

# Emotional Interaction in a Hybrid Conversation Group

Ido A. Iurgel

ZGDV e.V. (Computer Graphics Center)

Darmstadt, Germany

[iurgel@zgdv.de](mailto:iurgel@zgdv.de)

Key-Words: Applications of lifelike animated agents, Implementations of lifelike agents, Verbal and non-verbal behavior in animated agents, Emotion expression in animated agents, Theories and models of emotion and personality

## Abstract

*This paper presents a hybrid conversation group of talking heads, and discusses some solutions concerning emotions and social-emotional signs in more detail. A hybrid conversation group consists of real and virtual members, which meet to discuss some specific theme. The use of such a group, together with digital storytelling techniques, is a promising way of engaging the user in the exploration of complex and multi-faceted knowledge.*

*An import issue in a hybrid conversation group are the emotions and the social-emotional signs of the virtual characters, because those signs are necessary for the development of sub-plots in which the characters establish friendship and respect relationships with the user. Reciprocity - "Tit for tat" – turns out to be one important basic measure for the calculation of social-emotional signs of virtual characters, with the second measure being norm levels of exchange. Parametric variations of measurement of equity considerations of the virtual characters already comprise a wide range of personality trait differences.*

## 1. Introduction

One of the lessons we can learn from the overwhelming international success of "Sofie's World" is that there is considerable interest in complex themes as philosophy; at the same time, we see that an adapted presentation is essential for success.

In this paper, I present a first study on a hybrid virtual-real human conversation designed to present interactively even complex and multi-faceted knowledge areas such as philosophy, art or politics<sup>1</sup>.

An application of this sort, when elaborated in collaboration with experts and artists, should have a considerable surplus for the user, comparing to old-fashioned books. Think of a virtual, interactive talk show with virtual characters of pronounced personality, where the dialogue sequences are sometimes amazingly insightful – because they were authored in advance by experts, of course.

A conversation group allows to distribute theses, points of views and interests among different personalities, as well as a dialogic development of the theme<sup>2</sup>. The use of a group – instead of a single artificial dialogue partner – alleviates shortcomings of natural language understanding, since in group conversation it is quite normal that contributions of the participants do not fit exactly to each other, or that contributions simply go unnoticed.

The virtual characters in the conversation group, however, do not have full autonomy because, in part, they have to present sequences of contribu-

tions prepared beforehand. These limitations in autonomy arise from "digital storytelling" considerations, given that there is a specific story to be told, a message to transmit, and that these elaborate structures cannot "emerge" alone from the free intercourse of user and virtual characters, but have to be authored by an artist or expert.

But where is the place for user interaction if the main content is basically fixed? Certainly, interaction allows for

- Navigation – "tell this again";
- Adaptation to the user, e.g. to different knowledge levels;
- Urging the user to form an own opinion;
- Deploying sub-plots.

The notion of a sub-plot is borrowed from cinematography, where a main story branch is often accompanied by secondary branches that do not influence the main story, but contribute to the dramatic effects. One kind of interactive sub-plot of peculiar importance here involves socializing with the virtual characters. The hypothesis adopted in this study is that the user will appreciate to be able to insult or to praise his artificial counterpart, besides being involved into the rational deployment of thoughts and opinions, and that he will appreciate it the more if social sub-plots disclose personality traits of the virtual participants. This kind of deviation from the main line will allow the user to investigate the personality traits of the virtual characters and, by doing so, to associate this traits to the opinions defended by the character. In this way, an opinion receives a "face" of someone the user "knows". Hence, a social sub-plot does not change the content of the thematic dialogues, but this sort of deviation from the main line of deployment should provide an important additional source of motivation and interest for the user.

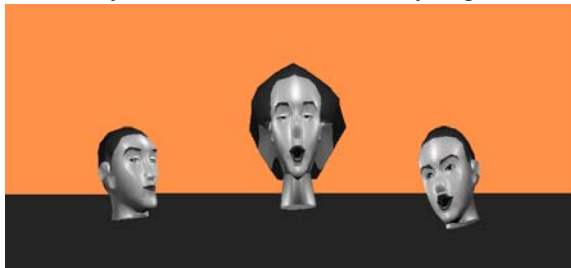
<sup>1</sup> This study is the first step into a 3-year project funded by the EC that will begin in September 2002, in which the discussion theme will be art. For the moment, however, philosophy is easier to handle because no additional multimedia is required.

<sup>2</sup> Cf. André et al. (2000).

Another lesson derived from mass media, besides the importance of the kind of presentation and of deviations from a single line of deployment, is that the personalities of a show should be appealing and pronounced. We all know that talk shows with famous personalities fascinate us also because of the idiosyncrasies of the participants. The consequence of this for the design of a hybrid conversation group is that it is important not to use fixed, though possibly “believable” character expressions, but to give designers of virtual characters the tools to construct virtual characters with the expressive, emotional and cognitive characteristics they feel to be appropriate for the settings, possibly employing exaggeration or simplification.

The example application I want to present here revives Plato’s “Symposium”, where Plato explains the nature of beauty by allegedly reporting a dialogue between Socrates and the priestess Diotima. The plot consists of Socrates, Diotima, Fritz, and an user avatar.

The role of Fritz is notable: Fritz and User, together, play roughly the role Sofie has in “Sofie’s World”; it is their task to ask the right questions that push the investigation forward. Whenever a certain key word is needed for the story to go on,



**Socrates, Diotima, and Fritz. The User Avatar is not shown.**

and the user does not provide it, Fritz will do it. That is, Fritz collaborates with the user as a kind of mediator between real and virtual world, standing in if the system does not get along with the human contributions, but who also explains, if necessary, what is being said in more detail to the user.

In what follows, I shall give a snapshot of ongoing work, rather than entering much into technical details.

## 2. Related Work

André et al. (2000) developed a presentation group starting from a similar motivation. Their emphasis is, however, not on the optical appearance and on digital storytelling, but on dialogue and natural language generation in dependence of personality traits of the virtual participants. The application is not interactive, the themes presented are quite different, and no episodes that were authored in advance are used.

I. Poggi and C. Pelachaud (2000) already pointed out the necessity of decomposing facial semantics

to represent changing mental states, and we follow this hint to a certain extent when advocating for a componential approach to facial expression.

B. Hayes-Roth and van Gent (1997) developed improvisational agents in a story telling environment, but they were not designed for interactive discussion. We follow Mateas and Stern (2000) in denying the agents strong autonomy, because of the necessity of having a story and a deployment, but their focus lies on artistic experience, whereas the present application aims at edutainment, which seem to imply less variability due to interaction (But cf. Spierling 2002).

Several researches investigated emotions in virtual characters (e.g. Elliot & Brzezinski 1998, Lester & Stone 1997). Prendinger and Ishizuka (2002) and El-Nasr et al. (2000) already considered affective interaction history in their work. However, they used a rather behaviorist framework, and we expect a model of more deliberate emotional reactions to produce results that are more appropriate to “intelligent” discussion partners in a virtual group.

## 3. Technical Overview

Every virtual character consists of three modules.

### Agentive person<sup>3</sup>

This module is responsible for rule-based, deliberate decisions. We use the production system Jess, in which a BDI-style<sup>4</sup> cycle of perception-deliberation-action is implemented. Jess is especially suitable for fast testing of ideas, allowing for changes in the rule set even at run-time.

### Bodily person

The word “body” means here that arousal level and reactive processes are simulated. Situation memory is also located here.

### Manifest person

The manifest person is the module that displays agentive decisions, emotions and processes of the “body”.

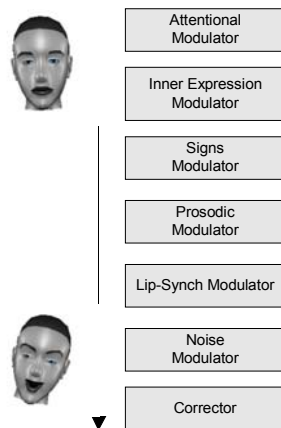
The virtual characters communicate through a blackboard. All aspects of the content related language capabilities are delegated to a central dialogue manager to ensure the correct sequences of arguments and counterarguments. The virtual characters are responsible for voice, emotional signs, timing and other non-content related language elements, as they are for every non-verbal action or state. The user is assisted by an avatar, which, technically, is almost identical to the virtual characters of the other participants.

The face is constructed in an assembly-line-like process, which is intended to mimic real dependencies of facial functions. Special authoring tools allow a character designer to freely define facial expressions that the assembly-line will handle. That

<sup>3</sup> A similar distinction was made by Elliot 1992.

<sup>4</sup> C.f. Rao and Georgeff, 1995.

is, the system at no point does associate any specific face to any semantic function, the mapping being in the responsibility of the artists.



**Assembly-line for faces**

To handle interaction, we distinguish interactive and non-interactive episodes in the course of a conversation session. Currently, the distinction is quite strict. In a non-interactive episode, the virtual characters do not give turn to the user, and react to user input only non-verbally, signalling e.g. “do not disturb”. The virtual characters initiate the interactive parts pro-actively. We use many instances of the GNU-licensed chatbot ALICE<sup>5</sup> in parallel to interpret user contributions and generate answers in interactive episodes. ALICE employs a smart pattern matching mechanism and is integrated here not only to allow free dialogue sections, but also to filter out navigation instructions like “please go on”, “repeat this” etc. The ALICE-instances are integrated into the dialogue manager, which decides about which character is to react verbally to user contributions and about whether a non-interactive episode should follow<sup>6</sup>.

In the next section, I want to describe the emotions and the social-emotional signs of the virtual characters in more detail.

#### **4. Reciprocity and Norm Levels in Virtual Social Interaction**

In the introduction, I mentioned the notion of a socializing sub-plot that does not change the thematic dialogues. How should we represent social relationships and their evolvment? Usually, psychology of the individual is taken to provide the theoretical base for emotions in believable agents. Without denying the importance of modelling individual emotional states, I still would like to stress the necessity of handling social psychological insights as well. The reason for this, on one hand, is

that social structure is one central trigger for emotions (cf. Kemper 1984), and on the other, that individual emotions and emotional signs also aim at changing social structures – they are not only causal, but also strategic (cf. Frank 1988). That is, believable agents should also have some representation of social relationships and act upon those representations. In order to understand the role a social psychological approach plays in modelling a group of virtual characters, I first abstract completely from individual psychology and try to produce as much believable behavior as possible on a social-structural basis alone.

Probably the most basic explanatory term in social interaction is reciprocity, or “tit-for-tat”. “You love me, I love you”, “you respect me, I respect you”. Reciprocity is easy to handle numerically – “you hate me 5, I hate you 10”. By implementing reciprocity measures, we enable the agents to

- react in the right measure, not too tenderly and not too offensively;
- react cross-modally, e.g. alternatively with a reproaching gaze and/or verbally;
- handle “cold” social-emotional signs of low arousal level, e.g. a benevolent smile
- “pay back in small amounts” until “the bill is met”.

It is appealing to focus on reciprocity for controlling social-emotional signs and socially relevant verbal contributions in our settings because this does not involve uncontrolled emotions and high arousal levels, which in most cases would not fit to the personalities of virtual philosophers. The emotional expressions became rather rational and deliberate.

The virtual characters are able to compute equity because every facial sign and every speech act carries meta-information about their social effects. E.g. if the user says “I hate you”, the dialogue manager will, with the help of ALICE, evaluate this speech act to have a value of  $-20$  and inform the virtual character, who, in order to restore balance, could counter e.g. with several facial signs of value  $-5$ .

Interestingly, the virtual characters can be given first personality trait differences solely on the basis of variations of “social book keeping”. Some examples of parameters we use are

- weighting according to whether a value is positive or negative; that is, a certain character might ignore remarks of sympathy, but react heavily to signs of antipathy;
- limiting maximum reaction values, e.g. a character might well be offended by the user, but react through a series of indignant facial signs, rather than giving back the insult.

To be sure, reciprocity and balance thriving alone do not lead to believable social-emotional behavior. The most obvious aspect missing is that we need, in addition, a representation of how social structure is modified by the exchange. Here, it should be

<sup>5</sup> Cf. [www.alicebot.org](http://www.alicebot.org)

<sup>6</sup> It is also possible to use ALICE-resources to force dialogue strategies by defining steps of following themes, but we still did not exploit this feature sufficiently.

stressed that after e.g. an exchange of insults, the relationship between the participants changes, even if reciprocity is reestablished. Undoubtedly, it would be strange for the participants to go on as if nothing happened. I account for this by introducing an additional “social distance level” for each exchange dimension. The social distance level increases with every positively valued contribution, and decreases with every negative contribution. E.g., if you say “I hate you” to virtual Socrates, the social distance between you and Socrates is likely to change. Note that the such distance levels are also assumed by a number of social structural theories<sup>7</sup>.

Finally, another point we can learn from social sciences is that there are individual norms for social distance: Socrates might be unsatisfied if you become very intimate, but also if you do not express any sympathy at all. By implementing social distance norms as goals of believable agents, the virtual characters get two competing, an sometimes incommensurable goals: to maintain a certain social distance and to maintain reciprocal balance. The reaction to this conflict certainly depends on their personality settings. To this aim, we provide the character designers with sliders and buttons they can use to set several parameters concerning the preferences of the virtual characters.

## 5. Display of Inner States and of Facial Signs

This section will explain how the facial expression is build, based on equity and distance level considerations.

To begin with, we distinguish between an “expressive face” and “facial signs”. The expressive face displays cognitive or emotional “inner” states. The “facial signs” are signs that “overlay” and “mix” with the expressive face. The idea standing behind this is that facial signs change their appearance according to the inner state of the character; if the virtual character has a certain way of showing interest, but he is unhappy, his unhappiness must remain visible even when he displays an interest-sign. Technically speaking, the expressive state is a direct mapping from properties of the individual situation to static facial expressions, whereas facial signs are short animations triggered by the agentive person due to e.g. social balance calculations.

### Expressive face

What are the “inner states” of a virtual character which are relevant to facial expression, and how are these states mapped to the face? If you start from AI, it seems natural to begin with basic emotions words, since one can reify them in the rules of a rule based system. However, there is a psychological and a graphical reservation to this.

From a psychological point of view, basic emotional faces reflect violent feelings<sup>8</sup>, which are not always the most appropriate expressions for a conversational group. From a graphical point of view, emotional base faces sets are difficult to combine without distorting their meaning. How do you “mix” graphically e.g. a face that expresses fear with a face that expresses anger, to create a face that is not only believable, but expresses exactly “fear and anger”?

The componential approach to emotional expression is a welcome alternative. This approach tries to directly assign mental states to facial features, e.g. attentional activity to raised brows, or pleasantness to the raising of lip corners. One core benefit of this approach should be the possibility of clearly displaying ambivalent faces.

In the compilation of several theories in the article of Smith and Scott (1997), the mental states associated with componential facial actions are (i) pleasantness, (ii) goal discrepancy (iii) anticipated effort, (iii) attentional activity (iv) certainty, (v) novelty, and (vi) personal agency. Those states are almost exactly the same dimensions which Ellsworth (1994) claims to be the basic emotion eliciting conditions. That is, using emotion rules as an additional level of indirection is not necessary, at least not if you are handling “calm” emotions, since eliciting conditions can be mapped directly onto facial action. In case of both theoretical approaches, the task of defining what situational properties are emotionally relevant is the same.

Space limitations prevent me from reporting our experiments with different emotional antecedents for the componential approach in more detail. As a single example on how faces that are ambivalent in a meaningful way, and therefore interesting, can be built, I want to mention that ambivalence can be construed by assigning “pleasantness” a value inversely proportional to social distance, and “goal discrepancy” the numeric value of the difference between current goal and state, where the current goal is determined by values of discrepancies, personality weighting and some factor of inertia to avoid behavior-aliasing. For example, if you exchange a number of vows of reciprocal admiration with Socrates, and his personality is tuned to pursue more social distance, his face would look like this:



<sup>7</sup> C.f. Kemper 1984.

<sup>8</sup> Cf. Scherer 1984.

It is clearly understandable that Socrates feels at the same time both pleased and dissatisfied, and it is not surprising if he soon begins to show that he does not really love you so much as you thought.

### Facial signs

The designer assigns to every facial sign he created a positive or negative value in the dimensions “love” and “status”, as well as a certain discursive function, e.g. turn taking or feedback. The rules for using discursive signs are kept in the agentive person module. If a particular discursive sign is due, the virtual character chooses the one that best fits his social-emotional goals, e.g. he will listen to you “with love” or give you the turn “with contempt”. Every sign made immediately changes the equity and the exchange level values<sup>9</sup>. The example of the following section might clarify more the use of facial signs.

## 5. Example

First, I want to give a textual example that explains Fritz role in the setting. An interactive episode is just beginning:

*Socrates, to User: “Do you agree on Diotima's opinion on sensual love?”*

(Remark that Socrates is being pro-active, thus narrowing the scope of possible answers.)

*User, to Diotima: “How does abstract love evolve from sensual love?”*

(An instance of ALICE finds a matching pattern, which is underscored here.)

*Fritz, to Diotima: „Indeed, today most people think of sensual love as a source of lust, not of wisdom”*

(The matching patterns cause Fritz to “interpret” the user and Diotima to answer to this “interpretation”. Diotimas next contribution is a fixed part of Fritz remark.)

*Diotima, to User: “Maybe they have a general disability to perceive their own life as a continuous progress”.*

(It is insecure if Diotima’s answer is an appropriate answer to what the user said, but it is certainly appropriate to Fritz’ interpretation.)

Now, I want to exemplify reciprocity based social signs, starting with the same interactive episode.

*Socrates, to User: “Do you agree on Diotima's opinion about sensual love?”*

*User, to Diotima: “It is completely stupid!”*

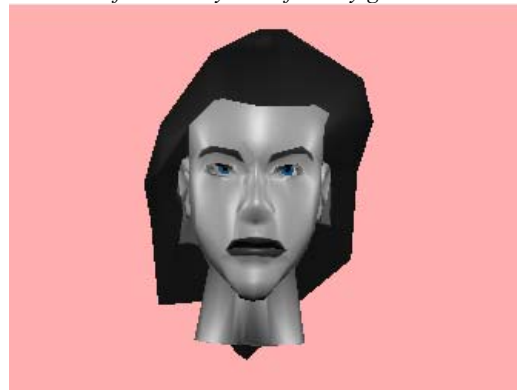
(This was offensive! Diotimas brows and lips show goal discrepancy and unpleasantness, respectively.)



*Diotima, to User: “Please say something substantial!”*  
(Diotima answer is ELIZA-like and fits to almost any user contribution. Because of parametric settings, she does not replicate immediately to the offence.)  
*User now says something to Diotima. Diotima listens by a nodding...*



*...which is followed by an unfriendly glance:*



(For a priestess, this is already a hard reaction. And she will continue to send small reproaching signals until the users offence is paid back for. Then, her brows will cease to show goal discrepancy, but her mouth will become friendly only if the user praises her competence.)

## 6. Conclusion

In this paper, I presented a snapshot of a work on virtual characters in a conversation group. This group is intended to present interactively even difficult themes like philosophy and art to the user.

<sup>9</sup> Cf. Bente et al. (2000) for more details on social-emotional signs in general and their – problematic – theoretical understanding in human sciences.

Because elaborated thematic deployments presuppose non-interactive parts, much of the charm of interactivity must be shifted to sub-plots, where social interaction becomes important. It was argued that, for the creation of virtual characters with pronounced personalities, we have to simulate deliberate, controlled emotional behavior, where “emotional” does not mean “loss of control”. At the same time, reciprocity considerations are an important aspect of deliberate emotionality, and parametrizations of individual exchange standards already cover some important personality trait differences. I hope to have shown that the proposed social-psychological framework provides an important theoretical basis for the generation of emotional display and social-emotional signs of virtual personalities. Of course, many other aspects and modules of human behavior could not be considered in this report.

Many issues are still unsettled. E.g., there is a need for some sort of reconciliation ritual, because the values of the conflicting goals of the model do not converge – as they do not in reality. Casualty and intention attributions are difficult to handle, but are essential to the choice of an emotional expression – the emotional reaction must vary depending on the reason of e.g. an insult.

For many questions concerning the details of emotional control in virtual characters, there is little scientific guidance. What appears to me the most promising way out is to systematically integrate artistic intuition, that is, to provide intuitive tools for character designers who shall decide, then, on the basis of their intuition and by try-and-error, which solution might produce the most appealing virtual personality.

## 7 . References

André, E.; Klesen, M.; Gebhard, P.; Allen, S., & Rist, Th.. Exploiting Models of Personality and Emotions to Control the Behavior of Animated Interface Agents. In: Jeff Rickel (eds.), *Proceedings of the workshop on "Achieving Human-Like Behavior in Interactive Animated Agents"*, in conjunction with the Fourth International Conference on Autonomous Agents, pp. 3-7, Barcelona, June 2000.

Bente, G.; Krämer, N. C.; Trogemann, G.; Piesk, J. & Fischer, O. (2000). Conversing with electronic devices - An integrated approach towards the generation and evaluation of nonverbal behavior in face-to-face like interface agents. In: *Proc. IMC 2000*, Rostock-Warnemünde, Germany, November 9-10.

El-Nasr, S. M.; Yen, J. & Ioerger, T. R.: FLAME - A fuzzy logic adaptive model of emotions. In: *Autonomous Agents and Multi-Agent Systems*, 3(3), 2000.

Elliot, C.: The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System, 1992. PhD thesis, Northwestern University. The Institute for Learning Sciences, Technical Report No. 32, 1992.

Elliot, C. & Brzezinski, J. Autonomous agents as synthetic character. *AI Magazine*, 19(2), 13-30, 1998.

Ellsworth, P.C. : Some reasons to expect universal antecedents of emotion. In: P. Ekman & R. J. Davidson (Eds.). *The nature of emotion: Fundamental questions* (150-154). New York: Oxford University Press, 1994.

Frank, R. *Passions Within Reason: The Strategic Use of Emotions*. New York 1988.

Gaarder, J. *Sophie's World: A Novel About the History of Philosophy*. Paulette Moller (Translator). Mass Market Paperback 1996

Hayes-Roth, B. und van Gent, R Story-Making with Improvisational Puppets. In W. Lewis Johnson, ed., *Proc. of the First International Conference on Autonomous Agents*, 1-7, 1997.

Kemper, Th.: Power, Status, and Emotions. In: K. Scherer und P. Ekman (Hrsg.), *Approaches to Emotion*, Hillsdale 1984, 369-383, 1984.

Lester, J., & Stone, B. Increasing Believability in Animated Pedagogical Agents. *Proceedings of the First International Conference on Autonomous Agents*. Marina del Rey, CA, USA, 16-21, 1997.

Mateas, M. & Stern, A. "Towards integrating plot and character for interactive drama", *Proceedings AAAI Fall Symposium on Socially Intelligent Agents: The Human in the Loop*, 2000.

Poggi, I., & Pelachaud, C. Emotional Meaning and Expression in Animated Faces. In: *Affective Interactions: Towards a New Generation of Computer Interfaces*. Ed. Ana Paiva, Springer 2000.

Plato. *Symposium*. Translated by Christopher Gill. Cambridge 1980.

Prendinger, H., & Ishizuka, M. Evolving social relationships with animate characters. *Symposium of the AISB-02 Convention on "Animating Expressive Characters for Social Interactions"*, London, UK, 2002, 73-78.

Rao, M. und Georgeff, M.: BDI Agents from Theory to Practice, Technical Note 56, AAIL, April 1995.

Scherer, K. R.: On the nature and function of emotion: A component process approach. In: Scherer, K. R., and Ekman, P., eds., *Approaches to emotion*. London 1984.

Smith, C. A., & Scott, H. S.: A componential approach to the meaning of facial expressions. In: *The Psychology of Facial Expression*. James A. Russell und J.-M. Fernandez-Dols (eds.). Cambridge 1997.

Spierling, U.; Braun, N.; Iurgel, I.; Grasbon, D.: Setting the scene: playing digital director in interactive storytelling and creation. In: *Computers & Graphics* 26 (2002)31–44.