Culture-Aware Cloud Services for Autonomous Conversation

Carmine Tommaso REcchiuto
DIBRIS
University of Genoa
via all'Opera Pia 13, Genoa, Italy
carmine.recchiuto@dibris.unige.it

Antonio Sgorbissa
DIBRIS
University of Genoa
via all'Opera Pia 13, Genoa, Italy
antonio.sgorbissa@unige.it

***Abstract*—Cultural competence - i.e., the capability to adapt verbal and non-verbal interaction to the user's cultural background - may be a key element for social robots to increase the user experience. However, designing and implementing culturally competent social robots is a complex task, given that advanced conversational skills are required. In this context, Cloud services may be useful for helping robots in generating appropriate interaction patterns in a culture-aware manner. In this work, we present the design and the implementation of the CARESSES Cloud, a set of robotic services aimed at endowing robots with cultural competence in verbal interaction, and we offer preliminary insights about their integration with robots in real scenarios.**

***Index Terms*—Autonomous Conversation, Social Robots, Human-Robot Interaction, Culture-Aware Technology**

I. Introduction

In the robotic domain, research on cultural factors has mainly focused on non-verbal aspects such as facial expressions [1] [2], greeting gestures [3], and interpersonal distance [4], which have been subjects of specific cross-cultural studies involving robots and virtual agents. As a general outcome, the results of these works suggest that cultural aspects may affect the robot's (or virtual agent's) likeability, acceptance, and persuasiveness, in the sense that people tend to prefer an artificial agent that conforms to the social norms of their own culture, both in the verbal and non-verbal behaviour. However, although underlying the importance of blending cultural factors into the Social Robotics domain, all the aforementioned approaches integrated a very small set of features that distinguish cultures from each other, and little work has been reported on how to build robots that can be easily adapted to the cultural identity of the user.

The CARESSES project has represented the first attempt to implement a robot that adapts its behavior according to the cultural identity of the person with whom it interacts, in terms of actions, actions' parameters (e.g., social distance, speech volume) [5], and dialogue patterns. In the context of the project, the evaluation of whether and how a robot using the CARESSES dialoguing framework is perceived as culturally competent has been carried out in a six-month experimental campaign, involving older persons belonging to different cultural groups, their informal caregivers, and the humanoid robot Pepper [6] [7].

Recently, the system has been refactored as Cloud services, with the final aim of a general dialoguing framework for culture-aware social robots. This abstract briefly summarizes the key concepts of the system, also discussing its integration with robots and artificial agents in real scenarios.

II. Cloud Services for verbal interaction

A. Conversational Framework

CARESSES has been conceived as a knowledge-driven, conversational framework for culturally competent robots, which relies on two core elements:

I An Ontology for storing all concepts of relevance, cultural information and statistics, person-specific information and preferences;

II An algorithm for building culture-aware dialogue patterns, relying on I.

More in detail, the core of the framework is a Description Logics Ontology, a formal representation of objects, concepts, and other entities, assumed to exist in some area of interest, and the relations that hold among them [8]. Knowledge about concepts and their mutual relations are stored in the terminological box (TBox) of the Ontology, while knowledge that is specific to instances belonging to the domain is stored in the assertional box (ABox). From an implementation perspective, the TBox is composed of classes and properties, which include Data Properties, relating instances of a class to literal data (e.g., strings, numbers), and Object Properties, relating instances of a class to other instances; the ABox stores instances of classes and instances of properties.

In the specific contexts, the Ontology is ideally divided in three layers:

- A layer that stores the terminology (TBox) required to represent all the information that may play a role in a culture-aware conversation, but without being specific for a given culture: beliefs, values, habits, preferences, objects, norms, among the others.
- A layer that stores the assertions (ABox-I), required to represent culture-specific information.
- A layer that stores the assertions (ABox-II) required to represent the unique cultural identity, preferences, social and physical environment of the assisted person.

Instances of this layer may encode the actual user's attitude about concepts, and they may be updated during the verbal interaction by collecting his/her feedback, or specific knowledge about the user (e.g., name, town of residence) explicitly added during setup.

The dialogue tree is built starting from the Ontology structure: instances of the Ontology (i.e. of the ABox-I layer) are seen as a *conversation topic*, i.e., a node of the tree. The relation between topics is borrowed from the structure of the Ontology: specifically, Object Properties and the hierarchical relationships among instances, are analyzed to define the branches of the dialogue tree. On the other side, sentences or chunk of sentences possibly used during verbal interaction, and the probability with which each node may be selected by the system are encoded as Data Properties, in the ABox-I and ABox-II layers.

Based on the dialogue tree, the policies for knowledge-driven conversation can be briefly summarized as follows [9]. Each time a user sentence is acquired:

1 A keyword-based Language Processing algorithm is applied to check if the sentence may trigger one of the topics in the tree.
2 If no topics are triggered, the conversation follows one of the branches of the tree, depending on the probabilities of each node.
3 Whatever node has been chosen, the system:
   (i) proposes some of the corresponding sentences;
   (ii) acquires the user's feedback that can be used to update the Ontology with user-specific instances (ABox-II layer) and/or determine the next node to move to.

B. Cloud Services

The conversational framework has been recently refactored to offer a portfolio of Cloud services, having the possibility of processing cultural information encoded in the Ontology using remote computing services (including massively-parallel computation), investigating collective learning strategies for a run-time expansion of the knowledge base, and integrating the framework with a wide range of robotic devices, such as humanoid robots, table-top robots, smartphone applications, and voice assistants.

Some aspects should be here remarked:

- The culture-aware Cloud services may easily work in conjunction with existing Language Processing systems (onboard or anyway integrated with the robot), with the final aim of complementing them. In other words, the developed Cloud services are not meant at substituting the conversational capabilities that the robot is already equipped with, but rather provide a backup plan whenever onboard language processing techniques are not able to interpret what the user is saying. Usually, in commercial products, the backup plan foresees a connection to Wikipedia or other websites: the proposed Cloud services provide a culturally competent alternative, allowing any robot or system to talk with the user about thousand of different topics.
- The system produces a mixed-initiative verbal interaction [9]: both the user and the robot can take the initiative in leading the conversation. This is aimed at providing dialogues that overcome the typical *command-only barrier* [10], which is the main limitation of home assistants and most robotic systems, typically expecting a command from the user (*Robot, tell me the weather report!*) which is then reactively executed. On the opposite, the proposed Cloud services are able to produce rich goal- and knowledge-driven dialogue patterns, making the robot able to execute different kinds of speech acts [10].

On the implementation perspective, at first, the client program running on the robot is supposed to send all information about the user: background culture, language, and personal data, such as the first name and the family name. During the interaction, sentences pronounced by the users should be handled locally by a parallel mechanism, if they are directly aimed at starting an activiy (e.g., *Play some music*, *Show me the weather report*), or sent to the Cloud in case the robot does not know how to reply. In this case, the user's sentence is analyzed on the Cloud side, and an appropriate conversation topic is chosen: The Cloud takes the initiative of the conversation: replying with a sentence that shows the robot's cultural awareness, or asking a question related to the user's attitude with respect to the current conversation topic. After the user's reply, the Ontology may be updated depending on the feedback received, and the dialogue may possibly continue by exploring the branches of the dialogue tree more deeply, always taking into account the user's cultural identity, through a sequence of educated guesses.

Please notice that the user may also take back the initiative at any moment, by saying something that triggers a different conversation topic on the Cloud, or by directly asking for a robot's task.

III. Integration with Social Robots and Artificial Agents

The culture-aware services have been already integrated with four robotic platforms and devices: the Pepper humanoid robot, realized by Softbank Robotics[1], Pillo, a robotic pill dispenser with social abilities developed by Pillo Health[2], Professor Einstein, a small social robot developed by Hanson Robotics[3], and a custom Android smartphone app (Figure 1). Examples of the interaction of users with ths sytem may be seen in the following videos[4,5,6]. Concerning the integration of the system with different robots, research work has been focused on the following areas:

[1]https://www.softbankrobotics.com/us/pepper
[2]https://pillohealth.com
[3]https://www.hansonrobotics.com/professor-einstein
[4]https://www.youtube.com/watch?v=0at2oZt_rag
[5]https://www.youtube.com/watch?v=Zfr1YUX6BJw
[6]https://www.youtube.com/watch?v=lhnVs-V2Ssg

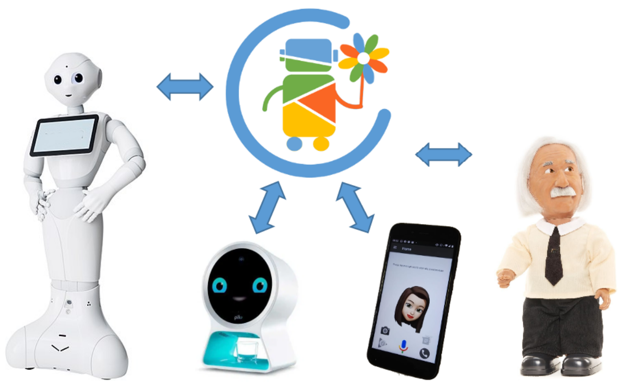

Fig. 1. The CARESSES Cloud services may be integrated with multiple robots and devices. Currently, the humanoid robots Pepper and Professor Einstein, the robotic pill dispenser Pillo and a custom Android application have been connected to the conversational framework.

- By integrating Pepper with the system, a six-month experimental campaign, involving older persons belonging to different cultural groups, and their informal caregivers, has been carried out [6] [7], with some interesting results, which are the subject of an ongoing publication. Overall, the robot relying on the CARESSES conversational framework was perceived by users as culturally competent, and the interaction with the robot improved the mental health of participants, in particular in terms of emotional wellbeing.
- The custom smartphone application connected with the proposed Cloud services has allowed for investigating the effects of the embodiment of conversational agents [11], showing that embodiment may have a positive effect on some aspects of the interaction, such as user satisfaction and perceived speech quality, while interacting with a robot is perceived as something more difficult to learn, and physical embodiment does not seem to impact the user's engagement.
- The integration of the system with the robotic pill dispenser Pillo has required the development of an additional Representational State Transfer (REST) server, which may act as a bridge between the developed Cloud services and artificial agents [12]. Preliminary data concerning the integration of the system with the pill dispenser Pillo are the subject of an ongoing publication.
- Finally, a thorough analysis of the architecture's feasibility in terms of communication and data processing delays has been performed [13], showing that communication latencies are satisfactory for verbal interaction in the Social Robotics domain, being also negligible with respect to widespread Speech-To-Text Cloud Services.

IV. Conclusion

This extended abstract briefly summarizes the recent work carried out for the implementation of culture-aware Cloud services for autonomous conversation, conceived for social robots and artificial agents. The framework has already proven to be reliable, being integrated with different robots and devices, and it has been used in a 6-months experimental campaign involving older people belonging to different cultural groups and the humanoid robot Pepper, giving positive insights about the potentialities of these novel Cloud services.

References

[1] C. Chen, L. B. Hensel, Y. Duan, R. A. Ince, O. G. Garrod, J. Beskow, R. E. Jack, and P. G. Schyns, "Equipping social robots with culturally-sensitive facial expressions of emotion using data-driven methods," in *2019 14th IEEE Int. Conf. on Automatic Face & Gesture Recognition (FG 2019)*. IEEE, 2019, pp. 1–8.

[2] T. Koda and Z. Ruttkay, "Eloquence of eyes and mouth of virtual agents: cultural study of facial expression perception," *AI & society*, vol. 32, no. 1, pp. 17–24, 2017.

[3] G. Trovato, M. Zecca, M. Do, Ö. Terlemez, M. Kuramochi, A. Waibel, T. Asfour, and A. Takanishi, "A novel greeting selection system for a culture-adaptive humanoid robot," *Int. Journal of Advanced Robotic Systems*, vol. 12, no. 4, p. 34, 2015.

[4] G. Eresha, M. Häring, B. Endrass, E. André, and M. Obaid, "Investigating the influence of culture on proxemic behaviors for humanoid robots," in *2013 IEEE RO-MAN*. IEEE, 2013, pp. 430–435.

[5] A. A. Khaliq, U. Köckemann, F. Pecora, A. Saffiotti, B. Bruno, C. T. Recchiuto, A. Sgorbissa, H.-D. Bui, and N. Y. Chong, "Culturally aware planning and execution of robot actions," in *2018 IEEE/RSJ Int. Conf. on Intel. Robots and Systems (IROS)*. IEEE, 2018, pp. 326–332.

[6] C. T. Recchiuto, C. Papadopoulos, T. Hill, N. Castro, B. Bruno, I. Papadopoulos, and A. Sgorbissa, "Designing an experimental and a reference robot to test and evaluate the impact of cultural competence in socially assistive robotics," in *2019 28th Int. Conf. on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2019.

[7] C. Papadopoulos, T. Hill, L. Battistuzzi, N. Castro, A. Nigath, G. Randhawa, L. Merton, S. Kanoria, H. Kamide, N.-Y. Chong *et al.*, "The caresses study protocol: testing and evaluating culturally competent socially assistive robots among older adults residing in long term care homes through a controlled experimental trial," *Archives of Public Health*, vol. 78, no. 1, pp. 1–10, 2020.

[8] N. Guarino, D. Oberle, and S. Staab, "What is an ontology?" in *Handbook on ontologies*. Springer, 2009, pp. 1–17.

[9] C. Recchuto, L. Gava, L. Grassi, A. Grillo, M. Lagomarsino, D. Lanza, Z. Liu, C. Papadopoulos, I. Papadopoulos, A. Scalmato *et al.*, "Cloud services for culture aware conversation: Socially assistive robots and virtual assistants," in *2020 17th Int. Conf. on Ubiquitous Robots (UR)*. IEEE, 2020, pp. 270–277.

[10] N. Mavridis, "A review of verbal and non-verbal human–robot interactive communication," *Robotics and Autonomous Systems*, vol. 63, pp. 22–35, 2015.

[11] L. Gava, L. Grassi, M. Lagomarsino, C. Recchiuto, and A. Sgorbissa, "Physical embodiment of conversational social robots," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, pp. 456–463.

[12] C. Pautasso, "Restful web services: principles, patterns, emerging technologies," in *Web Services Foundations*. Springer, 2014, pp. 31–51.

[13] C. T. Recchiuto and A. Sgorbissa, "A feasibility study of culture-aware cloud services for conversational robots," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6559–6566, 2020.