Demo: Have a Chat with Sensitive Artificial Listeners

Marc Schröder, Sathish Pammi,[1] **Hatice Gunes, Maja Pantic, Michel Valstar,**[2]
Dirk Heylen, Mark ter Maat,[3] **Florian Eyben, Björn Schuller, Martin Wöllmer,**[4]
Elisabetta Bevacqua, Catherine Pelachaud, Etienne de Sevin[5]

Abstract. Sensitive Artificial Listeners (SAL) are virtual dialogue partners based on audiovisual analysis and synthesis. Despite their very limited verbal understanding, they intend to engage the user in a conversation by paying attention to the user's emotions and non-verbal expressions. The SAL characters have their own emotionally defined personality, and attempt to drag the user towards their dominant emotion through a combination of verbal and non-verbal expression. The demonstrator shows the publicly available, fully autonomous SAL system.

1 INTRODUCTION

The SAL demo supports sustained emotionally-coloured machine-human interaction using non-verbal expression. The system aims to engage the user in a dialogue by paying attention to the user's non-verbal expression, and reacting accordingly. It focuses on the "soft skills" that humans naturally use to keep a conversation alive.

To simplify the challenge somewhat, the SAL system avoids task-oriented dialogue. Instead, it models the type of interaction found at parties: you listen to someone you want to chat with, and without really understanding much of what they are saying, you exhibit all the signs that are needed for them to continue talking to you. Similarly to the Rapport agent [2], SAL characters show non-verbal listener signals; in addition, they can also speak to engage the user in a simple dialogue. The approach has been test-run using Wizard of Oz setups at various stages of maturity [1]. This has allowed us to fine-tune the scripts used by the various characters, in order to react to the emotional state of the user in plausible ways despite the lack of verbal knowledge. The system is freely available from the project website `www.semaine-project.eu`.

2 THE DEMONSTRATION SETUP

In the SEMAINE demonstrator, one human user is sitting in front of a computer screen showing the face of an Embodied Conversational Agent (ECA). The user is wearing a headset for voice analysis and is recorded by a video camera for facial expression analysis. The ECA is speaking through loudspeakers, and is showing both verbal and non-verbal behaviour. A second computer screen shows a system monitor, displaying graphically the current information flow in the system.

Figure 1. The four SAL characters represent the four quadrants of arousal-valence space: Spike is aggressive; Poppy is cheerful; Obadiah is gloomy; and Prudence is pragmatic.

The user can speak to one of the four SAL characters at a time (see Figure 1). Each character will try sustain the conversation by non-verbal feedback and verbal utterances. After a while, the user or the agent can switch to a different character. In the lab, sessions typically last for around 20 minutes; in the demo session, much shorter sessions with changing users are anticipated.

Technically, the demonstrator system is a multimodal interactive system with components integrated across programming languages and operating systems by means of a standards-based framework for building emotion-oriented systems, the SEMAINE API [3]. Details on the technological setup and the individual processing components are described in a set of project deliverable reports available from the project website.

ACKNOWLEDGEMENTS

REFERENCES

[1] E. Douglas-Cowie, R. Cowie, C. Cox, N. Amir, and D. Heylen, 'The sensitive artificial listener: an induction technique for generating emotionally coloured conversation', in *LREC2008 Workshop on Corpora for Research on Emotion and Affect*, pp. 1–4, Marrakech, Morocco, (2008).

[2] J. Gratch, N. Wang, J. Gerten, E. Fast, and R. Duffy, 'Creating rapport with virtual agents', in *Intelligent Virtual Agents*, 125–138, (2007). doi:10.1007/978-3-540-74997-4_12.

[3] M. Schröder, 'The SEMAINE API: towards a standards-based framework for building emotion-oriented systems', *Advances in Human-Computer Interaction*, **2009**(319406), (2009). doi:10.1155/2009/319406.

[1] DFKI GmbH, Saarbrücken, Germany, email: firstname.lastname@dfki.de

[2] Imperial College London, UK, email: firstname.lastname@imperial.ac.uk

[3] Universiteit Twente, The Netherlands, email: (d.k.j.heylen |maatm)@ewi.utwente.nl

[4] Technische Universität München, Germany, email: lastname@tum.de

[5] CNRS, Telecom Paristech, France, email: firstname.lastname@telecom-paristech.fr