

EVALUATION OF CONVERSATIONAL AGENTS: MARIÁ AND ET

Dra. Soraia Raupp Musse
Technological Sciences -
Postgraduate Program applied
to Computation
Universidade do Vale do Rio
dos Sinos – UNISINOS
Unisinos Avenue, 950 - Cristo
Rei – ZIP Code 93.022-000 -
São Leopoldo - RS - Brazil
+55 (51) 35901100 – extension
1701
soraiaarm@unisinos.br

Dra. Eliane Schlemmer
Postgraduate Program applied
to Education
Universidade do Vale do Rio
dos Sinos – UNISINOS
Unisinos Avenue, 950 - Cristo
Rei – ZIP Code 93.022-000 -
São Leopoldo - RS - Brazil
+55 (51) 35901100 – extension
1133
elianes@unisinos.br

Dra. Susane Garrido
Distance Education Coordinator
Universidade do Vale do Rio
dos Sinos – UNISINOS
Unisinos Avenue, 950 - Cristo
Rei – ZIP Code 93.022-000 -
São Leopoldo - RS - Brazil
+55 (51) 35901100 – extension
1959
susane@bage.unisinos.br

ABSTRACT

This paper describes ET and Mariá, two conversational agents with different characteristics. Mariá presents a more realistic character; however she is not an interactive agent, since we only provide short movies with predefined questions and animations. ET is an interactive agent, which can recognize keywords (in Portuguese) and visual react emotionally and textually. We describe the development of those agents as well as results obtained with interaction with subjects. We also present details about MECA, the system we developed to Model Embodied Conversational Agents.

KEY WORDS

Conversational agents, interaction, recognize keywords and visual react emotionally and textually.

1. INTRODUCTION

From imaginary friends to the most advanced artificial intelligence robots, the idea of representing the human behavior means to realize a paradigm of rational and irrational attitude. Indeed, it is comparable to a speaker in front of an audience, where sometimes certain ideas can be transmitted through a single body postures with no need of words. The Embodied Conversational Agent (ECA) (Cassel, 2000) is a specific type of synthetic autonomous agent that aims to establish friendly dialogues with the user, enough to be utilized as system interfaces, entertainment applications or computer aided professional and teaching tools.

This paper presents Mariá and ET, two different conversational agents. Mariá is a more realistic character in the visual point of view; however, she is not interactive since she is presented through the exhibition of short movies. On the other hand, ET is a very simple character, he is not human, but he provides interaction with the users. More than that, ET changes humor as a function of established user-dialogue.

Figure 1. Snapshot of Mariá



MECA System used to develop ET is also described in this paper. Section 2 presents some few related works on the domain of ECAs, while section 3 presents details about MECA. Section 4 describes experiments applied with volunteers concerning interaction with ET and Mariá, as well as results evaluation by using Piaget theory.

Figure 2. Snapshot of ET



2. RELATED WORK

Many works have discussed and present ECAs in order to describe conversational agents with different purposes. Some of them are mainly focused on Internet applications in order to provide easily and friendly communication (Tatai, 2003). Some ECAs include humor variation and emotions to provide a more believable and fun interaction (Barsoux, 1993) (Morke, 2000). Interactive agents are also used for performing tasks required by users, normally using natural language (Hildebrand, 2003). Cassel (2001) presents BEAT (Behavior Expression Animation Toolkit), which proposes visually expressive conversational behavior during the user-interaction.

In this paper, we present two conversational agents. First of all, Mariá is not really interactive, since she is presented through short movies. ET, on the other hand, was modeled by using our MECA system (described in next section) which is used to model conversational agents that can react by applying animations and by answering textually questions made by users. The main goal of this paper is to evaluate the impact of interactions with the users.

3. MECA SYSTEM

The System M.E.C.A. (*Modeling Environment for Embodied Emotional Communicative Agents*) is a tool that aims to aid the designer to model an ECA. MECA has three main functions. The first one is viewing and animating the model. A light structure is required in order to provide real time frame rates. That is why we chosen to use a pre-defined animation format, where each frame is stored in a MD2 file format, the same developed and used in the game Quake II (<http://www.idsoftware.com/games/quake/quake2/>). The

model offers the advantage of striped vertex geometry, where the instructions to represent the model are reduced

until one third of complexity.

The second function describes the possibility of agent to textually answer the users' questions. Indeed, ET has a database including all possible verbal and visual information that can be used in the dialogue with the user. It encompasses all entries about questions, answers, keyword and emotion-related animations, and images related to the interaction context.

All keywords can be extracted from a text file, from which is possible to exclude the most relevant words through a filter previously defined. The animations representing the agent visual behavior are linked to determined emotions defined by the ECA's developer. In addition, it maintains a decision module responsible for choosing the next action to be performed.

While we talk, our mind is flooded by the meaning of the words, which are metaphorically linked with another signs and so on, and we search for information in this senses beyond our intellect to formulate what we want to say. After we had formulated what we wanted to say, the unconscious takes the control of our expressions. Not even an experienced orator has idea of all kinds

of gestures or facial expressions one can perform during a speech (Cassel, 2000). Even actors, specialists of the artistic representation of human emotions, suffer of this unconscious cryptography.

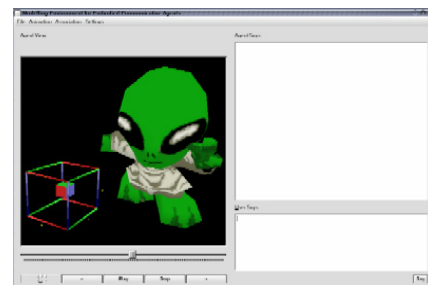
In this work, the execution of actions is described in a communicative agent emotional perspective, through simple parameters that give meanings to emotions. In spite of many works existent concerning emotion in ECAs, we chosen to implement a simple one where emotion is based on three simple parameters that interact forming an emotional triple, which can be visualized in MECA. Also, the user can call each one of parameters in the triple as he/she wants. Indeed, this triple forms the ECA emotional state, or simply *Zen* (as we have called). The user can modify it during the dialogue depending on the submitted expressions.

The third function offers the simulation module, when a model with a properly set reality, can be tested by the user. Thus, the user could perceive the behavioral transformations on ET's verbal and visual answers during the interaction. The dialogue and the ET's emotional changes can provide information to be analyzed aiming to evaluate quality in the relationship between users and ECAs.

3.1 Emotional Model

When we assimilate a given emotion, we are recording its meaning in our memory, and respectively all the reasons related about this expression. Our muscles, our entire motor coordination, our speech and analysis capacity, are linked to emotional meanings as part of the human cerebral reflexes.

Figure 3. Screenshot of the MECA system main window



In MECA, we associated keywords and animations to emotions. We defined an animation as a visual representation of an emotional status (*Zen*) of an ECA. Indeed, these associations are used in the agent communication model, when ET should decide what to do next. When the user interacts with ET, and depending on the submitted text, ET's *Zen* can be affected, in a positive or negative way. These definitions are pre-defined by the ECAs developer using MECA. An interface was designed in the system in order to visualize the current *Zen*.

Figure 4. The emotional state representation cube

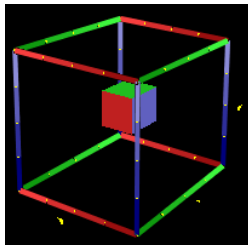


Figure 4 shows the three-dimensional voxel formed by emotions, based on model proposed by Darwin (1872). We defined each dimension as one emotion that can be changed by the MECA developer ($x = \text{anger}$, $y = \text{joy}$, $z = \text{fear}$). Each axis is divided in four segments, representing a qualification or degree of a determined emotion, which varies from zero (0) to one (1). For instance, in MECA's default the joy is represented by the blue axis, varying between sadness, neutral, happiness and its higher degree that is the agent's ecstasy or exhilaration. The anger (red axis), floats between calm, neutral and anger, and can reach the absolute wrath in its full intensity. The third and last axis is green, represents the agent's fear degree, and gives meanings to the stages of acceptance, comfort, and attention, and may reach its higher level, like panic or terror. In addition, there is a little cube that represents the current ECAs emotional state.

3.2 Answering a question

To answer a question sent by user, the ECA takes the expression typed by the user and performs a search in a verbal information database. In the database, there are several possible predefined questions and their associated answers. The questions and the answers are grouped by their context, i.e., questions and answers of same meaning or about some determined subject are categorized under a same topic. Each context has emotional factor that changes the emotional state of ECA during the interaction, being able to be positive, negative or zero, in each one of three dimensions. The answers for each context have an interval of emotional state of the agent, defined as minimum and maximum *Zen*. In the case of the system do not find the appropriate context to answer; the agent will use one of the standard expressions, stored in a list. Those answers may change the ECA emotional state as well.

The ECA also memorizes all reached contexts, having a k -sized memory and a patience limit of repetition $l < k$. The patience limit of repetition increases since user wants to insist on same sentence or expression. In this case, the agent will have its *Zen* altered by the specified value in the database.

Into a specific context, if there is more than one answer selected for next interaction, the system randomly chooses the answer to be sent. The verbal answer is performed by comparing of the user expressions with the predefined possibilities (stored in the database), without any natural language processing. Thus, one can define in MECA that any typing error means sadness and patience lose of the agent, changing it *Zen* to an undesired emotional state.

3.3 Animating the agent

Since a verbal answer is selected, the system finds an animation list to be executed in the moment of returning the expression to the user. These animations are linked to the answers as well the ECA *Zen* state. ET has several animations extracted from a MD2 model file format, which contains the three-dimensional geometric information, with indexed distinct animations frames. This geometric information are stored in the computer memory at the model load time, and used for visual representation of the model. Each defined animation is stored in a database, having an id, the minimum and maximum *Zen* factors, the show time duration of the animation (giving the notion of velocity or emotional intensity) and a linked keyword list. The verbal answer is broken into single words. Each word may have iconic meanings like an object size or place local indication representation. At the end of the answer (animation execution), the agent returns to its idle state, based on its final emotional state.

3.4 Images and action!

At last, MECA can store an images list in the database to be loaded during the dialogue with the user. Like the animations, keywords can be linked to images to build the agent memory. Images can be useful to illustrate an answer. The images have an identifier, a keyword list, and must be in the JPEG or Bitmap format, as well as to have symmetric dimensions in the power of two, standard to the OpenGL textures.

After selecting the textual answer, the animation and an image from the images list, the agent sends the answer to the user, showing the images like background slides during the execution. It can be seen in figure 5, and in this case, user asks for location of UNISINOS' library.

Figure 5. ET indicates the place of UNISINOS library building



4. ET AND MARIÁ: ASPECTS OF PERCEPTION AND REPRESENTATION

Mariá is a conversational agent developed at CROMOS Lab/UNISINOS (Schlemmer, Borba, Bica, Musse, Fagundes, Amorim, 2004) with the main goal to present news to the internal community. Mariá is a visually appealing female model that acts in a more realistic way, especially if compared to ET. However, Mariá does not offer interactive dialogues. Mariá's interaction occurs through the exhibition of short movies (normally presenting a specific subject). Figures 6 and 7 show some images of Mariá.

Figure 6. Mariá shows the location of a future student dormitory at University



The methodology used for this analysis consists in a qualitative evaluation performed by fifty (50) students (Human sciences courses) during a dialogue established between them and the conversational agents. For Mariá, the experience was a visit to her website; for ET, students executed MECA and interacted with the agent. Students described in questionnaire information about their sensorial perception, representation, psychological relationship (established with agents) and information about esthetic and cultural traces.

Piaget offers the theoretic basis for the analysis described in this section, in which the main idea is that knowledge is not a copy of reality, meaning that the beginning of the intelligence is not due to the perception. This concept breaks with the empiric tradition. Piaget and Inhelder (1993) mention that beyond of existing sensations and of the reason in mental life; we cannot be negligent with action importance: “the perception constitutes, in fact, in a private case of sensorial-motor activities”. (Piaget and Inhelder, 1993).

Under this perspective, the representation is an extension of the perception what includes the sensations from interaction environment-subject, as well as the conscious constructions each subject did as a function of his/her experience, memory and mental scheme (mental structures developed by people during the life¹). The perception and the representation suppose an abstraction that can occurs in different levels, and can be originated from processes of empirical abstraction (reactive behavior) as well as from reflexive abstraction (cognitive or deliberative behavior). According to Piaget, all new knowledge supposes an abstraction. To abstract in a large sense consists to differentiate; to separate a characteristic, means to isolate some characteristic in relation of others.

In this investigation, some subjects, during the interaction with ET and Mariá, perceive and represent this interaction through processes of empirical abstraction what consist in an abstraction merely descriptive, referring at an immediate establishment; it is everything what the subject extracts directly from the observation. However, it is relevant to remind that according to Piaget, in order to abstract the properties of an object, the subject needs (firstly) instruments of assimilation what actually means to have “a meaning” for him/her. The subject needs to establish relations and theses come from sensorial-motor or conceptual schemes constructed before the abstraction and are not given by

object. The empirical abstraction was verified with the following kind of questions mentioned by the subjects

1. Representation of the physical environment;
2. Physical representation of the agents ET and Mariá;
3. Changes in the physical representation of the ET and Mariá;
4. Description of elements (in general) which called more attention;
5. Description of elements (in general) which called less attention;
6. Suggestions to ET and Mariá.

Some subjects during the interaction with ET and Mariá perceive and represent it from processes of reflexive abstraction (Piaget, 1995), which support actions and operations coordination, i.e. the subject activity. Always when an object is modified by subject actions, it is enriched by the properties extracted from his/her coordination. In other words, the process of reflexive abstraction happens when the subject extracts data which is not explicit in the object. Consequently, what the object informs after subject’s action, describes information, which does not exist before the subject’s action. It was evidenced in following users’ questions:

1. Feelings about the interaction with the agents;
2. Relations established with the agents;
3. Questions of memories, experiences, interest;
4. Differences perceived between human-human conversations and human-ECA conversations
5. How subjects’ ideas concerning ECAs could be used in educational processes to contribute to the learning.

The most important differentiation found in this investigation, if we compare both ECAs, consists in the communication attributes. Concerning ET, even if he is much less realistic than Mariá, the communication attribute composed by immediate interaction using written language, expressions changing and humor fluctuation result in a better interaction with the subject. On the other hand, Mariá does not present this attribute, and consequently does not reflect the interaction aspects if compared with ET. Considering perception and representation, aspects of communication and language reflect intensively in the cognitive construction of the subjects, altering these same subjects in different ways.

Most people who interacted with ET could abstract in a reflexive way. The opposite happened with Mariá. It means that even if Mariá is visually much more realistic than ET, the perception and communication process established during the interactions with ET were much more efficient in the process of mental representation. The subjects’ considerations about Mariá describe much more the physical context of the agent. Consequently, people preferred to interact to ET than to Mariá.

Figure 7. Snapshots of Mariá



5. DISCUSSION AND FUTURE WORK

Concerning MECA, the system seems able to be used, since the learning and representation processes were described for subjects. Even, if the tested agent (ET) is not at all visually realistic. We intend, as future work, to prepare a more realistic agent to be tested in MECA platform, probably Mariá will be used for this task.

On the other hand, the visual aspect without an efficient interaction does not seem to be a good strategy to present a kind of information that users should create representations and reflexive abstraction.

As future works, we intend to test more ECAs and ask for more subjects to be investigated in order to enlarge our sample and improve our conclusions. Also, we are interested in using other interactive platforms as discussed by Garrido (2003,2004) and Schlemmer (2002).

REFERENCES

- [1]Cassel, J.; Sullivan, J.; Prevost, S.; Churchill, E. *Embodied Conversational Agents*, MIT Press, 2000.
- [2]Taitai, Gábor; L. László. The chatbot who loved me. Proc. *ECA Workshop of AAMAS 2003*.
- [3]Barssoux, J. L. *Funny Business: Humour management and business culture*. Cassell, New York, USA, 1993.
- [4] Morkes, J. Nass, C.I., Kernal, H.K. *Effects of Humour in task-oriented human-computer interaction and computer-mediated communication: A direct test of SRCT theory*. Human-Computer interaction 14 (4):395-435.
- [5] Hildebrand, M.; Eliëns, A.; Huang, Z.; and Visser, C. Interactive Agents Learning their Environment. Proc. *IVA 2003*, LNAI 2793, pp.13-17, 2003.
- [6] Cassel, J.; Vilhjálmsón; H. Bickmore, T. BEAT: the Behavior Expression Animation Toolkit. Proc. *SIGGRAPH '01*, L.A. CA, p. 477-486. Aug 2001.
- [7] Schlemmer, E.; Borba, M.; Bica, F.; Musse, S.; Fagundes, T.; Amorim, A. Mariá, um agente comunicativo. Proc. *SIGRADI2004* – São Leopoldo, Brazil, 2004.
- [8] Piaget, J., Inhelder, B. *A psicologia da criança*. 12.ed. Rio de Janeiro: Bertrand, 1993.
- [9] Piaget, J. *Abstração Reflexionante: relações lógico-aritméticas e ordem das relações espaciais*. Porto Alegre: Artes Médicas, 1995.