Model of Integration of Verbal and Nonverbal Expressions for Credible Virtual Agents

Jazmín Guadalupe Ramírez-Medina, María Lucila Morales-Rodríguez, Nelson Rangel-Valdez

Tecnológico Nacional de México,
Instituto Tecnológico de Ciudad Madero,
Mexico

{G20073011, lucila.mr, nelson.rv}@cdmadero.tecnm.mx

Abstract. For the development of a credible virtual agent, it is important that this agent has characteristics to create a more immersive experience for users and increases the effectiveness of communication. One of these characteristics is the ability to interact through verbal and non-verbal communication. That is why in this work an integration model of dialogue, facial and gestural expressions are proposed, where it is essential to consider the attributes of emotions and personality because these directly influence the expressions. This paper proposes the use of an optimization method named ELECTRE III for the process of selection of dialogue, facial and gestural expressions, using a corpus characterized by diverse criteria based on the influence of personality and emotions.

Keywords: virtual agent, verbal expressions, nonverbal expressions, selection, ELECTRE III.

1 Introduction

The moment we enable a face-to-face conversation aspects like verbal and nonverbal language are involved. If we develop a virtual agent that can participate in conversations and has the appearance of a human, it would be expected that it acts like one. For this behaviour to be congruent it is important that it communicates through dialogue, also with facial and corporal expressions as well. For the integration of the different elements that implicate verbal and nonverbal expressions in a virtual agent, it is proposed a selection model based on an outranking method called ELECTRE III [6]. This will allow us to create a credible virtual agent that can respond to social interactions with a natural behavior similar to a human. This way it will be able to make the user immerse in the virtual environment and develop a feeling of social presence.

The development of a credible virtual agent involves elements that need to be considered so that the communication of the agent is coherent and it interacts in a similar way to the human behavior. To accomplish this, it is necessary to develop a process of integration, in which the following areas are taken into

consideration: face-to-face conversation where language is involved, nonverbal behavior, emotions and personality, because human behavior is influenced by these areas [7]. This way the agent will be able to hold a fluent conversation using dialogue, facial and corporal expressions simultaneously, in a natural way and accordingly to the presented situations.

2 Related Work

This paper will provide a brief description of the background of projects related to credible virtual agents and how these impact the development of this project.

The work of Gratch [7] was one of the first ones where different elements that need to be considered to build a credible virtual agent were formalized. For the author, an agent must look and act like a person and at the same time, be able to enable a conversation. The main goal of this work is to develop a modular architecture and interface standards that allow future researchers to study each one of the three key areas of virtual human development, which consist of: face-to-face conversation, emotions and personality, animation of the human figure. All these elements raised by the author will serve as a basis for the realization of this work.

The work of Rosis [11] and colleagues was to develop an agent that has a natural appearance and behavior, for this very reason a prototype of a talking 3D head named "Greta" was designed. To reach this goal a cognitive model was presented, which was named "mind", where its beliefs, desires and intentions were established. This model is used to simulate how the agent reacts to the events that occur during a conversation. The second model that was intended to complement the cognitive model, was the "body". This model has access to a repertoire of signals that the agent can use in its communication process, for example, the facial expressions, head movement, sight direction, etc. The union of mind and body modules allows that the different contents of the message are expressed by the agent.

Greta was designed as a BDI [5] (Beliefs, Desires, Intentions) agent, whose state of mind integrates a representation of the beliefs and goals that drive the feeling of emotions and the decision to show or hide them. This characteristics will allow the user to have the impression of communicating with a person and not with a computer. As it was mentioned, for an agent to have a natural appearance, similar to the one of a human being in a face-to-face conversation, it is important to automatically generate the nonverbal expressions from the dialogue.

Because of the nonverbal behavior being affected by the personality, it is important to generate expressions that go accordingly with the personality of the agent. The work of Ishii [9], shows how the personality can help to improve the prediction of the nonverbal behavior for the whole body, this is what we refer to head movement, eyes, hands and posture. For the making of this study, verbal and nonverbal information was recollected, personality traits of a human conversation were recollected as well.

The model that was used to represent the personality was Big Five [8] (approaching, conscience, extroversion, likeability and neuroticism), which is an indicator of personality traits and it is one of the most accepted in academic psychology. To evaluate the significance of the Big Five model in nonverbal language, models of behavior generation were implemented with and without the Big Five model. According to this experiment, it was proved that the Big Five model is useful to generate nonverbal expressions. With this work, it is possible to identify the importance and effect that personality produces in the generation of nonverbal expressions.

The work done by Delgado [3] was the development of an agent which integrates the element of personality in the process of selection of phrases from a corpus characterized through acts of speech, using the ELECTRE III selection method. Delgado proposes an architecture with five principal components that are required in a conversational agent: perception, memory and knowledge, behavioral model, answer selection and action. From these modules the author incorporates the influence of the personality for the selection of phrases from an agent, providing more realism and credibility.

At the same time another work that was analyzed, was the one made by Reyes [10], which goal was the development of a software mechanism that integrates the selection of phrases and nonverbal expressions in assistants and/or video games based on web platforms. This work is based on Delgado [3] for the selection of phrases and uses the character "Kathia" as a representation of the developed animations [12]. This paper proposes a deliberative agent architecture and the analysis done for their development.

The agent personality influences the process of selecting speech, facial and gestural expression from a characterized corpus while measuring the similarity based on outranking relationships and the impact of personality in the preferences of the user as a measure of expression selection.

3 Theoretical Framework

In this section, it will be presented the theoretical framework of the development of a conversational deliberative agent and the role of the personality in the selection process of expressions using the method ELECTRE III.

3.1 Conversational Deliberative Agent

A Conversational Deliberative Agent (CDA) [5] has to emulate human behavior within speech, facial and gestural expressions, so it needs its own traits and reasoning that defines its responses during a conversation. This kind of agent needs to integrate beliefs, desires and intentions, alongside its perception of the conversation environment to select an expression that responds appropriately to the context of the conversation. The selection of dialogue, facial and gestural expressions in this virtual agent is based in an outranking method called ELECTRE III.

3.2 ELECTRE III

The ELECTRE model consists of a family of methods that belongs to the area of multi-criteria assistance of decision making [6]. The goal of the ELECTRE III model is to order different alternatives in a problem, leading to the generation of diverse criteria. This model allows the incorporation of the imperfection and uncertainty that is always present in a decision-making process while defining the parameters of preference and indifference [2]. To achieve this is necessary to use a ponderation for the criteria.

Traditional methods of classification come from preference and indifference. In ELECTRE III for every pair of alternatives, there is a measurement of concordance and discordance, the goal of the model is to combine these two measurements to produce a measurement for the degree of over-classification, in other words, an index of credibility, which validates the process of decision-making of choosing one alternative over the other one in a problem [6]. For the application of the ELECTRE method, a matrix of performance needs to be developed [2]. Where each linguistic variable needs to be indicated and also needs to receive a weight of importance.

After establishing the values of these variables, making a matrix that compares the values of the alternatives against the criteria that were previously established for the problem and lastly making a matrix that takes into consideration the level of concordance, a decision can be made about which alternative is the best for solving the problem presented.

3.3 Personality

To take into consideration the personality attributes for the selection of the nonverbal expressions in the agent, the MBTI [1] (Myers-Briggs Type Indicator) model was used. The MBTI is a psychometric test designed to evaluate the personality type of people from four dichotomies that are defined as polar opposites in eight categories. Each category is represented by a letter and because of this the results are shown in 16 possible combinations of four letters. Based on these combinations, four big groups can be identified: analysts, diplomats, sentinels and explorers [4].

This paper will only show the description of the characteristics of the personality type that was selected for the agent: **IFNJ Lawyer-Diplomatic**, the people with this profile tend to look for meaning and connection in ideas, relations and material possessions. They want to understand what motivates people and they are really perceptive about others. They are conscious and committed to their values. They develop a clear vision about what is best to serve the common good, usually organized and decided on the implementing their vision [1].

4 Proposed Architecture for the Integration of the Selection of Verbal and Nonverbal Expressions of a CDA

This paper proposes the architecture presented in Fig. 1 as a means of building a CDA that integrates verbal and nonverbal communication. This architecture is constituted by five principal components: perception, behavioral model, internal state of the agent, selection model and animation. All of these components must exist in a virtual agent to be credible:

- Perception: through this module, the agent will be able to perceive the state of the dialogue that develops with the user, which means the perception of all the data of the environment that will serve to characterize said context.
- Behavioral model: this model is in constant change because the emotions and intentions of the agent are established here.
- Internal state of the agent: in this component, the parameters such as personality and motivation are defined, which will be maintained during the whole process.
- Model of selection of verbal and nonverbal expressions: this module is integrated by the management of dialogue and a kinesic model, which are influenced by the behavioral model and the internal state of the agent. this kinesic model has two phases, the first phase is the emotional expression, that as its name indicates is the one that solves the emotional behavior, this first phase will have the responsibility of categorizing and choosing the nonverbal expressions of the agent, which means the gestural, postural and facial expressions. Subsequently, the second phase is the model of integration, which will have the responsibility of choosing in a definitive manner and combining the verbal and nonverbal selections to generate the animation. This module is where the contribution of this paper is centered.
- Animation: the last module will have as an input the verbal and nonverbal expressions previously selected and will be reported to the user.

4.1 Agent Personality

To model the personality expression and identify the facial and corporal expressions a video analysis of the interactions of the 4 personality type was done.For example, was observed that a person with a diplomatic type, is a person that smiles on a few occasions, which is one of the characteristics of an introverted person and, according to the personality type description it is a characteristic trait in them, however, in spite of showing seriousness it is an empathetic person that shows caring for others.

Regarding the corporal movements, they were almost null, there was no arm movement, even if there was movement from one side to the other. The video analysis of the interaction was used associated with the literature to establish the characteristic expressions of each of the personality types and thus obtain the characterization of the corpus according to the personality.

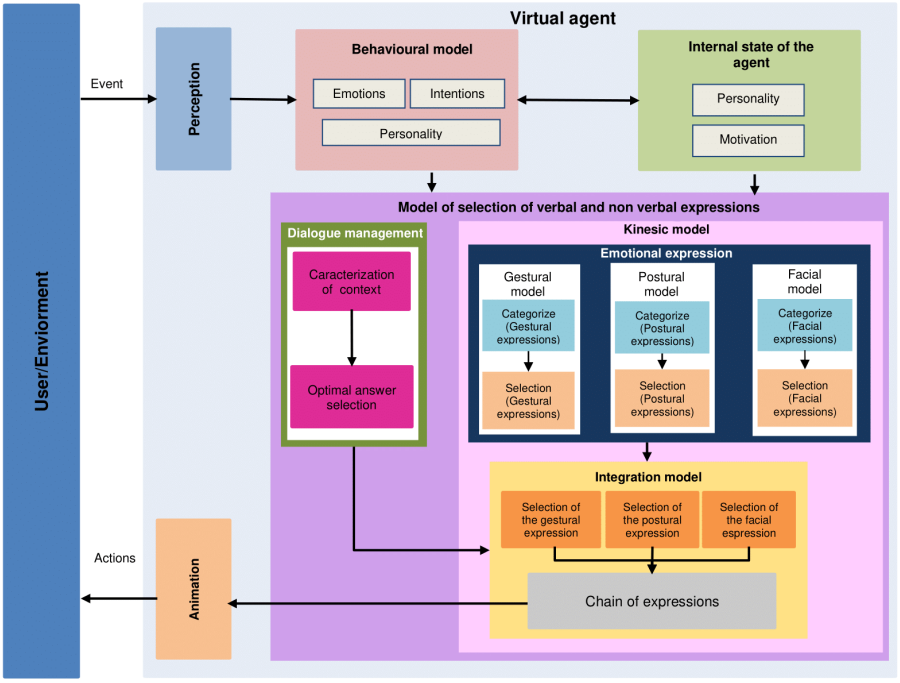

Fig. 1. Proposed architecture.

4.2 Corpus

The process of selection requires a corpus characterized by diverse criteria. The criteria used to characterize the corpus dialogue was the one applied by Delgado [3]. 33 criteria were proposed for this work to characterize the nonverbal expressions of the agent. Table 1 shows the criteria grouped into 4 categories: intention (21), emotion (4), context (4) and personality (4).

Table 1. The criteria proposed as characterized features of the corpus of facial and gestural expressions.

item	Criteria groups	Criteria
1	Intention	Sorry, Complaint, Sentimental expression, Attitudes, Affirmation,Statement, Explanation, Report, Suggestion, Petition, Question, Order, Mandate, Baptize, Inaugurate, Name, Dismiss, Promise, Oath, Offer,Threat.
2	Emotion	Joy, Fear, Sadness, Angry
3	Context	Social, Historic, Cultural, Educational
4	Personality	Analytical, Diplomatic, Sentinel, Explorer

5 Experimentation and Results

In this section, the CDA architecture is validated through a case study based on a dialogue between a student and tutor to compare the forecast of the expressions given by architecture and the expressions of people. From the 4 persons recorded, it was selected the person with the diplomatic personality type. This person had the role of the tutor. The components of the body evaluated were for facial expression, eye, eyebrow and mouth movements, and for gestural expressions, arms. Table 2 shows the results of 12 different interactions.

Where we can see a comparison between the expected expression that reinforces the selected sentences, that is, the expression used by the person with the role of tutor, versus the expression that was assigned by the ELECTRE III method. Were *Fn* is the variable that represents facial expressions and *Gn* represents gestural expressions. A complete accuracy of 67% was obtained, and the percentage achieved for the partial accuracy of facial expression was 75% and gesture expression was 75%.

Table 2. Comparison of the expressions resulting from the experimentation of 12 interactions with a diplomatic personality type.

Interaction	Expected expression	Assigned expression
1	F5, G5	F5, G5
2	F5, G5	F5, G5
3	F5, G5	F5, G5
4	F5, G5	F5, G5
5	F5, G5	F5, G5
6	F2, G5	F2, G5
7	**F5, G5**	**F3,G1**
8	**F2, G5**	**F5, G2**
9	F5, G5	F5, G5
10	F5, **G5**	F5, **G2**
11	**F1**, G5	**F2**,G5
12	F1, G5	F1, G5

As can be seen in the table, the identifiers of the expressions in bold were the inconsistencies between the expected expressions and the assigned expressions.

6 Conclusion and Future Work

This paper proposes the use of ELECTRE III, an optimization method for the integration of dialogue, and nonverbal expressions that use a corpus characterized by various criteria based on the influence of personality and emotions. In this work, the dialogue, facial expression, and gestural expressions are independent processes, using the same method ELECTRE III. In the future, seeking better results, nonverbal expression will also be determined by the phrase selected to reinforce verbal expression and evaluated with the four personality types.

Acknowledgments. The authors thanks CONACyT for the support granted through the Scholarship for Postgraduate Studies with CVU 1086661, and the project of cathedra with number 3058 and basic science 2017-2018 with number A1-S-11012.

References

1. Analytics Limited, NERIS: Tipos de personalidad (2011), https://www.16personalities.com/es/descripcion-de-los-tipos
2. Cuc, I.: Aplicación del método Electre III en la clasificación de clústeres de artesanías. INGE CUC 7(1), 16 (2011)
3. Delgado-Hernández, X.S., Morales-Rodriguez, M.L., Rangel-Valdez, N., Cruz-Reyes, L., Castro-Rivera, J.: Development of conversational Deliberative Agents Driven by Personality via Fuzzy Outranking Relations. International Journal of Fuzzy Systems 22(8), 2720–2734 (2020), http://link.springer.com/10.1007/s40815-020-00817-w
4. Díaz, J.: MBTI y tipos de personalidad: qué, cómo y por qué (2020), https://javierdisan.com/2020/01/22/mbti/
5. Fernández, C.A.I., Ayestarán, M.G., Cristóbal, J.C.G.: Definición de una metodología para el desarrollo de sistemas multiagente. p. 322
6. Gento, A.M., Redondo, A.: Comparación del método ELECTRE III y PROMETHEE II: Aplicación al caso de un automóvil p. 11
7. Gratch, J., Rickel, J., Andre, E., Cassell, J., Petajan, E., Badler, N.: Creating interactive virtual humans: some assembly required. IEEE Intelligent Systems 17(4), 54–63 (2002), conference Name: IEEE Intelligent Systems
8. G.Y. Lim, A.: Big Five Personality Traits (2020), https://www.simplypsychology.org/big-five-personality.html
9. Ishii, R., Ahuja, C., Nakano, Y.I., Morency, L.P.: Impact of Personality on Nonverbal Behavior Generation. In: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents. pp. 1–8. ACM, Virtual Event Scotland UK (2020), https://dl.acm.org/doi/10.1145/3383652.3423908
10. Reyes Nájera, P.D.: Integración de Motor de Selección de Frases basados en Superación para Desarrollo de Asistentes Web. Reporte de Residencias, Instituto Tecnológico de Ciudad Madero, Cd. Madero, Tamps. (2020)
11. Rosis, F.d., Pelachaud, C., Poggi, I., Carofiglio, V., Carolis, B.D.: From Greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. International Journal of Human-Computer Studies 59(1-2), 81–118 (2003), https://linkinghub.elsevier.com/retrieve/pii/S107158190300020X
12. Villarreal Hernández, J.A.: Framework para el diseño y animación de personajes renderizables en entornos web. Tesis de Licenciatura, Instituto Tecnológico de Ciudad Madero, Cd. Madero, Tamps. (2017)