Gentlebots - Team Description Paper

Francisco J. Rodríguez Lera Francisco Martín Rico
Vicente Matellán Olivera
https://gentlebots.gsyc.urjc.es/

November 27, 2023

Abstract. The present report shows the capabilities of the Spanish team named Gentlebots and it is submitted for qualification in the Open Platform League in the 2024 RoboCup@Home international competition to be held in Eindhoven , Netherlands. Our lines of research focus on the intersection between Mobile Robotics and Social Robotics, aiming to develop intelligent behaviors for robots that acceptably help people. The team proposes a cognitive architecture that uses advanced concepts of knowledge representation, subsumption, and classical planning. In this architecture, we have integrated 3D perception based on Deep Learning, navigation using medium/long-term maps, visual attention, and dialogue. In addition to applying these developments in projects with the industry, we have applied them in competitions such as the European Robotics League (ERL). All our software is publicly available on GitHub1, and it works with ROS and ROS 2. In the last 4 years, we have published about twenty scientific articles in journals indexed in ISI-Thompson and in congresses on robotic competitions, all of them with experimental backgrounds based on RoboCup and ERL @home leagues

1 Introduction

Our team is composed of two universities with a great tradition of participating in robotics competitions: 6 RoboCup (2 in 4-legged, 3 SPL, 2 OPL, 1 SSPL), 5 German Open (4 4-legged and 1 SPL), 1 Dutch Open (4-legged), 1 Rome Cup (SPL) and 1 Latin American Open (4-legged), 4 RoCKIn[1] (camp and competition @Home), 4 ERL (@Home) and 1 SciRoc.

The scientific production related to these participations is summarized in 5 doctoral theses, 20 articles in high-impact journals, and more than 40 articles in national and international congresses. It is an excellent performance for a team that has never exceeded 10 members at a time.

[1] http://rockinrobotchallenge.eu/

Since 2012, our team has always developed open-source software using ROS and ROS 2 [1].Our developments have always been released as ROS packages and made available to the RoboCup community.

Our primary focus lies in the pursuit of scientific advancement and production. In alignment with this commitment, we participate in the project Reto-Hogar[2], funded by the Ministry of Economy and Competitiveness of the Spain government under grant TIN2016-76515-R. This project contemplates the inclusion of a robot in a home to help people with acquired brain damage. Furthermore, the team was the winner in the elevator challenge of SciRoC 2019, UK[3]

The team was also involved in organizing @Home competitions. Francisco J. Rodríguez was a member of the Organization Committee (OC) of RoboCup@Home. Vicente Matellán is one of the organizers of the ERL[4], making available to this competition an officially approved domestic environment of the league within his laboratory in Leon. He has already organized two editions of this competition. Francisco Martín Rico, is one of the main contributors to ROS community in packages such as NAV2 or PlanSys 2.

After facing obstacles at the 2023 RoboCup@Home, impacting our performance unfavorably, our team is determined to redeem ourselves in the next competition. Assessing our 2023 performance has given us crucial insights for improvement. We've thoroughly reviewed strategies, identified tech gaps, and refined methodologies to address weaknesses. With a strong history in prestigious competitions and a track record of innovation, we're confident that these adjustments will substantially boost our performance.

2 Innovative technology and scientific contribution

2.1 Cognitive Architecture for social robots

Figure 1 shows the design on layers of our architecture. It is a concentric layered design, with a transversal component called *Knowledge Graph*. The Knowledge Graph stores the internal and external knowledge of the robot, and it is accessible from any layer.

Tier 1 and 2 mainly use symbolic information. Tier 3 and 4 use subsymbolic information. The central part of the architecture, at Tier 2, is a symbolic planner based on PDDL. Using PDDL, we define what types, symbolic predicates, and actions can be used to solve a problem in a domain. This planner has a knowledge base, accessible from other levels, that contains the instances and predicates of the current problem.

The evolution of our architecture has progressed from the Behavior-based Iterative Component Architecture (BICA) [2] to MERLIN [3], and from ROSPlan

[2] http://www.rovit.ua.es/retogar-retorno-al-hogar-sistema-de-mejora-de-la-autonomia-de-personas-con-dano-cerebral-adquirido-y-dependientes-en-su-integracion-en-la-sociedad/

[3] https://sciroc.eu/e04-take-the-elevator/

[4] https://sites.google.com/site/erlsrleonsept2017/home

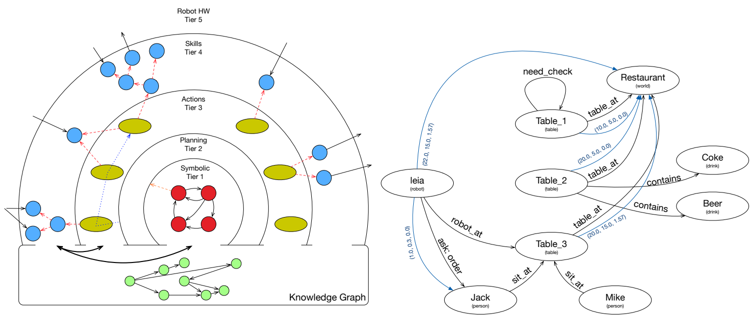


Fig. 1: Layered cognitive architecture.

[4] to PlanSys2 [5] at the planning level. Currently, our efforts are centered on enhancing the architecture's modularity and versatility to align with emerging technologies such as large language models and data-driven approaches.

Tier 1 starts the instances and predicates of the problem to be solved. This level contains hierarchical state machines that define the modes and behaviors of the robot at a high level. When a state machine at Tier 1 establishes a goal, the planner at **Tier 2** creates a plan using the content of his knowledge base. The plan is made up of a sequence of domain actions. The planner delivers the actions at Tier 3 one at a time. Each time an action indicates that it has been completed successfully, the next one is delivered until the plan is finished. **Tier 3** contains the implementation of the actions defined in the PDDL domain. Previously we defined this level and the following levels as subsymbolic. This level is the bridge between both paradigms. **Tier 4** contains skills that can be activated from actions. The skills can be reused from any of the actions. This level includes perceptual, attention, dialogue, and manipulation modules, among others. **Knowledge Graph** stores the information relevant to the operation of the robot. We have designed this shared representation of data to disengage some components of others, especially between different layers and we also have a tool for storing symbolic information easily using design patterns [6]. An action in Tier 3 uses the result of computing a skill in Tier 4 by reading it from the knowledge graph. Tier 1 can also use the symbolic information contained in the graph.

2.2 Perception

Our perception system is composed of a 3D object detector a visual attention system as well as a person and gesture detector.

Deep Learning detector The images taken by the robot's RGBD camera are the input of a convolutional neural network, which detects objects in the image. We have used Darknet-ROS[5] for this detection, and the output is a list

[5] https://github.com/leggedrobotics/darknet_ros

of bounding boxes labeled with the class and probability of the object detected. We have developed Darknet3D[6]. This component receives the 3D bounding boxes and a point cloud, generating 3D bounding boxes. In Figure 4 there is an example of a 3D bounding box for a person.

When a skill or an action requires receiving this information, instantiate a Darknet3D client, passing it a configuration. The configuration includes in which classes they are relevant, which is the working frame, the range of its coordinates, and size, among others. The client filters and transforms the information received and made it available for use. When an object is detected, the agreement is to add it to the knowledge graph with a `sees` edge. It also adds an edge TF that indicates its position.

Visual attention When the robot wants to detect objects or track an element, it uses the visual attention system shown in Figure 2. The system deployed here[7] is controlled using the knowledge graph, adding edges `want_see` from the robot to a node. For each type of target node, there is a specialized care component. If the node is of a type table, it is understood that you want to search for items on the table. If it is an area, the space that delimits this area is searched. On the other hand, if it is a person or an object, it is assumed already perceived and it is understood that it is desired to track it. Each specialized care component provides 3D points that the robot must attend. The attention system visits in round robin every 3D point that is provided.

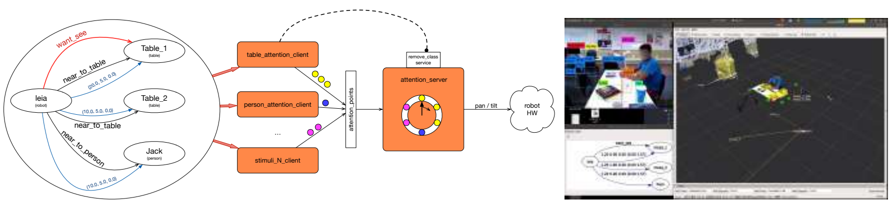
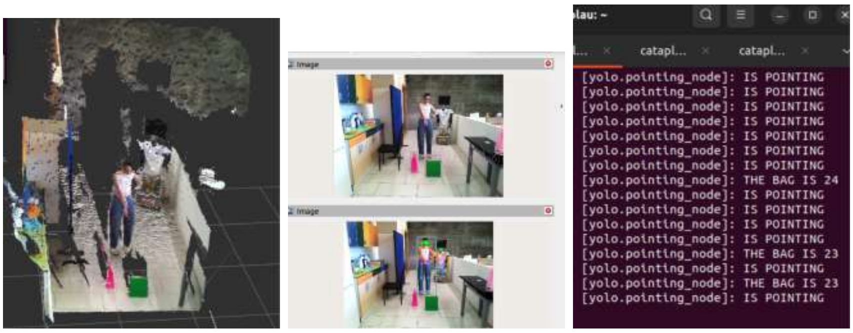


Fig. 2: Proxemic creation and path generation.

Person and gesture recognition The primary focus of the competition lies in executing tasks within home-like settings, often involving human-robot interactions in dynamic environments. This is made possible through the utilization of the Ultralytics library, enabling the extraction of a person's skeletal structure. These key points are subsequently aligned with the point cloud data the RGBD camera information. This framework also facilitates the tracking of identified individuals. By utilizing projected keypoints, heuristic algorithms can then calculate gestures; for instance, when the hip and knee points are in close proximity, indicating that the person is likely seated.

6 https://github.com/IntelligentRoboticsLabs/gb_visual_detection_3d

In the 'Carry my Luggage' task, this framework plays a pivotal role in identifying the suitable bag. By employing a specialized ray tracing algorithm for each extended arm, when the ray intersects with an object resembling a bag in proximity, it is able to determine the correct bag as seen at Figure 3



(a) person pointing to the bag (b) extracting person skeleton and keypoints (c) system logs, infering the correct bag

Fig. 3: Person and gesture system used for Carry My Luggage task

2.3 Social Navigation

Starting from the standard ROS navigation [8]and NAV 2 [9] packages , we have modified the map and location system to have a system that, based on a map of walls and doors, adds the permanent objects that the robot perceives during its operation. This is called long-term navigation [10], and is based on the fusion of several maps that reflect the presence of obstacles over time.

Our approach builds a static map starting from the construction plans of a building. A long-term map is started from the static map, and updated when adding and removing furniture, or when doors are opened or closed. A short-term map represents dynamic obstacles such as people. This approach is appropriate for fast deployment and long-term operations in office or domestic environments, able to adapt to changes in the environment.

We combine our long-term mapping approach with a social navigation algorithm for robots that is acceptable to people. Robots will detect both the personal areas of humans and their areas of activity, to carry out their tasks, generating navigation routes that have less impact on human activities. The main novelty of this approach is that the robot will perceive the moods of people, to adjust the size of proxemic areas. Figure 4 shows how we create a proxemic zone (right) and the route calculated (left), where the blue path respect this proxemic area.

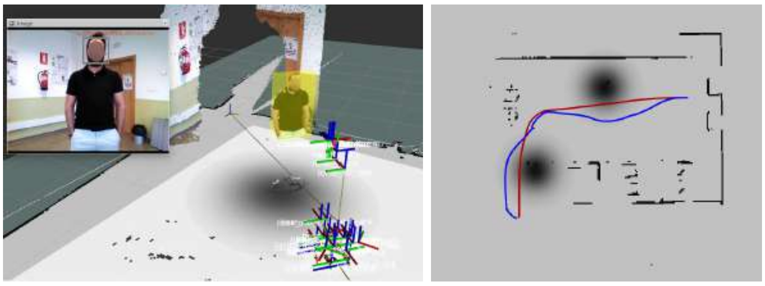


Fig. 4: Proxemic creation and path generation.

2.4 Human-robot Interaction

Our robot performs multi-modal communication with humans. Use three simultaneous forms of communication to avoid communication errors that may block the robot at some point.

The main form of communication is through a natural voice dialogue. Our system uses DialogFlow and Google Speech. DialogFlow has a web interface where, as in a chatbot, instances, contexts, and the elements necessary for a sophisticated dialogue are encoded. It is the only component of our software that requires access to a service outside the robot. This is why we have added the other communication systems, to avoid relying on wireless connectivity.

The dialogue can be in two ways: Unidirectional and Bidirectional. When a skill or action adds a `say: {text}` arc, the dialogue system causes the robot to pronounce the text and deletes the arc. On the other hand,, when a skill or action adds a `say: {context}` arc, the dialogue system causes the robot to initiate a dialogue with a person until the answer is obtained. This answer is in an arc in the opposite direction `answer: {response}`.

The second form of communication is the touch screen. There is shown, in addition to the state of the robot, the text that the robot is saying. The third form of communication is through phrases and words coded in a deck of QR symbols.

2.5 Explainability and Security

Since our goal is to use our systems in real environments, the team develops new software modules with security mechanisms [11] that prevent intrusions, attacks and not allowed manipulations. Besides, the team is also involved in developing eXplainability solutions for Autonomous Robots (XAR) [12].

3 Re-usability of the system for other research groups

There are two main repositories associated with Gentlebots Team. Robotics solutions are mainly developed in C++ or Python for ROS or ROS 2.

- URJC: https://github.com/IntelligentRoboticsLabs/
- ULE: https://github.com/uleroboticsgroup

4 Applicability in the real world

These are the videos that you can watch to see the application of team solutions in real-world scenarios.

- Qualification video for RoboCup@Home: https://www.youtube.com/watch?v=G82iv95paFs
- Winners of RoboCup@Home Education Challege (Guimaraes 2022): https://youtu.be/RXjXPLBWDHk

Related to general research:

- URJC: https://www.youtube.com/@fmr1c0/videos
- ULE: https://www.youtube.com/@roboticaunileon4750/videos

5 Conclusions and future work

This TDP described the research developments of our team that would suit the 2024 RoboCup@home competition. It also presents the team trajectory in @Home competitions, work done and published. It is expected to successfully face the challenges for the OPL. The team is committed to open source, and all our work is available before, during and after the competition.

Acknowledgements

The team is partially funded by:

- DMARCE - PID2021-126592OB-C21: Gobierno de España. Proyectos de Generación de Conocimiento 2021 Ministerio de Ciencia e Inovación, Agencia Estatal de Investigación.
- Horizon Europe CoreSense Project associated to HORIZON-CL4- 2021-DIGITAL-EMERGING-01-11
- UPF4ROS2 under AIPlan4EU H2020 project (https://aiplan4eu-project.eu) that is funded by the European Commission under grant agreement number 101016442.

References

1. Morgan Quigley, Ken Conley, Brian P. Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y. Ng. Ros: an open-source roboft operating system. In *ICRA Workshop on Open Source Software*, 2009.

2. Francisco Martín, Carlos Agüero, José María Cañas, and Eduardo Perdices. Humanoid soccer player design. *Robot Soccer*, pages 67–100, 2010.
3. Miguel Á González-Santamarta, Francisco J Rodríguez-Lera, Claudia Álvarez-Aparicio, Ángel M Guerrero-Higueras, and Camino Fernández-Llamas. Merlin a cognitive architecture for service robots. *Applied Sciences*, 10(17):5989, 2020.
4. Michael Cashmore, Maria Fox, Derek Long, Daniele Magazzeni, Bram Ridder, Arnau Carrera, Narcis Palomeras, Natalia Hurtos, and Marc Carreras. Rosplan: Planning in the robot operating system. In *Proceedings of the international conference on automated planning and scheduling*, volume 25, pages 333–341, 2015.
5. Francisco Martín, Jonatan Ginés Clavero, Vicente Matellán, and Francisco J Rodríguez. Plansys2: A planning system framework for ros2. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 9742–9749. IEEE, 2021.
6. Miguel Á González-Santamarta, Francisco J Rodríguez-Lera, Francisco Martín, Camino Fernández, and Vicente Matellán. Kant: A tool for grounding and knowledge management. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pages 452–461. Springer, 2022.
7. Francisco Martín, Jonatan Ginés, Francisco J Rodríguez-Lera, Angel M Guerrero-Higueras, and Vicente Matellan Olivera. Client-server approach for managing visual attention, integrated in a cognitive architecture for a social robot. *Frontiers in Neurorobotics*, 15:630386, 2021.
8. Eitan Marder-Eppstein, Eric Berger, Tully Foote, Brian Gerkey, and Kurt Konolige. The office marathon: Robust navigation in an indoor office environment. In *International Conference on Robotics and Automation*, 2010.
9. Steve Macenski, Francisco Martín, Ruffin White, and Jonatan Ginés Clavero. The marathon 2: A navigation system. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2718–2725. IEEE, 2020.
10. Jonathan Ginés, Francisco Martín, Vicente Matellán, Francisco J. Lera, and Jesús Balsa. Dynamics maps for long-term autonomy. In *IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, 2017.
11. Francisco J. Rodríguez Lera, Camino Fernández Llamas, Ángel Manuel Guerrero, and Vicente Matellán. Cyber-security of robotics and autonomous systems: Privacy and safety. *Robotics - Legal, Ethical and Socioeconomic Impacts*, 2017.
12. Francisco Javier Rodríguez-Lera, Miguel Ángel González-Santamarta, Ángel Manuel Guerrero-Higueras, Francisco Martín-Rico, and Vicente Matellán-Olivera. Towards explainability in robotics: A performance analysis of a cloud accountability system. *Expert Systems*, page e13004, 2022.
13. Jordi Pages, Luca Marchionni, and Francesco Ferro. Tiago: the modular robot that adapts to different research needs. In *International workshop on robot modularity, IROS*, volume 290, 2016.
14. Francisco Martín, Francisco J Rodríguez Lera, Jonatan Ginés, and Vicente Matellán. Evolution of a cognitive architecture for social robots: Integrating behaviors and symbolic knowledge. *Applied Sciences*, 10(17):6067, 2020.
15. Miguel Á. González-Santamarta, Francisco J. Rodríguez-Lera, Vicente Matellán-Olivera, and Camino Fernández-Llamas. Yasmin: Yet another state machine. In Danilo Tardioli, Vicente Matellán, Guillermo Heredia, Manuel F. Silva, and Lino Marques, editors, *ROBOT2022: Fifth Iberian Robotics Conference*, pages 528–539, Cham, 2023. Springer International Publishing.
16. Amanda Coles, Andrew Coles, Maria Fox, and Derek Long. Forward-chaining partial-order planning. In *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 20, pages 42–49, 2010.

Annex: TIAGo Description

Hardware description

In the 2023 edition of the RoboCup@home, we will participate using the TIAGo Iron robot from PAL Robotics [13]. This robot is composed of:

- A holonomic mobile base with a maximum speed of 1m / s.
- Battery autonomy 8 - 10h.
- A 180º laser with a range of 5.60m.
- A torso that can rise.
- A 15-inch touch screen.
- Head with 2 DoF (pan/tilt).
- RGB-D camera.
- 1x 5W audio speaker.
- 2-microphone array with stereo output 50-8000 Hz.
- An Intel i7 Haswell 16 GB and 512 GB computer.
- An NVidia Jetson TX2 coupled to the robot.
- 2x Gigabit, 802.11 n/ac 2x2 Dual-Band Wi-Fi and Bluetooth 4.0.

Fig. 5: Robot TIAGo

Robot's Software Description

- Platform: Ubuntu 22.04
- Navigation: Nav2
- Face recognition: OpenCv Based
- Speech recognition: Dialog Flow
- Speech generation: espeak
- Object recognition: YOLO
- Control:
- Planning: PlanSys 2 / UPF;
- Behavioral FSM
- Arms control and two-hand coordination: MoveIt 2.

External Devices

TIAGo robot relies on the following external hardware:

- Laptop (on top of robot's back, connected via eth)

Cloud Services

TIAGo connects the following cloud services:

- Dialog: Google dialogflow.

Robot software and hardware specification sheet