A listening agent exhibiting personality traits

Elisabetta Bevacqua, Maurizio Mancini, and Catherine Pelachaud

University Paris 8, 140 rue de la Nouvelle France 93100, Montreuil, France
INRIA Rocquencourt, Mirages BP 105, 78153 Le Chesnay Cedex, France

Abstract. Within the Sensitive Artificial Listening Agent project, we propose a system that computes the behaviour of a listening agent by encompassing the notion of personality. In our system agent's behaviour tendencies are defined by a set of parameters. The system selects the signals to be displayed by the agent depending on its behaviour tendency and its communicative intentions. In this paper we focus on the listener's behaviour.

1 Introduction

The work presented in this paper is set within the Sensitive Artificial Listening Agent (SAL) project. It is part of the EU STREP SEMAINE project (http://www.semaine-project.eu). This project aims to build an autonomous talking head able to exhibit appropriate behaviour when it plays the role of the listener in a conversation with a user. Four characters, with different personalities, invite the user to chat trying to induce her/him in a particular mood. Within SAL, we aim to build a real-time Embodied Conversational Agent (ECA) endowed with recognizable personality traits while interacting with users in the role of the listener. We propose a model that provides a static definition of such an ECA on the base of: (i) the preference the agent has in using each available communicative modality (e.g., head orientation, eyebrows movements, voice and so on) and (ii) a set of parameters that affect the qualities of the agent's behaviour (e.g. wide vs. narrow gestures). We call such a definition of the agent the *baseline* and we fix it depending on the agent's personality traits. The proposed work incorporates a pre-existing system for the generation of distinctive behaviour in ECAs [MP07,MP08]. The result is a system capable of determining the agent's communicative intention in real-time, according to the speaker's behaviour, and computing the verbal and non-verbal behaviours that the agent has to perform on the basis of both its *baseline* and its communicative intentions. For example, an extrovert agent can show its agreement through wide and fast head nods and an open smile on its face.

2 Background

2.1 Personality

A big challenge that must be faced in the design of virtual agent is the issue of credibility, not only in the agent's aspect but also in its behaviour. Users tend to react as if in a real human-human interaction when the virtual agent behaves in a natural human-way manner. Researches have sown that agents that exhibit personality traits are more believable. In particular, Nass et al. showed that people react to agents endowed with personality characteristics in the same manner they would react to humans with similar personalities [CN94].

2.2 Behaviour tendencies

Argyle [Arg88] states that there is an underlying tendency which is constantly present in each person's behaviour: for example people that look more tend to do so in most communicative situations, that is, there is a certain amount of consistency with the person's general tendency. Also Gallaher [Gal92] found consistencies in the way people behave: she conducted evaluation studies in which subjects' behaviour style was evaluated by friends, and by self-evaluation. Results demonstrated that for example people who are fast when writing have a tendency to be fast while eating; if a person produces wide gestures then she also walks with large steps.

2.3 Listener's behaviour

To assure a successful communication, listeners must provide responses about both the content of the speaker's speech and the communication itself. Through verbal and non verbal signals, called *backchannel signals*, a listener provides information about the basic communicative functions, as perception, attention, interest, understanding, attitude (e.g., belief, liking and so on) and acceptance towards what the speaker is saying [ANA93,Pog05]. For instance, the interlocutor can show that s/he is paying attention but not understanding and, according to the listener's responses, the speaker can decide how to carry on the interaction: for example by re-formulating a sentence.

3 Related work

Past researches on ECAs have provided first approaches to the implementation of a backchannel model. K. R. Thórisson developed a talking head, called Gandalf, able to produce real-time backchannel signals during a conversation with an user [Thó96], while Cassell et al. [CB99] developed the Real Estate Agent (REA) which is able to understand the user requests in real-time. The Listening Agent [MGM05], developed at ICT, produces backchannel signals based on real-time analysis of the speaker's non verbal behaviour (as head motion and body posture) and of acoustic features extracted from the speaker's voice [WT00]. Kopp et al. [KSG07] proposed a model for generating incremental backchannel. The system is based both on a probabilistic model, that defines a set of rules to determine the occurrence of a backchannel, and on a simulation model that perceives, understands and evaluates input through multi-layered processes. All the models described above do not take into account the agent's personality;

backchannel signals are emitted according to the speaker's behaviour but they are not linked neither to the agent's communicative intentions nor to its personality traits. Moreover these systems do not consider agents with different behaviour tendencies.

4 Agent Definition

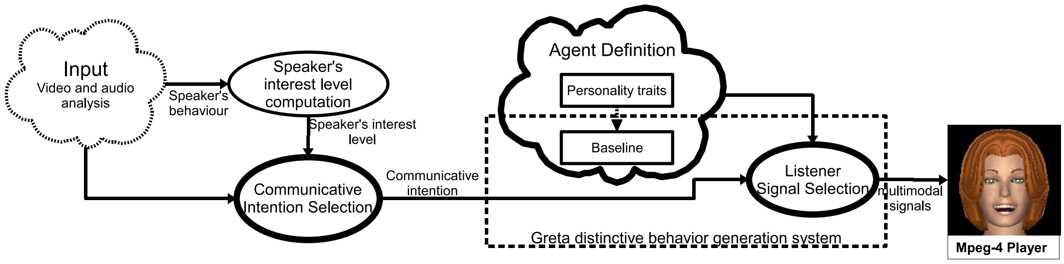

Fig. 1. System diagram: we highlight the three parts we focus on in this paper: the Agent Definition, the Communicative Intention Selection and the Listener Signal Selection. We base our work on the distinctive behaviour generation system of the Greta agent [MP07,MP08].

As shown in Figure 1, we define an agent by fixing its personality traits, from which we determine its behaviour tendencies. We introduce these two concepts in the following Sections.

4.1 Personality traits

To provide our agent with personality traits we will use an existing model like the Big Five model or the model proposed by Wiggins. As a first approach, we aim to simulate the personality of the four characters proposed in SAL: Poppy (who is outgoing), Obadiah (who is depressing), Spike (who is argumentative) and Prudence (who is pragmatic). For example, based on the personality parameters on the Big Five model, Spike could be defined by setting the dimensions extraversion and neuroticism to high values, agreeableness and conscientiousness to low values and openess to a medium value. We use the personality parameters to determine the agent's *baseline*, defined in Section 4.2. In fact, researches have shown there exists a link between personality and behaviour. Wickett and Vernonb found that ambiverts move slower than introverts and that extraverts are the fastest [WV00].

4.2 Baseline: modality preference and behaviour expressivity

We define the agent's *baseline* as a set of numeric parameters that represents the agent's behaviour tendencies. In the baseline we represent two kinds of data: the agent's modality preference and the agent's behaviour expressivity.

People can communicate by being more or less expressive in the different modalities: a person can move a lot her arms while another one can produce many facial signals and so on. The reasons behind such personal differences can be due for example to differences in personality, emotional state, mood, sex, age, nationality [WS86]. In our static definition of an agent, we implemented the *modality preference* to represent the agent's degree of preference in using each available modality. If, for example, we want to specify that the agent has the tendency to mainly use hand gestures during communication, we assign a high degree of preference to the gesture modality, if it uses mainly the face, the face modality is set to a higher value, and so on.

To define the behaviour tendencies of an agent we also defined and implemented a set of parameters that allow one to alter the way the agent expresses its actual communicative intention [HMBP05]. The agent's *behaviour expressivity* is defined by a set of 6 parameters that influence the quality of the agent's movements: the frequency (OAC parameter), speed (TMP parameter), spatial volume (SPC parameter), energy (POW parameter), fluidity (FLD parameter), and repetitivity (REP parameter) of the nonverbal signals produced by the agent.

5 Communicative Intention Selection

To display a believable listener behaviour, a virtual agent must be able to: (i) decide *when* a backchannel signal should be emitted and (ii) select *which* communicative intentions the agent should transmit through the signal. In our system these tasks are performed by the *Communicative Intention Selection* module of Figure 1. Past researches have shown that backchannel signals are often emitted according to the verbal and non verbal behaviour performed by the speaker [WT00,MGM05]. On the base of these results, our system evaluates video and audio data to select user's behaviours that could elicit a backchannel from the agent. Once a speaker candidate behaviour is found, the system computes the probability that the listener has the intention to communicate something to the speaker. Such a probability depends on the speaker's level of interest, an emotional state linked to the speakers goal of obtaining new knowledge [Pet05]. In the system the interest level is a parameters that is calculated evaluating her/his gaze, head and torso direction within a temporal window. As a consequence the agent displays less and less backchannels if the speaker wants to stop the conversation.

When the probability of communicate something to the speaker is higher than a threshold, our system selects the communicative intentions that the agent wants to transmit. We define the listening agent's communicative intentions by a set of 12 parameters, called *listener intentions.* Such parameters correspond to

12 communicative intentions that we chose from the literature: agree, disagree, accept, refuse, believe, disbelieve, interest, not interest, like, dislike, understand, not understand. Each parameter can have a value between 0 and 1 that represents the normalized percentage of the importance that the agent gives to the transmission of a certain communicative intention. Therefore their sum must be 1. For example, when a backchannel signal must be emitted and the values of the parameters *agree* and *understand* are respectively 0.8 and 0.2, the agent decides to transmit the communicative intentions of agreement and understanding, and that the transmission of the intension *agree* is far more important. To vary the *listener intentions* parameters a virtual agent should be provided at least with a cognitive module that interprets the speaker's words and compares the agent's believes with the speech content. In fact, in human-human conversations the interlocutor's communicative intentions depend on several factors like the content of the speaker's speech, her/his own beliefs, the relationship between the two parties and so on. Similar modules are still very hard to implement and usually limited to specific domain. Such a module is under development within the SAL project.

6 Listener Signal Selection

In this Section we describe the process of performing the selection of the nonverbal behaviours that the listener has to produce in order to convey its communicative intention. This task is performed by the *Listener Signal Selection* (*LSS*) module of Figure 1, which is an extended version of the corresponding one we implemented for the distinctive behaviour generation system of the Greta agent, presented in [MP07,MP08].

6.1 Behaviour sets

In the LSS module, all the possible listener's communicative intentions are associated with the multimodal signals that can be produced by the listener in order to convey them. Each of these associations represents one *entry* of a lexicon, called *behaviour set*. A behaviour set is defined by the following parameters:

- the *name* of the corresponding communicative intention. For example, to define the behaviour set corresponding to the communicative intention *refuse*, we set the *name* of the set to *refuse*.
- the set S, containing the name of the signals produced on single modalities that can be used to convey the intention specified by the parameter *name*. For example the intention *refuse* can be conveyed by: shaking the head, saying "*no!*" and so on. This parameter does not precise *how* and *if* these signals can be combined. The next two parameters will specify this information.
- the list of signals that are mandatory to communicate the intention corresponding to the behaviour set; for example to communicate *refuse*, the listener *MUST* shake its head.

– a set of logic rules like *if A then B* where A is a condition involving both the parameters of the agent definition (see Section 4) and the signals contained in the set S and B is a subset of S. Let us consider an example of one of these rules:

$$\texttt{if}(\texttt{head.OAC.value} > \texttt{threshold1})\texttt{producesignal}(\texttt{s1}); \quad (1)$$

With rule 1 we specify that if the value of the head Overall Activation parameter (OAC) is higher than a given threshold then the listener has to produce the signal $s1$. In this case we are referencing to a value in the agent's baseline.

6.2 Performing the Listener Signal Selection

As shown in Figure 1, the Listener Signal Selection process takes as input the agent definition (that is, its personality traits and its baseline) and the agent communicative intention and computes the multimodal behaviour that the agent has to perform. The process consists of some steps of computation: first the system looks for the behaviour set corresponding to the agent's communicative intention and computes all the possibile combinations of the signals contained in the set S (see Section 5); it discards the combinations that do not contain the mandatory signals and checks the logic rules contained in the behaviour set, discarding the signal combinations that do not verify these rules; finally it prioritizes the signal combinations depending on the agent modality preference (see Section 4.2) and chooses the signal combination with the highest priority.

7 Conclusion and future work

We propose to develop a system that computes the behaviour of a listening agent by encompassing the notion of personality. This system is part of the SAL project aiming to implement an agent exhibiting realistic behaviour when playing the role of a listener during a conversation. In our system the agent's personality traits are used to manually determine the agent's behaviour tendencies, that is, the preference the agent has in using each modality and also the expressivity of behaviour. Then we propose to implement the real-time computation the agent's communicative intention. Finally, we provide the information about the agent's definition and communicative intention as input to a pre-existing module which computes the agent's multimodal behaviour. The output of such a system is the listener multimodal behavior which depends on the listener's personality, because: (i) we consider the listener's personality traits in manually defining its behavior tendencies and (ii) we include the listener's personality traits in the process going from the listener's communicative intention to the the generation of its multimodal behavior.

In the future we aim to automatically determine the agent's behaviour tendencies starting from its personality traits. Moreover we will add the cognitive module that is under development in SAL to interpret the speaker's speech and to automatically compute the *listener intentions* parameters.

8 Acknowledgement

This work has been funded by the STREP SEMAINE project IST-211486 (http://www.semaine-project.eu) and the IP-CALLAS project IST-034800 (http://www.callas-newmedia.eu).

References

[ANA93] J. Allwood, J. Nivre, and E. Ahlsn. On the semantics and pragmatics of linguistic feedback. *Semantics*, 9(1), 1993.

[Arg88] M. Argyle. *Bodily Communication.* Methuen & Co., London, 2nd edition, 1988.

[CB99] J. Cassell and T. Bickmore. Embodiment in conversational interfaces: Rea. In *Conference on Human Factors in Computing Systems*, Pittsburgh, PA, 1999.

[CN94] E. R. Tauber C. Nass, J. Steuer. Computers are social actors. In *CHI*, pages 72–78. 1994.

[Gal92] P. E. Gallaher. Individual differences in nonverbal behavior: Dimensions of style. *Journal of Personality and Social Psychology*, 63(1):133–145, 1992.

[HMBP05] B. Hartmann, M. Mancini, S. Buisine, and C. Pelachaud. Design and evaluation of expressive gesture synthesis for embodied conversational agents. In *3th International Joint Conference on Autonomous Agents & Multi-Agent Systems*, Utretch, 2005.

[KSG07] S. Kopp, T. Stocksmeier, and D. Gibbon. Incremental multimodal feedback for conversational agents. In *IVA*, pages 139–146, 2007.

[MGM05] R. M. Maatman, J. Gratch, and S. Marsella. Natural behavior of a listening agent. In *5th International Conference on Interactive Virtual Agents*. Kos, Greece, 2005.

[MP07] M. Mancini and C. Pelachaud. Dynamic behavior qualifiers for conversational agents. In *Intelligent Virtual Agents*, pages 112–124, 2007.

[MP08] M. Mancini and C. Pelachaud. Distinctiveness in multimodal behaviors. In *Conference on Autonomous Agents and Multiagent System*, 2008.

[Pet05] C. Peters. Direction of attention perception for conversation initiation in virtual environments. In *International Working Conference on Intelligent Virtual Agents*, pages 215–228, Kos, Greece, September 2005.

[Pog05] I. Poggi. Backchannel: from humans to embodied agents. In *AISB*. University of Hertfordshire, Hatfield, UK, 2005.

[Thó96] K. R. Thórisson. *Communicative Humanoids: A Computational Model of Psychosocial Dialogue Skills.* PhD thesis, MIT Media Laboratory, 1996.

[WS86] H. G. Wallbott and K. R. Scherer. Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, 51(4):690–699, 1986.

[WT00] N. Ward and W. Tsukahara. Prosodic features which cue back-channel responses in english and japanese. *Journal of Pragmatics*, 23:1177–1207, 2000.

[WV00] J. C. Wickett and P. A. Vernonb. Replicating the movement timeextraversion linkwith a little help from iq*1. *Journal Personality and Individual Differences*, 28:205–215, 2000.