Argumentation based dialogue for Cultural Heritage presentations with Virtual Humans

Antonio Origlia[1,2], Martina Di Bratto[2,4], Maria Di Maro[1,2], Sabrina Mennella[2,3,4], and Francesco Cutugno[1,2]

[1] DIETI - University of Naples Federico II
[2] URBAN/ECO Research Center
[3] Dept. of Humanties - University of Catania
[4] Logogramma S.r.l
{antonio.origlia, martina.dibratto, maria.dimaro2, sabrina.mennella, cutugno}@unina.it

Abstract. In this work, we describe a framework for the development of Argumentative Conversational Recommender Systems for Cultural Heritage. Specifically, we describe our research about the management of Virtual Humans in museum-like 3D environments. We describe the resulting technological framework, the development methodology linked to it, and we summarise the results that we obtained with it. Also, we present the innovations that are currently being introduced to support more complex interactions for the task of artworks recommendation using a combination of different AI techniques for dialogue management.

Keywords: Virtual Humans · Graph databases · Conversational AI

1 Introduction

Nowadays, Conversational Recommender Systems (CoRS) are acquiring a fundamental role in information seeking and retrieval. Typically, the main task of such systems is to point users to potential items of interest [3]. Recommendation dialogues are characterised by two or more participants who disclose their preference and make recommendations that should satisfy the requirements retrieved during the communicative exchange. CoRS, in the same way, aim at finding or recommending the most relevant information based on textual or spoken dialogues [2]. Furthermore, CoRS can be seen as persuasive social actors since a recommendation can be persuasive when it attempts to change or affect people's mind or behaviour by employing various persuasive strategies [6]. A conversation where two or more interlocutors (humans or not) aim to resolve a conflict of opinion, can be considered as a form of persuasion dialogue leveraging on argumentation

[7]. For this reason, CoRS can be framed in the field of argumentation-based dialogue. This considers the problems arising from dialogues involving different agents and whose information are shared and distributed among them.

Although argumentation-based dialogue has long been studied, there is still no shared framework in the community. One work-in-progress attempt is represented by a structured graph-based modelling and analysis of such dialogues in argumentation-based dialogue systems, as presented in [5]. Argumentative Conversational Recommender Systems (A-CoRS) are typically designed to negotiate with human users the shared adoption of specific *beliefs* in the respective representations of the common ground. In their case, the goal of the system is to *persuade* the interlocutor that a specific item corresponds to their manifested interests. In the domain of cultural heritage, A-CoRS are of particular interest to capture user preferences about emotional experiences that are often hard to link to categorical representations. As a further difficulty, users may be partially unaware of the available categories.

Commercial recommender systems often present simple statistical arguments based on their inner workings using statements like "people who bought X also bought Y". In A-CoRS, conversely, conversational approaches guide users towards the selection of interesting items and, with the introduction of argumentation capabilities, have the potential to reinforce recommendations while making users aware of the features motivating the recommendation. In fact, A-CoRS must be capable of exploit the knowledge collected through the dialogue to propose an appropriate item in a claim justified by specific features. These features, that characterise an item, are retrieved through conversational and domain knowledge and are used to built the argument supporting the claim.

In this work, we summarise the research studies we conducted in the field of Virtual Humans as guides for Cultural Heritage from different aspects related to embodied dialogue management. Specifically, we describe a linguistically motivated framework for A-CoRS development using a) Machine Learning for perception and behaviour synthesis tasks, b) graph-based technology for Knowledge Representation and deductive reasoning, c) Behaviour Trees for task prioritisation and organisation and d) Bayesian Networks for decision making.

2 Embodied A-CoRS for Cultural Heritage

The work conducted in the framework of the national 2015 PRIN project Cultural Heritage Resources Orienting Multimodal Experiences (CHROME) produced the Framework for Advanced Natural Tools and Applications with Social Interactive Agents (FANTASIA) [4], which was developed to integrate a modern game engine (Unreal Engine) and its development toolkit with ease of access to AI capabilities. FANTASIA introduces a series of connectors, in the Unreal Engine, that allow to use AI services, it integrates Bayesian Networks and it provides access to a graph database. FANTASIA also supports the integration of graph-based reasoning and Bayesian decisions with the Unreal Engine AI framework built using Behaviour Trees. The combination of all these different

approaches to AI implementation in a single development experience provides the foundations for an harmonic formalisation of the different processes involved in managing A-CoRS and is deployed for cultural heritage presentations.

On this basis, research involving Virtual Humans for Cultural Heritage has used FANTASIA to focus on different communication areas, from gestures and camera movements generation to Argumentation-based Dialogue Management for recommendation. In this work, we will describe how such a system is deployed and how different AI methods have been applied to build a linguistically motivated recommendation dialogue management system for Cultural Heritage.

Concerning the management of the Virtual Human and its relationship with respect to the environment, we rely on the concept of semantic annotations of 3D models, as described in [1]. Semantic annotations are imported in applications developed in the Unreal Engine 5 and provide a connection between concepts expressed in texts and geometrical information provided by 3D models. During presentations, semantically annotated concepts in the texts can be retrieved in the 3D world and guide both pointing gestures by the avatar and camera movements to support orienting approaches during pre-visit experiences.

Concerning dialogue management, our approach is based on the use of a graph database to keep track of the dialogue and incrementally build a belief system (defined as a *belief graph*) connecting the dialogue evolution to the knowledge domain. The belief graph is, then used, to iteratively select a subset of the knowledge domain on which the system reasons to either select an item to recommend (exploit) or ask a question to improve the user model (explore). Relevant subgraphs extracted from the database are converted into a Bayesian Network to estimate the probability that an item (i.e. a painting) or its features (its creator, movement, style, etc...) will be of interest for the user. By selecting appropriate questions, the system attempts, at each step to either reduce the size of the domain to investigate or to reduce the entropy of the Bayesian Network. Once an item has been selected for recommendation, that is when the probability of it being appreciated is high and the entropy of the distribution is low, it is presented to the user along with the most relevant features that support the proposal. These emerge primarily from the dialogue itself, to promote the perception of personalised recommendation. Figure 1 shows the graph resulting from the interpretation of an utterence, by the user, expressing a positive opinion about Caravaggio. The expressed intent implies the belief that the speaker *likes* the entity connected to the *Subject* slot so, after interpreting the utterance, the system models starts *believing* that the subject likes the entity Caravaggio, as represented in the knowledge graph. This information can be used to select new sub-graphs on which to reason and to select relevant features to build argumentation-based recommendations.

References

1. Cera, V., Origlia, A., Cutugno, F., Campi, M.: Semantically annotated 3d material supporting the design of natural user interfaces for architectural heritage. In: Proc of the AVI-CH Workshop (2018)

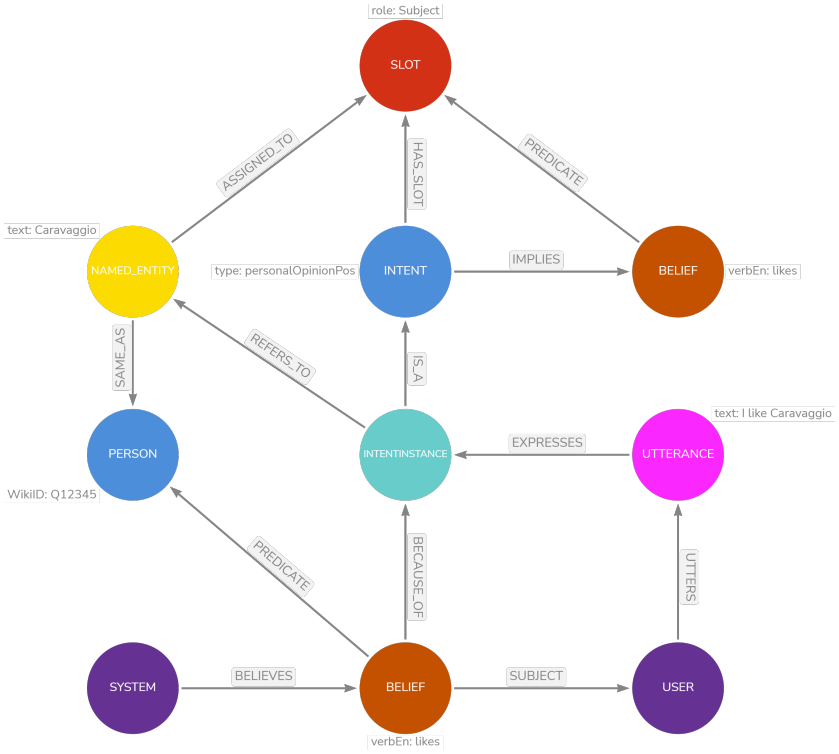

Fig. 1. The Belief Graph resulting from a user expression of a positive opinion

2. Fu, Z., Xian, Y., Zhang, Y., Zhang, Y.: Tutorial on conversational recommendation systems. In: Fourteenth ACM Conference on Recommender Systems. pp. 751–753 (2020)
3. Jannach, D., Manzoor, A., Cai, W., Chen, L.: A survey on conversational recommender systems. ACM Computing Surveys (CSUR) **54**(5), 1–36 (2021)
4. Origlia, A., Cutugno, F., Rodà, A., Cosi, P., Zmarich, C.: Fantasia: a framework for advanced natural tools and applications in social, interactive approaches. Multimedia Tools and Applications **78**(10), 13613–13648 (2019)
5. Origlia, A., Di Bratto, M., Di Maro, M., Mennella, S.: Developing embodied conversational agents in the unreal engine: The fantasia plugin. ACM Multimedia (Technical Demo) (2022)
6. Shi, W., Wang, X., Oh, Y.J., Zhang, J., Sahay, S., Yu, Z.: Effects of persuasive dialogues: testing bot identities and inquiry strategies. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–13 (2020)
7. Walton, D.N.: Logical dialogue-games and fallacies (1984)