Socially Aware Virtual Characters: 
The Social Signal of Smiles

At first, machines were mainly used to solve complex mathematical problems. Today, they often embody more and more roles typically endowed by humans, such as a tutor in a virtual learning class or an assistant for virtual task realization. Moreover, the paradigm highlighted in [1] reveals that the user’s relation with the computer is intrinsically social, with high similarities compared to interpersonal relationships. For instance, people have no problem using communicative means (verbal and nonverbal behavior) to interact with artificial entities and apply polite rules. One particularly interesting aspect of the human–machine interaction is the bidirectionality of the relation: users exhibit social behaviors to the computer, and the machine’s behavior affects users. Reeves and Nass [1] have shown that people tend to react naturally and socially to computers, as they would do to another person. In such a human–machine, which can be seen as a particular social context with their corresponding meaning; this social signal may be displayed by a virtual character in a way that the meaning would be perceived correctly by the user—in other words, as positive and warm. A socially aware virtual agent should be able to select the social signals suitable to the situation and to the message that it wants to convey. For instance, if the user’s utterance is in contradiction with the virtual character’s beliefs, it may decide to communicate disagreement.

To respond in a social manner, in synchrony with the user, a virtual character should be able to display, in a timely manner, social signals that can be perceived by the user. For instance, to communicate disagreement, a raise of eyebrows may be shown. The value of a signal depends on its context, whether this corresponds to other signals displayed by the speaker or by the listener. A smile shown on a sad face may be perceived as a smile of despair. However, if a virtual character smiles when the interlocutor shows sadness, it may be perceived as cold and heartless. For this purpose, the virtual character should be endowed with a repertoire of expressions with their corresponding meanings; this meaning being considered within a given discourse context.

In this article, we aim to study in particular the social signal of a smile. The smile has interesting characteristics, both in terms of social functions and in terms of perception. A smile may convey totally different meanings—such as amusement, embarrassment, or politeness—depending on subtle characteristics of the face. Moreover, one’s smiling behavior may affect the other’s perception, his motivation, enthusiasm, and even the realization of a task. Both during human–human and human–machine interactions, the impact of smiling behavior might vary from positive to negative depending on when and which smiles are expressed, and in response to what. In the next section, we explore the different meanings that such a social signal can transmit depending on subtle characteristics displayed on the virtual character’s face.

Last but not least, a virtual character should be aware of the impact that its social signals have on the users. Indeed, one’s overall social behavior during interplay may convey a social attitude. For instance, someone always expressing smiles during a dialog might be perceived as positive and warm. A socially aware virtual agent should be able to select the social signals appropriate for the social attitude that it wants to convey.

VIRTUAL SOCIAL SIGNAL REPertoire

To create a socially aware virtual agent, the first problem is to identify how a social signal may be displayed by a virtual character in a way that the meaning of the social signal would be perceived correctly by the user—in other words, how a repertoire of social signals with their corresponding meaning can be...
created. Different approaches have been explored. A traditional approach consists of exploiting the empirical and theoretical studies in psychology that have highlighted the social signals associated with specific social meanings. For instance, Ekman and Friesen [3] brought out some morphological characteristics that enable one to distinguish between smiles conveying happiness or politeness. Another approach to construct a repertoire of social signals is based on real data corpus analysis. To capture social signals expressed by people, several methods may be used. One method consists of recording videos of actors having the instructions to express specific social signals. Another method consists of collecting spontaneous expressions by putting people in situations triggering the social signal studied. For instance, a common method used to generate frustration is to simulate a glitch in a computer program. The second step is the annotation of the corpus to attribute labels to expressions, and to find out the morphological and dynamic characteristics of the social signals to create the virtual social signals repertoire.

Following a user-perceptive approach, another method to create repertoire of social signals for virtual characters consists of collecting a corpus of virtual character's expressions directly created by users. This method breaks with the traditional approach used to create a repertoire of expressions: instead of asking people to label existing expressions, users are at the heart of the creation process of virtual character's expressions. To create a repertoire of virtual character's smiles, we implemented a Web application, named E-smiles creator (Figure 1), that enables a user to easily create different smiles on a virtual character's face. Among all the smiles that have been studied and reported in empirical and theoretical research in human and social sciences, we turned our attention to three types of a smile: an amused smile (also called felt, Duchenne, enjoyment, or genuine smile), a polite smile (also called non-Duchenne, false, social, masking, or controlled smile) and an embarrassed smile. These smiles have the advantages of having been studied both from the encoder's point of view, i.e., the person who smiles (for instance, in [3]) and from the decoder's point of view, i.e., the person who perceives the smile (for instance, in [4]). The polite, amused, and embarrassed smiles have different morphological and dynamic characteristics that enable one to distinguish them. Morphological characteristics are, for instance, mouth opening or cheek raising. Dynamic characteristics correspond to the temporal unfolding of the smile such as velocity. Despite some specific muscle contractions associated with smile types, no consensus exists in the literature on the facial characteristics of amused, polite, and embarrassed smiles. Consequently, to identify the morphological and dynamics characteristics of the three smiles, we used the E-smiles creator application (Figure 1). Through radio buttons on an interface, the user could generate any smile by choosing a combination of seven parameters (amplitude of smile, cheek-raising, duration of the smile, mouth opening, symmetry of the lip corner, lip press, and the velocity of the onset and offset of the smile). We have considered two or three discrete values for each of these parameters (for instance, small or large for the amplitude of the smile). When the user changes the value of one of the parameters, a virtual character shows automatically the corresponding animation.

Considering all the possible combinations of the discrete values of the parameters, there are 192 different possible combinations, each one corresponding to a smile. The user was instructed to create one animation for each type of smile. Three hundred forty-eight participants (195 females) with a mean age of 30 years created smiles. We then collected 348 descriptions for each smile (amused, embarrassed, and polite). Based on this smile corpus and on a decision-tree classification technique, we have defined an algorithm to determine the morphological and dynamic characteristics of the smile types; that is, we investigated if we could characterize these smiles through common morphological and dynamic features. We have chosen to use decision-tree learning as this technique is well-adapted to qualitative data
and produces results that are interpretable. With the morphological and dynamic characteristics as input variables and the types of smile as target variables, we have obtained a decision tree in which the nodes correspond to the smile characteristics and the leaves to the smile types. In the resulting decision tree, ten leaves are labeled as polite smiles, seven as amused smiles, and three as embarrassed smiles. The advantage of such a method is to consider not only one single definition for each smile type but a variety of signals. That enables one to increase the repertoire of the virtual character’s expressions.

Based on this user-perceptive approach, a context-free repertoire of smiles with different meanings has been created. To ensure that these smiles are those that the virtual character is expected to display in a polite, amused, or embarrassed situation, a perceptive evaluation has been performed. The four best classified amused, polite, and embarrassed smiles have been selected. Different scenarios (of polite, embarrassed, and amused situations) were presented in text to the user. For each scenario, video clips of virtual character’s different smiles were presented. We asked users to imagine the virtual character displaying the facial expression while it was in the situation presented in the scenarios. The user had to rate each of the facial expressions on its appropriateness for each given scenario. The evaluation was conducted on the Web. Seventy-five individuals participated in this evaluation (57 female) with a mean age of 32. The evaluation revealed significant results showing that the generated smiles are appropriate to their corresponding context (for more details on the corpora of smiles, the proposed algorithm, and the perceptive evaluation, see [5]).

As highlighted in [6], the context cannot be ignored to create socially intelligent machines. In the study presented above, scenarios describing in text specific amused, embarrassed, and polite situations have been presented to participants. The participants have a European culture. Since the texts of the scenarios clearly describe the role of the virtual character and since they are presented to culture-specific users, several elements of the context are considered, such as the social role of the virtual character, the cultural conventions, and the social norms [6]. However, the main problem that remains is the involvement of the user. Indeed, the users read scenarios but don’t face the virtual character, and moreover do not interact directly with it. A social signal displayed by a virtual character may be interpreted differently depending, for instance, on the verbal message that comes with the social signals, i.e., the context of the discourse. For instance, a smile may be interpreted as ironic if there is a discrepancy between the negative verbal response and the positive behavior (e.g., a smile) of the virtual character to the user’s message. In the next section, we attempt to explore the user’s perception of virtual characters expressing a social signal during an interaction.

**PERCEPTION OF VIRTUAL CHARACTER’S SOCIAL SIGNAL EXPRESSIONS**

Beyond the synthesis of a social signal correctly perceived by users, a socially aware virtual character should be able to evaluate how users view it. Users build a model of the virtual character’s social attitude during the interaction. Such a model evolves dynamically with the user’s perception of the virtual character’s social attitudes.

First, the situations in which the virtual character may express a particular social signal during an interaction should be identified. For instance, a head nod may communicate agreement or emphasis. Some social signals are obviously associated, by their meaning, to particular communicative intentions. For instance, an amused smile may be accompanied by the verbal communication of a positive emotion, such as “I’m happy today!” The direct link between a communicative intention and a social signal is not always straightforward. For instance, a polite smile may be expressed for different communicative intentions: a greeting or a back channel to express agreement. In a virtual character, an explicit model of its communicative intentions enables one to easily integrate the associated social signal that may accompany the verbal message. The communicative intention corresponds to what the virtual character aims to communicate to the user, for instance its emotional state or its uncertainty. In the virtual character GRETA, the communicative intentions and the associated social signals are defined through an Extensible Markup Language (XML)-based language (for more details, see [7]). Using this XML-based language, the different smiles identified through the user-created corpus described previously have been associated to particular communicative intentions for given discourse contexts. For instance, polite smiles are associated with the intention to communicate encouragement and to a back channel to show understanding.

Second, a social signal may or may not be displayed when the virtual character performed a communicative intention. For instance, a virtual character may decide to not express an amused smile when it communicates a joyful message. The expression of a social signal may impact the user’s perception of the virtual character’s social attitude. For instance, research has shown that smiles expressed both by a human or a virtual character enhances the social attitudes perceived by others. The effects on perception may depend on the social signal itself. For instance, the...
virtual character’s expression of a polite smile may have different impact compared to the expression of an amused smile [8].

The evaluation of the impact of a virtual character’s social signal expressions may be performed at different levels. A first level consists of studying the effect of the expressions of a social signal at an utterance level. One may compare the effects of the expression and the nonexpression of a social signal when a virtual character gives a message to the user. Both the impact on the user’s perception of the virtual character’s social attitude and the impact on the understanding of the message could be analyzed. For instance, in the context of smiles, we have conducted a study to measure the effects of the expressions of the polite and amused smile on the perception of the virtual character’s social attitudes and on the understanding of the verbal message transmitted to the user. Given the types of smiles considered, we have chosen situations in which the virtual character tells a joke to the user. Whereas the amused smile is displayed at the end of the joke, the polite smile is used to accompany the virtual character’s salutation at the beginning. A user watched videos of a virtual character telling a joke and indicated his perception of the following social attitudes: spontaneous, stiff, cold, warm, boring, and enjoyable. To measure the perception of what the virtual character said, we asked the user to indicate how well he understood the joke and if he liked them. To measure the effect of gender on the user’s perception of virtual character’s smiles, we have considered two different virtual characters: one male and one female (Figure 2).

![Bar chart of the mean of the perceptive warm of the virtual character depending on the smiles expressed (condition) and on the gender of the virtual character.](image)

<table>
<thead>
<tr>
<th>Gender</th>
<th>Mean of the Social Attitude Warm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Condition</td>
</tr>
<tr>
<td>No Smile</td>
<td>0.00</td>
</tr>
<tr>
<td>Only Amused Smile</td>
<td>1.00</td>
</tr>
<tr>
<td>Only Polite Smile</td>
<td>2.00</td>
</tr>
<tr>
<td>Amused and Polite Smile</td>
<td>3.00</td>
</tr>
<tr>
<td>Female</td>
<td>Condition</td>
</tr>
<tr>
<td>No Smile</td>
<td>0.00</td>
</tr>
<tr>
<td>Only Amused Smile</td>
<td>1.00</td>
</tr>
<tr>
<td>Only Polite Smile</td>
<td>2.00</td>
</tr>
<tr>
<td>Amused and Polite Smile</td>
<td>3.00</td>
</tr>
</tbody>
</table>

IEEE RECOGNIZES THE SPECIAL NEEDS OF SMALLER COLLEGES, and wants students to have access to the information that will put them on the path to career success. Now, smaller colleges can subscribe to the same IEEE collections that large universities receive, but at a lower price, based on your full-time enrollment and degree programs.

Find out more—visit www.ieee.org/learning
The results reveal significant effects of both the gender and the smiles on the user’s perception of the social attitudes and of the joke. For instance, Figure 3 illustrates the perception of the warm social attitude depending on the smiles expressed and on the gender of the virtual character (for more details on the result, see [5]).

One limit with the evaluation at the utterance level is the lack of interactivity with the user. Indeed, the user remains passive since he is not involved in the conversation with the virtual character.

A second level of evaluation consists of studying the effects of social signal expressions on the perception of the overall interaction. Note that the evaluation at this level is the most frequent in the literature. The objective in this case is to measure the effects of the expressions of a social signal on the perception of the user who is involved in an interaction with the virtual character. A questionnaire at the end of the interaction is generally used to collect the user’s overall perception of the virtual character and/or of the interaction. This level of evaluation enables one to capture the global perception of the user compared to the evaluation at the utterance level that measures the effect of a social signal expression at a given time $t$ of the interaction. In the case of smiles, to measure the global perception of a user interacting with a smiling virtual character, we have integrated the smiles identified based on the user-created corpus, in the platform [9]. This open-source platform enables users to naturally converse with virtual characters. The smiles of the virtual character are automatically selected depending on the communicative intention of the virtual character computing by the dialog module of the SEMAINE platform. We considered a male and a female virtual character (Figure 2). We asked to participants to interact for three minutes with each of two characters twice (in the no-smile condition and in the smiling condition). The user looked at a teleprompter, which consists of a semisilvered screen at 45° to the vertical, with a horizontal computer screen below it (Figure 4). The user saw the face of the virtual character.

The first results with 15 participants tend to reveal similar effects of the smiles at the global level compared to the utterance level. For instance, virtual characters tend to be more warmly perceived when they expressed polite and amused smiles compared to no smile. However, gender effect seems more pronounced in the context of an interaction. Also, one interesting aspect is that participants didn’t notice the difference between the no-smile version and the smile version, whereas significant effects on their perception appeared. The perception of virtual character’s social signals may be unconscious for the user during natural interaction.

CONCLUSIONS

To create socially aware virtual characters, a repertoire of smiles with different meanings has been produced starting from a user-created corpus of virtual smiles. The latter has been integrated in virtual characters and evaluated at different levels of interaction. Our next research step is to bring the agent and the user in an interactive loop. The agent ought to adapt to the user’s behavior in a timely synchronous manner. Moreover, the synchrony of social signals between partners of an interaction reflects the quality of the interaction and the partners’ affiliation [10]; this could be used to compute the user’s perception of the virtual character’s social attitudes.

ACKNOWLEDGMENTS

This research has been supported by the European Community Seventh Framework Program (FP7/2007-2013) under grant agreement 231287 (SSPNet). The authors would like to thank Paul Brunet, Gary McKeown, and Radoslaw Nieviadomski for their participation in the studies on virtual characters’ smiles.

**REFERENCES**


