Abstract

We aim at creating an Embodied Conversational Agent (ECA) that would exhibit not only a consistent behavior with her personality and contextual environment factors but also that would be defined as an individual and not as a generic agent. The behavior of an agent depends not only on factors defining her individuality (such as her culture, her social and professional role, her personality, her experience) but also on a large set of contextual and dynamic influences.

1. Introduction

We aim at creating an Embodied Conversational Agent (ECA) that would exhibit specific behaviors, by the choice of the signals to be displayed as well as by the expressivity the signal would be displayed. The determination and the expressivity of the agent’s behavior depends not only on factors defining her individuality (such as her culture, her personality, her social and professional role, her experience), on a large set of contextual influences (e.g., her interlocutor, the social conversation setting) and also on dynamic influences (such as her emotional states).

Most of the agents that have been created so far are very generic in their behavior type. Several studies have shown the importance to consider complex information [3] such as cultural factors, personality, environment setting when designing an agent. Personality, gender, age have effect on the way people interact with others, talk about things... These differences arise at different levels: the formulation of our thoughts as well as their expression [2]. Mapping such a reasoning to the agent world, we can say that all the factors that should be considered to create an individual agent act at different levels of the creation. They have influences on the surface generation and realization level during the dialog generation phase as well as on the selection of the non-verbal behaviors and of their expressivity [2].

Several systems have been developed aiming at creating agents whose behaviors may be modulated by different factors: culture [2], emotion, social relationship [4, 5], personality and so on. The other work that is related to our is by Ruttkay and Noot [6]. The authors developed a very complex representation language based on several dictionaries. In this language the authors embed notions such as culture, personality, gender but also physical information such as gesturing manner or tiredness. The authors modelled explicitly how factors such as culture and personality affect behaviors. We distinguish our work from them in the sense that we do not modelled such factors, rather we modelled the different types of influences that may occur and how they may modulate an agent’s behaviors.

In this paper we do not aim at modelling what culture or personality mean, nor do we aim at simulating dynamically expressive animations. In this paper we limit our scope at representing influences that would modify the behavior an agent will display to communicate a given meaning within a specific context and how it will be shown.

In section 2 we present a taxonomy of influences while in section 3 we turn our attention toward the modelling of expressivity. Section 4 describes our algorithm. Finally we conclude by describing our future work.

2. Taxonomy of Influences

At first, we consider that each human being has a set of the conscious and unconscious habits that intervene in the content of her discourse and that define her behavior. These habits, we call the intrinsic factors, derive, among others, from her personality, her age, her sex, her nationality, her culture, her education and her experiences. Several factors, depending on the context, may increase or decrease the effects of the intrinsic factors. The contextual factors may refer to the environment setting such as the light conditions, the sound intensity, the spatial layout or the function of the conversation site. These factor may affect some speech and behavior characteristics. Her relationships with...
her interlocutor modulates also her behavior: she does not behave in the same way with a friend, an unknown person, an employee, a child or a doctor [1]. These contextual influences, unlike the intrinsic ones, may vary during a dialog session. Dynamic factors embed variations due to the agent’s mental and emotional states. They also affect greatly the way the agent will behave. It modifies the prosody of speech, the amplitude of a facial expression, the movement tempo. For instance, a person does not talk and does not behave in the same way whether she is angry or not.

3. Expressivity modelling

We want to simulate that two different agents may behave differently in a same context and may express the same felt emotion differently. To this aim, we define a behavioral profile; that is a set of numerical values representing the variation with respect to a generic agent. This profile is defined as a set of numerical values, positive or negative. There is one value for each modality. These values may express that an agent has a very expressive face or that she barely uses wide arm movements. Moreover people may express their communicative acts by privileging one modality over the others. This instantiation depends also on the agent’s preferences. To model such a characteristic, we define a hierarchy of the modalities by associating to each of them a numeric value that represents their expressivity level.

Expressivity ought to be modelled differently depending on the modalities (face, gesture, gaze) it applies to. Indeed, expressivity when talking about facial expression is related to intensity (muscular contraction) as well as its temporal course; while when talking about gaze, it is related to factors such as length of mutual gaze, length of looking at the conversation partners; while when talking about gesture, expressivity is related to many parameters such as the strength of a movement, its tempo, its dynamism, its amplitude. Moreover expressivity of a behavior may be expressed not only by the intensity of the signals within a given modality but also by how the signals are dispatched over all the modalities.

4. Algorithm

The system uses a database that contains a large list of (meaning, signal) pairs. It takes as input:

1. agent’s definition: (a) the value of the agent’s global expressivity attributes (Overall activation, Spatial extent, Temporal, Fluidity, power/energy, Repetitivity) (b) the agent’s preference of modalities hierarchy (agent’s tendency to use one modality over another when communicating)

2. a text to be said by the agent annotated with Affective Presentation Markup Language (APML) tags (it specifies the agent’s behavior at the meaning level) [1].

3. meaning intensity: intensity value of each meaning tag

4. a set of values representing the intrinsic and contextual influences

At first the system selects, within the database of all possible nonverbal behaviors, which signals may be chosen to convey a given meaning. The selection is constraint by the value of the input parameters. Finally it sends this selection to the expressivity animation module that computes the expressivity parameters for each behavior within each modality.

5. Conclusion

In this paper we have presented a system that takes into account different types of influences to compute the behavior of an agent. The system determines not only the appropriate signals but also their expressivity. We aim at defining different strategies to combine signals (using redundancy, complementarity, substitution, and so on) while maintaining behavior coherency. Evaluation phase of how expressivity is perceived still has to be done.

References


