BRILLO: A Robotic Architecture for Personalised Long-lasting Interactions in a Bartending Domain

Alessandra Rossi, Mariacarla Staffa, Antonio Origlia, Maria di Maro, Silvia Rossi

University of Naples Federico II
Naples, Italy
{alessandra.rossi, mariacarla.staffa, antonio.origlia, maria.dimaro2, silvia.rossi}@unina.it

ABSTRACT

The use of robots for the automation of the supply of food and beverages is a commercially attractive and modern application of robotic technologies. Such innovative technologies are deemed helping to renew the image of a service and, in this way, stimulate people's curiosity. The novelty effect linked to the experience of a new technology, however, has a very limited duration in time, and it is not suitable for guaranteeing user loyalty. Consequently, the need for continuous renewal to keep the commercial proposal attractive becomes very expensive. In this paper, we present the architecture of a new project, called Bartending Robot for Interactive Long Lasting Operations (BRILLO), which aims to create a long-lasting operational robotic system that is able to have personalised multi-user interactions and to work as a bartender by performing different tasks according to the users' requests and preferences.

KEYWORDS

Personalised HRI, social robotics, multi-modal interaction, long-lasting collaborations

ACM Reference Format:
Alessandra Rossi, Mariacarla Staffa, Antonio Origlia, Maria di Maro, Silvia Rossi . 2021. BRILLO: A Robotic Architecture for Personalised Long-lasting Interactions in a Bartending Domain. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21 Companion), March 8–11, 2021, Boulder, CO, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3434074.3447206

1 INTRODUCTION

BRILLO (Bartending Robot for Interactive Long Lasting Operations) is a three years national project with the overall goal of creating an autonomous robotic system that is capable of working as a bartender and interacting naturally with customers.

In recent years, the field of service robotics is rapidly growing bringing the need of making such robots endowed with complex perceptual capabilities to achieve a proper context and human awareness [9] and, consequently, a natural interaction with humans. Humanoids service robots have been employed in a variety of application contexts to perform tasks in public environments (e.g., museums, restaurants, call-center, fairs, and exhibitions). The bartending domain is particularly challenging since it combines the ability of complex (eventually bi-manual) manipulation of objects, and the necessity of actively engaging and interacting with customers. The integration of a system that combines a "system of record" (e.g., analysing and storing past interactions, preferences to optimise sells with respect to the specific users) and a "systems of engagement" (e.g., aiming at facilitating and enhancing the experience via a personalised and natural interaction) plays a central role in customer relationship management. Hence, the BRILLO project aims to create a robotic platform to be deployed in a real environment that is capable of accomplishing the tasks related to the management of a bar counter in its multiple aspects of the interaction, manipulation and personalisation of the services offered. Topics related to the modeling of preferences and interactions between the user and the robotic bartender, in terms of customisation of this interaction, for an effective and lasting use of the services offered, will then be addressed.

In the literature, different projects addressed the robotic bartending domain from an automation perspective by developing complex manipulation capabilities, such as the automatic green tea serving robot [6], or a robot that is able to serve a large number of customers [7]. Other researches mainly focused on modeling and recognising the social interactions in a bartending domain, as for example the James project [3, 4] whereas a bartending robot was design to exhibit social behaviours in response to costumers' states, intentions, and affects as recognised from low-level sensor data during the interaction. The results of these projects showed that the interactions with a socially intelligent robot mixing task execution, dialogue and social interaction were somewhat more efficient. The project BRILLO aims at combining these efforts to obtain an easy to use robotic system that can be deployed in commercial environments, but also aims to rely on history of interactions and user models to personalise such interaction and making the use of the system attractive on long-term base.

In this paper, we provide a description of the BRILLO project's framework, where three interactive interfaces are considered (i.e., a bartending robot, a totem, a waiter robot) acting as a centralised system designed around costumers' needs, preferences and mood.

2 MULTI-MODAL AND MULTI-USERS SCENARIOS

A typical scenario of a BRILLO service point includes a human user and three interacting systems, i.e. a totem kiosk, the bartender robot, and a waiter robot.

HRI '21 Companion, March 8–11, 2021, Boulder, CO, USA

ACM ISBN 978-1-4503-8290-8/21/03…$15.00
https://doi.org/10.1145/3434074.3447206

Upon their arrival to the service point, users are welcomed by a totem kiosk where they are automatically recognised via bio-metric face recognition if they are already registered users, otherwise, a quick registration of their data in the system is carried out. Then, the users will be asked to order a drink (for example, shakes or cocktails). Once the order has been completed, users can decide to pick up their orders at the bar station, served by the two arms humanoid robot (i.e. the bartender robot), or at the table, served by the waiter robot.

At the bar station, the bartending robot detects the presence of the registered user, it proceeds with the reception and it continues the profiling and recommendation operations based on the customer's preferences rising from previous interaction experiences and orders. During the entire interaction phase, the bartending robot is able to entertain the customer by adapting its social response to the specific user' profile and needs. During the orders' preparation carried out by the robot, the customer can place additional orders or change their orders on the fly, for example by adding alcohol to the shake or changing alcohol ounces on previously chosen cocktails. Any performed changes to the orders affect the robot task execution in real-time. The bartending robot receives the new or modified order and, accordingly proceeds to the correct preparation. The bardending robot is also able to managing multiple orders and users, by opportunely scheduling and adapting its behaviors.

3 SYSTEM DESCRIPTION

The BRILLO system is a collection of software and hardware modules that allows the three involved interacting systems to perceive their users, learn about their preferences, behaviours, attitude and moods, and dynamically adapt to them for a long term and effective interaction.

Figure 1 presents an overview of the architecture and relevant implemented modules. This architecture is based on the same principles of common layered architectures (such as [1, 2, 5]). This project aims to develop a robotic system that is able to work as bartender, completing multiple and in parallel tasks while adapting its behaviours to its users' needs and personal differences, in terms of preferences, moods and traits.

According to the current literature that confirms the importance of verbal and non-verbal communication following a multimodal approach and supplemented by social skills for the improvement of service quality and loyalty, BRILLO's architecture will follow a user-centered approach. In particular, the design of the interfaces for accessing the functions of the bartending robot will be divided into a functional part, dedicated to identifying the services on which to evaluate performance, and a social part, dedicated to specifying the behaviours to be implemented in relation to the different user profiles.

Our system is a four-layers control architecture where each component is linked to the Beliefs, Desires, Intentions (BDI) model architectures [2, 5].

3.1 Hardware layer

The Hardware layer consists in the robots' hardware (i.e., for the bartender robot and the waiter robot), and in the graphical user interfaces (i.e., for the displays of the totem kiosk and tablet).

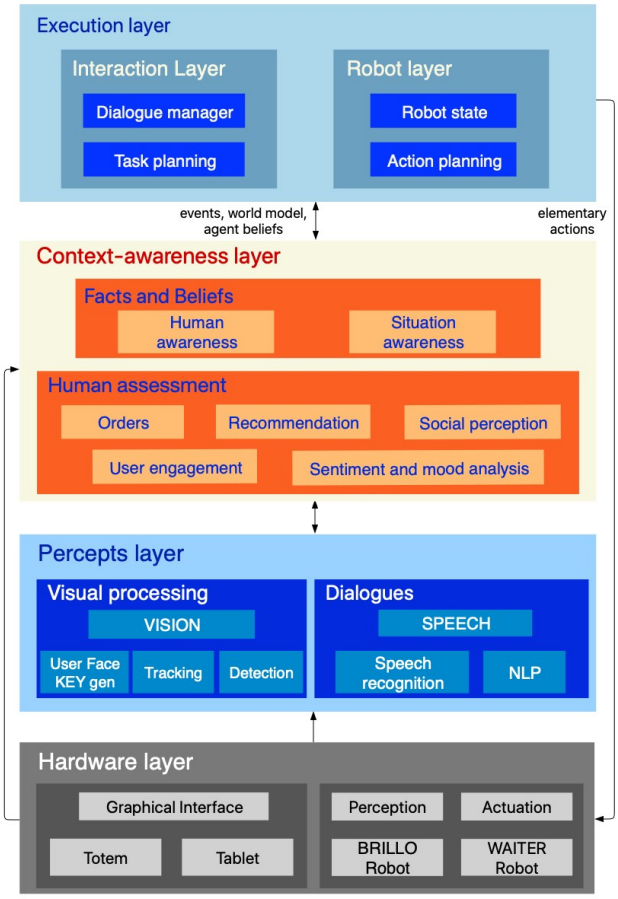

Figure 1: Overview of the BRILLO project's architecture. It is composed of four principal layers: 1) Hardware, 2) Percepts, 3) Context-awareness, and 4) Execution. Each layer has a collection of different modules.

3.1.1 Robots. An anthropomorphic structure is adopted for the bartender robot, since it induces users to ascribe social skills to the robot by so positively affecting humans trustworthiness and willingness to interact with the robot [10]. The bartender robot has a minimalist structure with two robotics arms on the torso and, in the future, it will have a screen as head (see Figure 2).

The waiter robot has a serving tray and wheels to autonomously and safely navigate between the tables of the service point (i.e., the bar counter).

The bartender and the waiter robots are equipped with a large variety of sensors that enable them to perceive and assess the environment, the customers, and the activities of the agents involved in the interaction.

3.1.2 Graphical User Interfaces. The graphical interfaces of the totem and the tablet on the tables provide the users with several functionalities aimed to make and manage the orders of drinks that will be, then, prepared by the bartender robot.

The totem is able to detect the presence of a user upon their arrival, and if they are returning customers, the system recognises

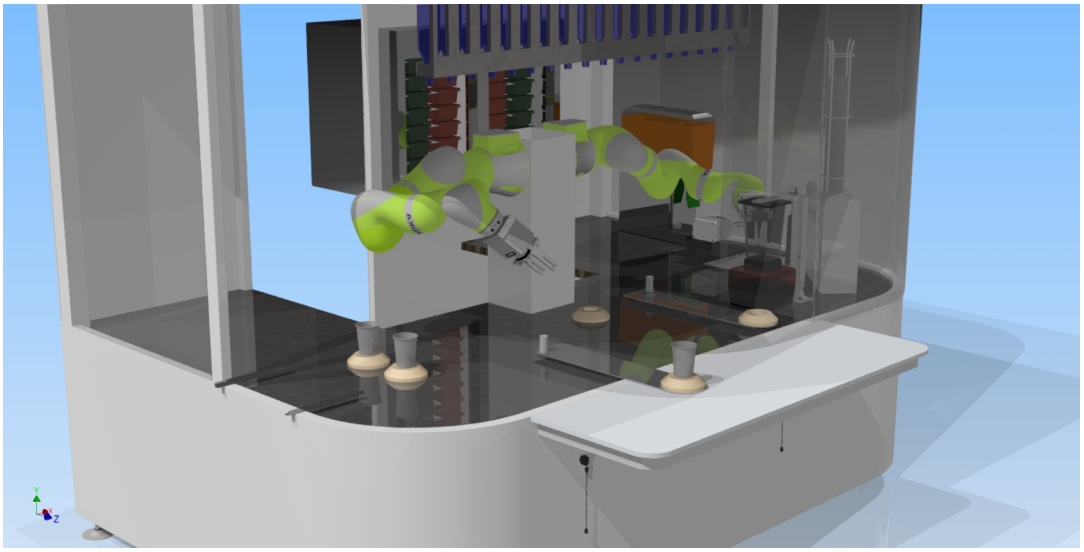

Figure 2: 3D representation of the bartending robot at the service point.

their facial features. The graphical interface of the totem allows the users to complete several tasks, such as registering them at their first use, making and managing an order. More precisely, the tablet on the tables allows the users to manage new or existing orders made at the totem.

3.2 Percepts layer

The Percepts layer manages the multimodal interaction capabilities of the robots which consists in the information obtained by the modules *Visual processing* and *Dialogues*.

3.2.1 Visual processing. 4-kinect cameras are used to detect and track users' faces in real-time both at the totem and at the bar premises. The vision module also derives costumers's focus of attention via their body and head pose, facial expressions recognition and gaze by suggesting the robotic platform to adapt it social behaviour accordingly.

3.2.2 Dialogues. The bartending robot is able to engage users in natural dialogue. It will use an automatic speech recognition system to obtain the speech transcription, and a machine learning approach for the Natural Language Understanding (NLP) for engaging the user through customised dialogues. Currently, the dialogue system supports interactions in Italian language. The robot is able to adapt the dialogues according to the personas (described in the next section) whom it is interacting with.

3.3 Context-awareness layer

This module builds the system's awareness of the environment, the context and the robot during the interactions. The communication between these modules is bidirectional to build the system's knowledge in a dynamic and more adaptive way.

The system is be able to assess users' profiles with the aim to adapt the robot's behaviours and interactive capabilities to the users' social expectations, preferences and level of engagement. For example, the bartending robot might adapt its behaviours according to the following personas:

- Worker on a lunch break: this user wants to access quickly to a service for getting a break from the fatigue of work. The system provides an interaction meant to physically and mentally relax them.
- Curious person: this user likes new stories to tell their friends. The system engages them in a variety of interaction topics (news, energy values of products, product popularity statistics, small talks, product history).
- Regular user: this user wants to access a service that brings entertainment or in between their activities. The system builds the interaction by personalising it based on previous interests and preferences.
- Working group: these users want to having working conversations within their group (business guests, colleagues) in an environment more relaxed than their office. The system provides them with a quick service that eases the interaction between the users (e.g. an user-initiative interaction not to disturb them).
- Group of friends: these users want to have some qualitative time with the other components of the group. The system provides a one-to-many interaction with general and non-individual dialogues. For example, it allows efficient ordering of multiple products (e.g. "How many apple smoothies?" "How many with pears?").
- Group of family members: as the group of friends, these users want to spend quality time with the loved ones. It is likely that they will be a group of people of different age, therefore the system adapts the interactions accordingly. For example, it takes orderings from adults and entertains children.

This module also serves the recommendation system according to previous interactions with the users, in terms of the users' previous orders and preferences of robot social behaviours.

Moreover, the *Context-awareness layer* builds the dialogues and robots' behaviours for engaging the users by analysing the users' moods and good predisposition to interact with the robot via natural language processing, text analysis and biometrics.

Finally, the modules in this layer provides the knowledge base for the system's higher level that is in charge of planning and executing tasks (*Execution layer*).

The knowledge based acquired by this module drives the robot in deciding the most appropriate social response inferred by the context. This allows the robot to have a social intelligence, which is needed to provide a positive relationship between users and robots [8].

3.4 Execution layer

The high level goals and plans, the instantiating, execution and monitoring of the actions of the robot are managed by the *Execution layer*. The communications between these modules and the *Context-awareness layer* are bidirectional to allow the robot to dynamically plan actions (single or multiple), and behaviours according to the user and the situational context. For example, the robot might engage the user in a casual conversation while preparing their drink for fostering a more enjoyable interaction. This module collects the users' *desires* of some actions from the dialogue manager, and decides whether the action planner and controller should transform it into a *goal* and, then, generates *intentions*.

4 FUTURE WORKS / CONCLUSIONS

In this paper, we briefly illustrated the current state of the robotic system, including hardware and software modules, used for the BRILLO project. The project aims to deploy a service robot that is able to engage human users in a dynamic and unsupervised interaction. The BRILLO's robots are not only able to autonomously complete multiple tasks in a efficient way, but they are also able to personalise the experience and adapt their behaviours (such as context and frequency of dialogues, movements and poses) to their users' moods, preferences and personal traits by showing a social intelligence. The presented architecture serves as a preliminary step towards an experimental prototype for in-lab studies and, then, the deployment in unstructured environments.

5 ACKNOWLEDGMENTS

This work has been supported by Italian PON I&C 2014-2020 within the BRILLO research project "Bartending Robot for Interactive Long Lasting Operations", no. F/190066/01-02/X44.

REFERENCES

[1] E. Burattini, A. Finzi, S. Rossi, and M. Staffa. 2012. Attentional human-robot interaction in simple manipulation tasks. *HRI'12 - Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction* (2012), 129–130. https://doi.org/10.1145/2157689.2157719

[2] Coşar, Serhan, Fernández-Carmona, Manuel, Roxana Agrigoroaie, Jordi Pages, Ferland, François, Feng Zhao, Shigang Yue, Nicola Bellotto, and Adriana Tapus. 2020. ENRICHME: Perception and Interaction of an Assistive Robot for the Elderly at Home. *International Journal of Social Robotics* 12 (07 2020). https://doi.org/10.1007/s12369-019-00614-y

[3] Mary Ellen Foster, Andre Gaschler, and Manuel Giuliani. 2013. How Can i Help You': Comparing Engagement Classification Strategies for a Robot Bartender. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction* (Sydney, Australia) *(ICMI '13)*. Association for Computing Machinery, New York, NY, USA, 255–262.

[4] Mary Ellen Foster, Simon Keizer, and Oliver Lemon. 2014. Towards Action Selection under Uncertainty for a Socially Aware Robot Bartender. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*. Association for Computing Machinery, New York, NY, USA, 158–159.

[5] Séverin Lemaignan, Mathieu Warnier, E. Akin Sisbot, Aurélie Clodic, and Rachid Alami. 2017. Artificial cognition for social human–robot interaction: An implementation. *Artificial Intelligence* 247 (2017), 45 – 69. https://doi.org/10.1016/j.artint.2016.07.002 Special Issue on AI and Robotics.

[6] T. Masuda and D. Misaki. 2005. Development of Japanese green tea serving robot "T-Bartender". In *IEEE International Conference Mechatronics and Automation, 2005*, Vol. 2. 1069–1074 Vol. 2. https://doi.org/10.1109/ICMA.2005.1626700

[7] H. Q. T. Ngo, T. P. Nguyen, and H. Nguyen. 2018. Investigation on Barbot to Serve Human in Public Space. In *2018 4th International Conference on Green Technology and Sustainable Development (GTSD)*. 300–305. https://doi.org/10.1109/GTSD.2018.8595543

[8] A. Rossi, K. Dautenhahn, K. Lee Koay, and M. L. Walters. 2020. How Social Robots Influence People's Trust in Critical Situations. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1020–1025. https://doi.org/10.1109/RO-MAN47096.2020.9223471

[9] Silvia Rossi, Alessandra Rossi, and Kerstin Dautenhahn. 2020. The Secret Life of Robots: Perspectives and Challenges for Robot's Behaviours During Non-interactive Tasks. *International Journal of Social Robotics* (2020). https://doi.org/10.1007/s12369-020-00650-z

[10] M. Staffa and S. Rossi. 2016. Recommender interfaces: The more human-like, the more humans like. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9979 LNAI (2016), 200–210. https://doi.org/10.1007/978-3-319-47437-3_20