Architecture of a Socio-Emotional Conversational Agent Capable of Identifying Intentions within a Socratic Dialogue

Héctor Rodríguez-Arteaga, Bárbara María Esther García-Morales,
María Lucila Morales-Rodríguez[*], Claudia Guadalupe Gómez-Santillán

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

{G16070821, D10070255, lucila.mr, claudia.gs}@cdmadero.tecnm.mx

Abstract. This article presents an architecture for developing a socio-emotional conversational agent to identify intentions within Socratic dialogues, using a combination of speech act theory and sentiment analysis. This proposal seeks to create a socio-emotional conversational agent capable of addressing the issue of how emotions associated with personality affect the recognition of intentions of expressions within the educational environment. To identify the appropriate intention, this approach considers the influence of personality, the polarity of the words used by the user, and the context in which the dialogue is situated. With this, it is expected to accurately identify communicative intentions, facilitating a natural and empathetic interaction between intelligent virtual agents and users, contributing to the field of artificial intelligence applied to education.

Keywords. Social-emotional conversational agents, speech act theory, sentiment analysis, Socratic dialogue, artificial intelligence.

1 Introduction

Nowadays, intelligent agents are used in various applications, such as virtual assistants (Siri, Alexa, Google Assistant), customer service chatbots, and personalized online tutors. These agents not only facilitate daily tasks but also promise to transform our interaction with technology radically. The integration of advanced concepts like Speech Act Theory (SAT) and Sentiment Analysis (SA) in these agents points towards an improvement in the ability to discern intentions and emotions, marking an advancement in creating dialogues that allow for more natural and empathetic interaction. According to [1, 5], Speech Act Theory (SAT) enables the processing and generation of language to reflect specific communicative intentions, which can be applied in the dialogue management models of conversational agents.

On the other hand, studies such as those by [4, 11] highlight the relevance of SA for capturing polarity and emotions in text, which would improve their understanding of the specific context established in a conversation. No research has been found in the literature review that explores how emotions specifically affect the determination of intentions in expressions within virtual educational environments.

This presents an opportunity to understand the intentions in complex conversational interactions that characterize human communication. The work of [3, 5] highlights the importance of integrating speech act theory into the design of agents. However, a deep exploration of how this approach can work with sentiment analysis in identifying intentions in educational contexts is still lacking. To address this issue, this article proposes an architecture for a socio-emotional conversational agent with the role of an educational coach.

This architecture combines speech act theory (SAT) and sentiment analysis (SA) to identify intentions under the influence of personality profiles and polarity in expressions within the context of educational coaching. This article presents several vital sections reviewing related works on socio-emotional conversational agents to establish the context and identify the need for architecture development. It then provides the background, where the fundamental concepts related to the techniques above are detailed.

It then describes how the architecture of the Agent with the role of virtual coach is composed and the functions of each module. Finally, the conclusions about the functionality of the Agent, its scope, and how it contributes to the advancement of socio-emotional conversational agents in educational environments are discussed.

2 Related Works

This section presents some related works in the area of socio-emotional conversational agent development that use speech act theory as a strategy to characterize phrases in intelligent virtual agents. In [9], relevance theory is adopted as the strategy for analyzing phrases, focusing on how human cognition and the relevance of the context assist in understanding communicative intention and its interpretation. Speech acts represent communicative intention and guide the actions of an intelligent agent in a dialogue.

The dialogue is organized into "communication games" and "phases"; these communication games help organize the speech acts used by speakers during the interaction, characterizing the phrases through perlocutionary acts that represent goals, illocutionary acts that define the strategy used in the game, and the conditions that can change it. In [3, 7], the use of SAT (Speech Act Theory) in selecting phrases from a specific corpus for virtual agents with personality is highlighted.

In [3], a manually characterized corpus of phrases was analyzed, including aspects of speech acts, emotions, personality, and other relevant attributes, using the MBTI model to simulate the personality of the virtual Agent. This allowed for fine-tuning in phrase selection, reflecting an accurate alignment with the Agent's personality traits.

In [7], previous work is improved by incorporating an automated phrase selection model for a virtual coach, which stands out for combining SAT with an inductive educational methodology. This model generates powerful questions inspired by Socratic dialectics, adapted to the coachee's personality through the MPBCD model (Personality Model Under a Decision Context). Despite the significant contributions of the mentioned works in applying Speech Act

Theory (SAT) and other theoretical frameworks for developing intelligent virtual agents, there is an issue in identifying intentions, as these studies assume that the intention behind the user's expressions is known.

However, no detailed mechanism for identifying them is available. This affects the agents' ability to understand and respond appropriately to the complexity and subtlety of dialogue since effective interaction crucially depends on recognizing not only what is explicitly said but also hidden intentions and communicative nuances.

Without understanding these subtle aspects, agents may respond inappropriately, making the conversation less natural and effective. This work incorporates a mechanism for identifying communicative intentions through sentiment analysis and the theory of speech acts.

This mechanism is situated within the architecture of a socioemotional conversational agent, which can identify the intentions in contrast to the user's personality profile and the context in which the conversation takes place, obtaining a deep and complete understanding of the context. With the integration of this mechanism, we intend to overcome the observed weaknesses by significantly improving the interaction between the conversational Agent and the user, allowing a more dynamic and empathic dialogue.

3 Background

To thoroughly understand the techniques that make up the virtual tutor capable of identifying intentions within a Socratic dialogue, it is essential to detail the fundamental concepts related to conversational agents, the theory of speech acts, the influence of personality models on socioemotional agents, Socratic dialectics in educational coaching processes and sentiment analysis.

3.1 Conversational Agents as Virtual Coaches

Conversational agents, also known as chatbots or virtual assistants, are artificial entities capable of emulating human behavior. As described by [5], these agents must be designed to interpret the intentions of conversations and understand the

interlocutors' questions and answers to decide what actions to take in real time. Natural language processing techniques, as well as machine learning methods and algorithms, are used to deal with that [9].

These agents incorporate the fundamental principles underlying most intelligent systems in their architecture. These principles include reactivity, the management of an internal state, the formulation of principles and goals, the capability for autonomy in decision-making, a predisposition towards sociability, and the ability of reasoning. These principles link the agents to the real world through beliefs, desires, and intentions.

The Agent analyzes its personality and context to define its identity and act according to the context in which it operates. The desires are closely connected with the goals that the Agent pursues. In this application case, they focus on formulating questions based on dialectics influenced by the Agent's personality [7].

3.2 Speech Act Theory for Phrase Characterization in Dialogue

Currently, conversational agents that employ speech act analysis and computational linguistics are available to deduce the meaning of words contained in expressions provided by the user [5].

Proposed by [1, 13], Speech Act Theory (SAT) is a branch of linguistic pragmatics focused on analyzing how people use language to perform actions in the world beyond merely communicating information. A mechanism to recognize the intention of phrases the user expresses is characterizing attributes belonging to speech acts [3, 6, 9].

According to [1], phrases can be characterized into three main types of acts (locutionary, illocutionary, and perlocutionary). Illocutionary speech acts play a fundamental role in determining the underlying meaning in a sentence, as pointed out by [13], illocutionary acts are divided into two categories: direct and indirect, and the main difference lies in that, in indirect acts, the locutionary and illocutionary acts do not overlap, while in direct acts, they do.

Furthermore, it's important to highlight that these illocutionary acts can be classified into five distinct categories: assertive, commissive, directives, declaratives, and expressive. According to [13], illocutionary acts are actions carried out by speakers when expressing words with communicative force. These acts allow the characterization of a sentence and identify the speaker's intention, which influences the receiver's interpretation.

Speech acts can characterize any type of dialogue, including a Socratic dialogue. In this type of dialogue, the phrases are direct acts, implying that the coachee's intention coincides with the coach's understanding. The illocutionary verbs in this type of dialogue can play locutionary or perlocutionary roles.

The powerful questions in this type of dialogue are classified as directive illocutionary acts aimed at interrogating [5]. This type of interaction in conversational agents can be linked with traits and personality types using personality models under decision contexts such as the MPBCD [7].

3.3 Personality Models for Virtual Coaches

To emulate a reflection of the coachee's personality within the educational coaching processes, it is necessary to resort to behavior modeling [7] to achieve effective realism. Therefore, it is essential to integrate elements such as emotions, personality traits, preferences, and motivations into the personality modeling functions of the virtual entity [12].

To achieve this, one can make use of the two major theories of personality on human behavior which are based on traits (FFM, OCEAN) and based on types (MBTI, KTS). These theories are grounded in psychological research dedicated to characterizing and describing individuals' personalities [2].

According to personality theory, the FMM and OCEAN models are based on personality traits. They are highlighted through five factors (openness, conscientiousness, extraversion, agreeableness, and neuroticism) representing personality [12]. These factors are derived from an analysis of questionnaires and adjectives that indicate the strengths and weaknesses of each trait through levels (high-low) [2].

The MBTI model is based on personality types, assigning individuals with abbreviated four-letter labels, thereby creating 16 different labels or
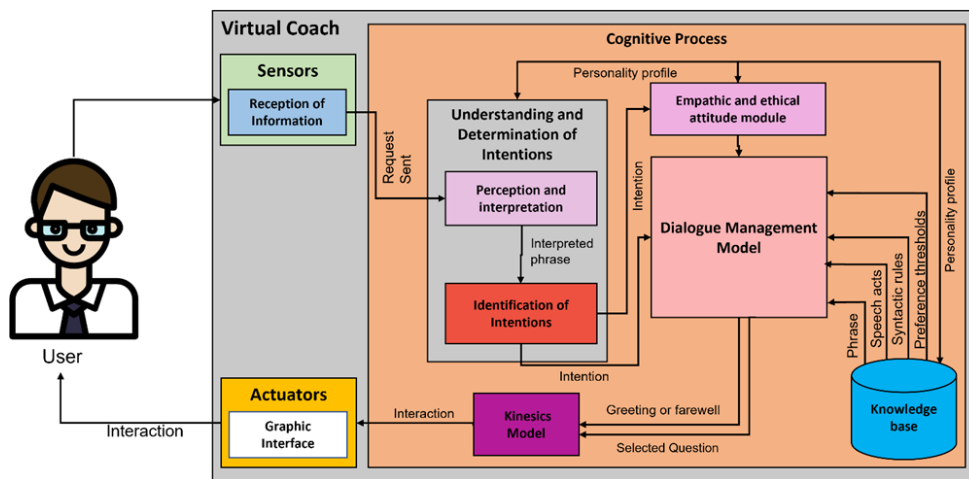
Fig. 1. Architecture of the conversational agent with the role of virtual coach

profiles. The labels are generated based on information obtained through a questionnaire. Each label describes the attitude of individuals in their interaction with their environment, which is manifested through dichotomies [2].

Another model based on personality types is the KTS model. That is a self-assessed questionnaire designed to help people understand their personality. KTS is based on MBTI, using the same combinations of items to establish the labels. The difference between this model and MBTI is the inclusion of temperament, role, and role variants [2].

The MPBCD (Personality Model Under a Decision Context) model proposes a relationship between the FFM, MBTI, and KTS models to integrate the traits and personality types approaches. Its primary function is to use information from individuals and generate a profile that illustrates their behavior in decision-making situations while examining the influence of personality on their preferences [2]. The MPBCD has been employed to develop agents based on Socratic dialectics [5].

3.4 Socratic Dialectics in Educational Coaching Processes

The Socratic dialectic, known as maieutic, involves conducting a strategic interrogation and counterpointing each answer with additional questions. Often, these questions are formulated neutrally to reach a truthful answer that synthesizes the partial truths of the previous responses [10].

In [6], the Socratic dialectic characterizes a dialogue that incorporates powerful questions and relies on two essential components: maieutics and irony. Its purpose is to foster self-knowledge through the cognitive opportunities that emerge from the responses generated in the context of Socratic maieutics.

According to the literature, the powerful questions posed in this context can be linked both to speech acts and to existing personality profiles, allowing the emulation of the coachee's personality in each interaction.

3.5 Sentiment Analysis (SA)

Personality-related characteristics can influence decision-making when generating new phrases. To analyze the emotions conveyed in each sentence spoken by the coachee, it is proposed to work with the relationship between the degree of polarity and the intentions through the technique of Sentiment Analysis (SA).

Sentiment Analysis (SA), known as Opinion Extraction, Opinion Mining, Sentiment Mining, or Subjective Analysis, is a computational technique for studying opinions, sentiments, and emotions expressed in texts. SA covers three processing

tasks: the binary classification of the attitude of a text, the self-assessment of a text according to its degree of polarity, and the identification of the aspects mentioned in a text, together with their associated sentiments [11].

The binary classification of attitude in a text is a natural language processing technique used to discern whether a text expresses a positive or negative attitude toward a particular topic [4]. The self-assessment of a text-based on polarity degrees involves classifying messages as positive, negative, or neutral.

This representation is the most basic and widely used in sentiment analysis [8]. The authors [4] comment that this task requires advanced natural language processing techniques, such as syntactic dependency analysis and the extraction of named entities.

These techniques are combined with sentiment classification algorithms specifically adapted for each identified aspect. Considering the tasks of SA, it is possible to use it in a Socratic dialogue to analyze the emotions associated with the intentions expressed through speech acts deeply.

4 Proposed Architecture for the Creation of the Conversational Agent

A detailed architecture has been conceptualized and designed to initiate the development of the conversational Agent capable of identifying intentions in a Socratic dialogue, as shown in Figure 1. This architecture, grounded in the BDI (Belief, Desire, Intention) agent model, is based on the architecture proposed by [5], which compiles all the essential elements for the construction of the virtual tutor.

This system consists of several interconnected modules designed to interpret and respond to user interactions in a virtual educational coaching environment. The interaction process starts with the Information Reception Module, which captures the user's phrase, identifying if the type of illocutionary act is expressive.

The response is generated from the knowledge base (phrases for greetings or farewells). In other cases, if this class is not expressive, it sends the phrase to the Understanding and Intention Determination Model, which contains the Perception and Interpretation Module. This module receives and detects the polarity of the phrase from a neutral perspective, considering the influence of the personality profile in the interpretation.

This allows for a more precise definition of the intensity of that polarity from the perspective of the Agent's personality, which becomes the factor for determining the illocutionary force of the phrase, resulting in the identification of intentions.

If more than one intention is detected, they are passed to the Intention Inference Module, which uses the ID3 (Iterative Dichotomiser 3) algorithm to choose an intention according to the context of the conversation, taking into account the value of the polarity associated with the description of the personality profile.

This obtained intention is sent to the speech acts module, thus continuing with the determination of a powerful question that will serve as a response to the coachee.

The Knowledge Base stores greeting phrases, grammatical rules, the description of personality profiles, and information criteria that guide the generation of responses. The dialogue management model comprises the greeting or farewell selection process and the Dialectic Modeling, which includes modules for analyzing speech acts and selecting questions that promote reflection, taking into account the personality and preferences of the virtual Agent.

The Phrase Characterization Process through Speech Acts evaluates the issue raised by the user and the coachee's personality to generate powerful questions that reflect the coachee's profile. The Question Selector analyzes the issue at a grammatical level to structure these questions, which are filtered within the Dialogue Management Module through the Deliberative Process. The questions are selected according to the preferences of the virtual tutor, using the ELECTRE III method to determine the most suitable solution.

5 Discussion and Future Work

In this article, an architecture for the development of a socio-emotional conversational agent that integrates Speech Act Theory (SAT) and

Sentiment Analysis (SA) was proposed through the identification of the polarity of phrases provided by the user, associating it with the personality profile and the characterization of illocutionary forces to precisely identify intentions in a Socratic dialogue, especially in the context of educational coaching.

Combining these two techniques provides a richer understanding of the interlocutor's expressions, allowing the virtual tutor to make better decisions in its interaction, enriching the interaction and feedback in the educational coaching process.

As a virtual tutor, this conversational Agent can significantly contribute to developing conversational agents and credible virtual entities in educational environments. This advancement not only has applications in educational coaching but can also extend to other areas of education and human-machine interaction.

This represents an important step towards creating more intelligent and empathetic virtual entities that can understand and adapt to their interlocutors' emotional and cognitive needs. In order to assess the performance of the socio-emotional conversational Agent, we will conduct quantitative and qualitative analyses.

In the quantitative approach, we will compare the intentions perceived by real people with different personality profiles to the results generated by the Agent that simulates these same profiles. Additionally, we will evaluate the coherence of the Agent's responses by using a dictionary to compare the definitions with the responses obtained.

The qualitative approach will include semi-structured interviews with real people. That allows us to explore how they perceive and justify identifying intentions and how this relates to their personality profile.

Acknowledgments

References

1. **Austin, J. L. (1975).** How to do things with words: second edition. Harvard University Press. books.google.com/books?id=u67VAAA AMAAJ.

2. **Castro-Rivera, J. A. (2017).** Modelado de personalidad en modelos preferenciales multicriterio a través de agentes virtuales inteligentes. Master's Thesis, Instituto Tecnológico de Ciudad Madero.

3. **Delgado-Hernández, X. (2021).** Aplicación de modelos de programación matemática a la selección de expresiones verbales en agentes virtuales socio-emocionales. Master's Thesis, Instituto Tecnológico de Ciudad Madero.

4. **Dubiau, L., Ale, J. M. (2013).** Análisis de sentimientos sobre un corpus en español: experimentación con un caso de estudio. Proceedings of the XIV Argentine Symposium on Artificial Intelligence, pp. 36–47.

5. **Garcia-Morales, B. M. E. (2021).** Modelo de selección de frases para un coach virtual que utiliza la técnica de la dialéctica. Master's Thesis, Instituto Tecnológico de Ciudad Madero.

6. **García-Morales, B. M. E, Morales-Rodríguez, M. L., Rangel-Valdez, N., García-Vite, P. M., González-Barbosa, J. J. (2022).** Caracterización de frases para un sistema conversacional inteligente en un entorno educativo virtual basado en las características de la dialéctica y los actos del habla. Research in Computing Science, Vol. 151, No. 6, pp. 105–118.

7. **García-Morales, B. M. E, Morales-Rodríguez, M. L., Valdez, N. R. (2023).** Conversational system based on Socratic dialectics in a virtual educational environment. Research in Computing Science, Vol. 152, No. 4, pp. 17–26.

8. **Gonzalvo, S. H. (2020).** Introducción a las técnicas de análisis de sentimientos. Bachelor's Thesis, Universidad de Granada

9. **Morales-Rodriguez, M. L. (2007).** Modèle d'interaction sociale pour des agents conversationnels animés. Application à la rééducation de patients cérébro-lésés

(Doctoral Thesis). Université Toulouse III: Paul Sabatier.

10. **Castillero-Mimenza, O. (2018).** Método Socrático: qué es y cómo se aplica en la psicología. Psicología clínica, Vol. 21, No. 43.

11. **Pang, B., Lee, L. (2008).** Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, Vol. 2, No. 1–2, pp. 1–135. DOI: 10.1561/1500000011.

12. **Pérez-Pinillos, D., Fernández, S., Borrajo, D. (2013).** Modeling motivations, personality traits and emotional states in deliberative agents based on automated planning. Communications in Computer and Information Science, Vol 271. DOI: 10.1007/978-3-642-29966-7_10.

13. **Searle, J. R. (1969).** Speech acts: An essay in the philosophy of language. Cambridge University Press. DOI: 10.1017/CBO9781139173438.