that include a list of all conversation participants, speaker, who is "looking at me" (for any given agent) and who is requesting turn at each given time. They also have configuration for urge-to-speak, yield tolerance, impatience for getting the turn. Although it supports multi-party conversation it does not consider the expression of attitudes and dynamic group formation. In [27], Ravenet et. al, presented a computational model for affective real-time turn-taking that allowed an agent to express interpersonal attitudes in a group. The multi-agent model was developed where multiple virtual agents could converse and display non-verbal behaviour, including turn-taking. It consists of a turn-taking component, a group behaviour component and a conversational behaviour component. The turns are modeled as a state machine and activate based on specific input values. Although, this model supports multi-party conversation and expression of attitudes, it does not consider the content of speech or different interruption strategies. These kind of simulations allows middle-level of detail for crowds, in which the characters are close enough to be visible, but are not the main characters in the setting. Their main role is not to interact with the user, but rather maintain the illusion of a realistic environment where the user is situated.

Ada and Grace are virtual museum guides with differing opinions and behaviours that are designed to provide useful information [34]. The two agents interact with the user and answer the questions when the user presses a push-to-talk button and speaks into a microphone. The interaction between the agents is limited to sharing the responses and do not interact with each other when no one is talking with them. Gunslinger is an interactive-entertainment application where a single participant can interact verbally and nonverbally with multiple virtual characters [14]. It includes speech recognition and visual sensing to recognize the multimodal user input. A statistical text classification algorithm selects the character's responses based on the user's utterance, a visual recognition module detects and tracks the motion of the user in the physical space over time and the visual understanding module infers higher level behaviours e. g., facing a character. In Mission Rehearsal Exercise (MRE) [36], Traum et. al, proposed a model for multi-party interactions in a 3D virtual environment. This model supports multi-floor dialogue. Agents can interact with a human user and as well as with each other and respond. The model consists of several layers i. e., Contact, Attention, Conversation (Participants, Turn, Initiative, Grounding, Topic, Rhetorical), Obligations, Negotiation. Each layer has an information state, dialogue acts and required signals and recognitions rules. The agents can make contact by going close (eye or ear shot) and break contact by walking away. In [6], a computational model is presented for multi-party turn-taking. It tracks the conversational dynamics to make turn decisions and for rendering decisions about turns into an appropriate set of low-level behaviours like, coordinated gaze, gesture and speech. In [2], Barange et. al, propose a multi-party coordination model that allows several virtual and human agents to dialogue and reason to perform a collaborative task.

3 RESEARCH APPROACH

As we have seen, there exists several embodied conversational agents that aid users and provide coaching in a specific aspect e. g., diabetes coach, meditation coach. Developing healthcare virtual agents to provide adequate support and advice on their health, habits and lifestyle could improve the overall quality of life among the elderly. To the best of our knowledge, this will be the first work where multiple embodied conversational coaches with specific domain expertise will be developed. Virtual humans need to be persuasive in order to promote behaviour change in human users. Design features of the virtual agent like role, verbal, and non-verbal communication is critical. The attitude that each agent aims at expressing towards the others determines the verbal and nonverbal behavior that are exhibited in such group interactions [27].

While several studies have focused on understanding the numerous aspects that influence the degree of persuasion, most of them are limited to dyadic interactions. Since we aim to have multiple virtual agents in the same environment, it is therefore important to understand the preferences of the users i. e., features related to the appearance style, group composition e. g., single or multiple agents in different roles. This will facilitate in developing coaches that will be able to handle conversations successfully and be effective in user behaviour changes. The attitudes expressed by the coaches will have a direct influence on the conversational structure i. e., an agent with dominant personality will have the tendency to interrupt or claim turns more frequently [7]. The first step in this research work will be to understand the user preferences of the agent in a multiparty setting. A user-based study is in progress which aims to understand the effects of gender and role on user's persuasion. Since this is the preliminary experiment, we make use of a non-interactive presentation format. Future studies will involve an interactive system which will include a behaviour profile of the users and aim to provide personalized assistance and appropriate advice related to health.

The current experiment is based on 2 x 2 x 3 design, where the variables include agent gender (male vs. female), role (authoritative vs. peer) and persuasion type (multiple agent user-based vs. multiple agent vicarious vs. single agent). For the initial step, we have chosen the topic of discussion on movies, which is one of the most common topics of discussion among the general population. The participants are randomly assigned to one of the 12 conditions specified above and will be asked to provide rating for three movies based on textual description of the movies. In the next step, the user will be shown a 60-90s clip where one/several virtual agents will present a persuasive dialogue about the movie which received the lowest rating and ask the participant to provide a rating again. The results from this study will be considered while designing the appearance of virtual agents and their personalities for the Council of Coaches project.

Further we will develop a theoretical model to handle multiparty turn taking. Existing models that handle turn-taking in multi-party conversations display the required non-verbal behaviour but they do not consider the verbal content and the dynamic nature of group conversations completely. Our aim is to:

- Extend the model proposed in [27] to consider the verbal content.
- Handle multi-topic management.
- Be able to switch floors when required and handle interruptions by user or other agents.
- Enable the model to dynamically adjust the turn-taking strategy when a participant joins/leaves the conversation.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant Agreement Number 769553. This result only reflects the author's view and the EU is not responsible for any use that may be made of the information it contains.

REFERENCES

[1] K. Anderson, E. André, T. Baur, S. Bernardini, M. Chollet, E. Chryssafidou, I. Damian, C. Ennis, A. Egges, P. Gebhard, et al. 2013. The TARDIS framework: Intelligent Virtual Agents for social coaching in job interviews. In *Advances in computer entertainment*. Springer, 476–491.

[2] M. Barange, J. Saunier, and A. Pauchet. 2017. Multiparty Interactions for Coordination in a Mixed Human-Agent Teamwork. In *Proc. International Conference on Intelligent Virtual Agents*. Springer, 29–42.

[3] T Bickmore, L Caruso, K Clough-Gorr, and T Heeren. 2005. 'It's just like you talk to a friend' relational agents for older adults. *Interacting with Computers* 17, 6 (2005), 711–735.

[4] Timothy Bickmore and Laura Pfeifer. 2008. Relational agents for antipsychotic medication adherence. In *CHI'08 workshop on Technology in Mental Health*.

[5] T W Bickmore, L M Pfeifer, and M K Paasche-Orlow. 2007. Health document explanation by virtual agents. In *International Workshop on Intelligent Virtual Agents*. Springer, 183–196.

[6] D. Bohus and E. Horvitz. 2010. *Computational Models for Multiparty Turn Taking*. Technical Report. Microsoft Research Technical Report MSR-TR 2010-115.

[7] A Cafaro, N Glas, and C Pelachaud. 2016. The effects of interrupting behavior on interpersonal attitude and engagement in dyadic interactions. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 911–920.

[8] J. Cassell, T. Bickmore, L. Campbell, H. Vilhjálmsson, and H. Yan. 2000. Conversation as a system framework: Designing embodied conversational agents. *Embodied conversational agents* (2000), 29–63.

[9] D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgila, J. Gratch, A. Hartholt, M. Lhommet, et al. 2014. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In *Proc. of the 2014 International conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1061–1068.

[10] D. DeVault, J. Mell, and J. Gratch. 2015. Toward natural turn-taking in a virtual human negotiation agent. In *Proc. AAAI Spring Symposium on Turn-taking and Coordination in Human-Machine Interaction. AAAI Press, Stanford, CA*.

[11] B. Farber. 2006. *Self-disclosure in psychotherapy*. Guilford Press.

[12] E G. Padilha and J Carletta. 2002. A Simulation of Small Group Discussion. (09 2002).

[13] J. Gratch, S. Kang, and N. Wang. 2013. Using social agents to explore theories of rapport and emotional resonance. *Social Emotions in Nature and Artifact* 181 (2013).

[14] A. Hartholt, J. Gratch, L. Weiss, et al. 2009. At the virtual frontier: Introducing Gunslinger, a multi-character, mixed-reality, story-driven experience. In *Proc. International Workshop on Intelligent Virtual Agents*. Springer, 500–501.

[15] T. Hiraoka, K. Georgila, E. Nouri, D. Traum, and S. Nakamura. 2016. Reinforcement Learning of Multi-Party Trading Dialog Policies. *Information and Media Technologies* 11 (2016), 264–277.

[16] D Jan and D R Traum. 2005. Dialog simulation for background characters. In *International Workshop on Intelligent Virtual Agents*. Springer, 65–74.

[17] D Jan and D R Traum. 2007. Dynamic movement and positioning of embodied agents in multiparty conversations. In *Proceedings of the Workshop on Embodied Language Processing*. Association for Computational Linguistics, 59–66.

[18] J Lawrence, M Snaith, B Konat, K Budzynska, and C Reed. 2017. Debating Technology for Dialogical Argument: Sensemaking, Engagement, and Analytics. *ACM Transactions on Internet Technology (TOIT)* 17, 3 (2017), 24.

[19] M. Leary and R. Kowalski. 1990. Impression management: A literature review and two-component model. *Psychological bulletin* 107, 1 (1990), 34.

[20] N. Leßmann, A. Kranstedt, and I. Wachsmuth. 2004. Towards a cognitively motivated processing of turn-taking signals for the embodied conversational agent Max. In *Proc. Workshop Embodied Conversational Agents: Balanced Perception and Action*.

[21] C Lisetti, R Amini, U Yasavur, and N Rishe. 2013. I can help you change! an empathic virtual agent delivers behavior change health interventions. *ACM Transactions on Management Information Systems (TMIS)* 4, 4 (2013), 19.

[22] G. Lucas, J. Gratch, A. King, and L. Morency. 2014. It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior* 37 (2014), 94–100.

[23] H Monkaresi, R Calvo, A Pardo, K Chow, B Mullan, M Lam, SM Twigg, and DI Cook. 2013. Intelligent diabetes lifestyle coach. In *OzCHI workshops programme*.

[24] Council of Coaches. 2017. The Project. (2017). http://council-of-coaches.eu

[25] F. Pecune, A. Cafaro, M. Chollet, P. Philippe, and C. Pelachaud. 2014. Suggestions for extending SAIBA with the VIB Platform. In *Proc. of the Workshop on Architectures and Standards for Intelligent Virtual Agents at IVA*. 16–20.

[26] S Provoost, H Lau, J Ruwaard, and H Riper. 2017. Embodied conversational agents in clinical psychology: A scoping review. *Journal of medical Internet research* 19, 5 (2017).

[27] B. Ravenet, A. Cafaro, B. Biancardi, M. Ochs, and C. Pelachaud. 2015. Conversational Behavior Reflecting Interpersonal Attitudes in Small Group Interactions. In *Proc. International Conference on Intelligent Virtual Agents*. Springer, 375–388.

[28] N Reithinger, P Gebhard, M Löckelt, A Ndiaye, N Pfleger, and M Klesen. 2006. VirtualHuman: dialogic and affective interaction with virtual characters. In *Proc. of the 8th international conference on Multimodal interfaces*. ACM, 51–58.

[29] T Rist, E André, S Baldes, P Gebhard, M Klesen, M Kipp, P Rist, and M Schmitt. 2004. A review of the development of embodied presentation agents and their application fields. In *Life-Like Characters*. Springer, 377–404.

[30] A Shamekhi and T Bickmore. 2015. Breathe with me: a virtual meditation coach. In *International Conference on Intelligent Virtual Agents*. Springer, 279–282.

[31] A Shamekhi, T Bickmore, A Lestoquoy, L Negash, and P Gardiner. 2016. Blissful agents: adjuncts to group medical visits for chronic pain and depression. In *International Conference on Intelligent Virtual Agents*. Springer, 433–437.

[32] M Smith, E Ginger, K Wright, M Wright, J Taylor, L Humm, D Olsen, M Bell, and Michael F Fleming. 2014. Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 44, 10 (2014), 2450–2463.

[33] G. Stratou, L. Morency, D. DeVault, A. Hartholt, E. Fast, M. Lhommet, G. Lucas, F. Morbini, K. Georgila, S. Scherer, et al. 2015. A demonstration of the perception system in SimSensei, a virtual human application for healthcare interviews. In *Proc. International Conference on Affective Computing and Intelligent Interaction (ACII), 2015*. IEEE, 787–789.

[34] W. Swartout, D. Traum, R. Artstein, D. Noren, P. Debevec, K. Bronnenkant, J. Williams, A. Leuski, S. Narayanan, D. Piepol, et al. 2010. Ada and Grace: Toward realistic and engaging virtual museum guides. In *Proc. International Conference on Intelligent Virtual Agents*. Springer, 286–300.

[35] K. Thórisson, O. Gislason, G. Jonsdottir, and H. Thorisson. 2010. A Multiparty Multimodal Architecture for Realtime Turntaking. In *Proc. Intelligent Virtual Agents*. Springer, 350–356.

[36] D. Traum and J. Rickel. 2002. Embodied Agents for Multi-party Dialogue in Immersive Virtual Worlds. In *Proc. of the First International Joint Conference on Autonomous agents and Multi-agent systems*. ACM, 766–773.

[37] L Vardoulakis, L Ring, B Barry, C Sidner, and T Bickmore. 2012. Designing relational agents as long term social companions for older adults. In *International Conference on Intelligent Virtual Agents*. Springer, 289–302.

[38] L Wanner, E André, J Blat, S Dasiopoulou, M Farrùs, T Fraga, E Kamateri, F Lingenfelser, G Llorach, O Martínez, et al. 2017. Kristina: A knowledge-based virtual conversation agent. In *Proc. International Conference on Practical Applications of Agents and Multi-Agent Systems*. Springer, 284–295.

[39] Q. Xu, L. Li, and G. Wang. 2013. Designing engagement-aware agents for multiparty conversations. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2233–2242.

[40] Z. Yumak, J. Ren, N. Thalmann, and J. Yuan. 2014. Tracking and fusion for multiparty interaction with a virtual character and a social robot. In *Proc. SIGGRAPH Asia 2014 Autonomous Virtual Humans and Social Robot for Telepresence*. ACM, 3.