UTBots 2025 Team Description Paper

João A. Fabro Marlon O. Vaz Bruno B. Silva
João H. C. Pires Gustavo F. Armênio David Segalle
Enzo H. Gaio Victor B. Errera Larissa B. L. da Silva
Luiz G. S. Greca Matheus S. P. de Lima Bruno E. Zenatti
Gabriel Spadafora Cristoffer N. Spring Gabriel T. O. Shinzato
Luiza R. R. I. Ocampo André S. de Oliveira

August 28, 2026

Abstract. Being a participant in the Brazilian league since its first edition in 2014, team UTBots achieved 3rd place in both 2014 and 2022. Several research projects have been developed over the years on a range of subjects, most involving robot perception and control. Some of our research has now advanced in complexity and abstraction, tackling problems such as Task Orchestration models with Behavior Trees and Household Environment Ontology.

Some of our novel approaches include developments in Natural Language Processing for better dialogue comprehension and decision-making in social scenarios, where the robot depends on verbal information for its behavior. As mentioned earlier, our approach to Task Orchestration involves the use of Behavior Trees to improve the robot's capacity to deal with reactivity in @Home tasks. HESTIA (Household Efficient Service Task-Integrated Assistant) is our new robot with a mobile base adapted from a commercial hoverboard, and Theseus is our open-source project for low-cost manipulators in the league.

Our contributions to the league include organizing the Latin American Robotics Competition in 2017, the VI Brazilian Workshop on Service Robotics, and the V Brazilian Workshop on Humanoid Robotics, as well as participation in the Brazilian league's technical committee. All of our software and some of our hardware—like the Theseus project—are documented and available on our GitHub. Additionally, our ROS tutorials in Portuguese have been instrumental in educating researchers and team members within the Brazilian league.

1 Introduction

The UTBots@Home[1] team has been dedicated to the ongoing development of service robots designed to assist humans in household tasks since 2014. Some no-

[1] HomePage of the Team: `https://laser.dainf.ct.utfpr.edu.br/doku.php?id=utbots_at_home`

table achievements in local ranks include the 3rd place in the first Latin American edition of the @Home challenge, in 2014, and, most recently, the 3rd place in the 2022 edition of the Latin American league. Driven by the challenges presented by the Robocup@Home competition, our efforts have advanced in both complexity and abstraction. Initially focused on basic control of robotic systems, we now integrate and orchestrate our hardware and software systems at a level of abstraction suited to the ontology of a household environment in line with the competition's characteristics.

Over the years, the competition has served as a platform for numerous academic research endeavors at UTFPR. Our research involves, but is not limited to, navigation with fuzzy logic [1][2], computational intelligence frameworks [3], human-robot emotional interface [4], fast methods for dataset capture for object detection and manipulation [5], methods for estimating an object's representative 3D point in space [6], and new methods for people recognition [7].

In the domain of software, our current work is primarily focused on Natural Language Understanding, Object Manipulation, and Task Orchestration techniques with behavior trees.

Additionally, several upgrades to our hardware helped the team achieve more control over our robot. HESTIA (Household Efficient Service Task-Integrated Assistant) is UTBots new robot developed from scratch, with a mobile base that integrates motors and controllers from a commercial hoverboard, inspired by the SHARK approach [8]. Furthermore, HESTIA has a new home-made and low-cost 5 DOF robotic manipulator with brushless motors and open-source position drivers, coupled with magnetic encoders – our open-source hardware project Theseus. Considering the many years of developing new capabilities, an opportunity arose to develop software to synchronize with a complex task orchestration system.

2 Scientific Research and Innovation

2.1 Task Orchestration with Behaviour Trees

In order to give the programmer high level access to the robot's capabilities, allowing them to code complex tasks with ease, two orchestration systems were developed and tested within the context of CBR2024 @Home. Firstly, State Machines were found to be a capable solution for level one and two tasks at CBR, this, however, is insufficient when trying to develop a system capable of solving any task with an arbitrary amount of reactivity required, such as Stickler for the Rules.

For tasks similar to Stickler for the Rules, which require reactivity, Behavior Trees and their added reactivity offer a better solution. One of the biggest advantages of Behavior Trees are not on the system itself. The development of ROS2's nav2 with behavior trees rendered several tools that facilitate the use of custom trees. In practice this means that with a well integrated system we are able to solve completely new tasks by only adjusting YAML and XML files, no source code required. A solution to a @Home task is found in Figure 3.

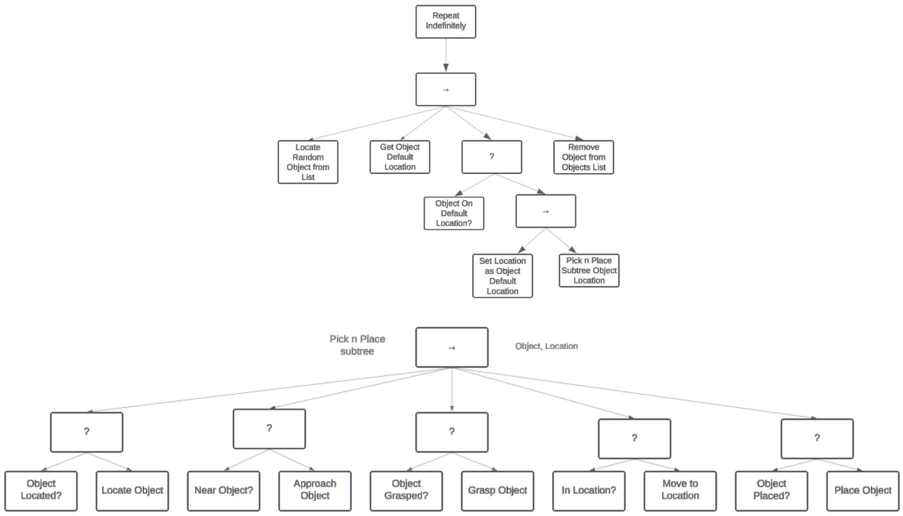

Fig. 1. Simplified Behavior Tree designed to solve the take out the garbage task.

2.2 Voice Command Interpretation and Context-Adaptable Aswering

To solve the Speech Recognition and QA task, we improved the STT (Speech-to-text) ant TTS (Text-to-Speech) pipelines. For improving the STT accuracy, we integrated the RNNoise [9] noise filter to the pipeline. The filter uses a series of digital signal processing techniques and Deep Learning to split the human speech from varying noises, and its sourced to be built in all platforms. It reveled great performance under the CBR 2024 atHome Speech-Recognition task. The next step is using the OpenAI's whisper-v3-turbo-q8 [10] to improve the speech translation to text. The bigger model now distilled and quantized offers more accuracy with less latency and less memory usage.

For dialog comprehension, we implemented a context-based RAG to query and retrieve context to feed the LLM (Large-Language-Model) so it can answer questions as the context provided. The question is embedded so we can query the k-Nearest Neighbors and ask the Meta-AI's Llama-3.1 [11] model to answer the question. To not surpass the 10 second deadline after the operator asks the question, we adjusted the prompt so it could generate the answers faster.

While LLMs are capable of providing an understanding of the semantics of a Voice Command regardless of syntax with a certain guiding, we also explored alternatives to Human-Robot dialog that require less computing resources. With RASA's Natural Language Understanding server [12], intention extraction with short-term memory helps guiding the dialog with correct answers and comprehension of Voice Commands for robot planning, specially in tasks such as GPSR and Enhanced GPSR.

2.3 Pointcloud Filtering for Object 3D Point Estimation

Presented in the 2023 Latin American Robotics Symposium, this work [6] proposed and tested methods for filtering depth points inside an object's detected bounding box, aiming for low-cost in computing resources for 3D point estimation. Considering the limitations of a rectangular bounding box to represent complex-shape objects, it is expected that some depth values of pixels – gathered with a RGB-D camera – inside of the bounding box are not related to the object's semantic, but from other elements in the scene. Filtering the points becomes even more relevant when using low-resolution RGB-D sensors such as Microsoft Kinect v1. Therefore, three methods were evaluated, a median-based one, an outlier removal approach and a novel method called Spiral Filtering. Results in several conditions with the Kinect v1 and YOLOv3 [13] for object detection indicated the proposed Spiral Filtering method as the most accurate, with a a mean absolute error of 4.256cm, and a maximum absolute error of 7.6cm in relation to ground truth.

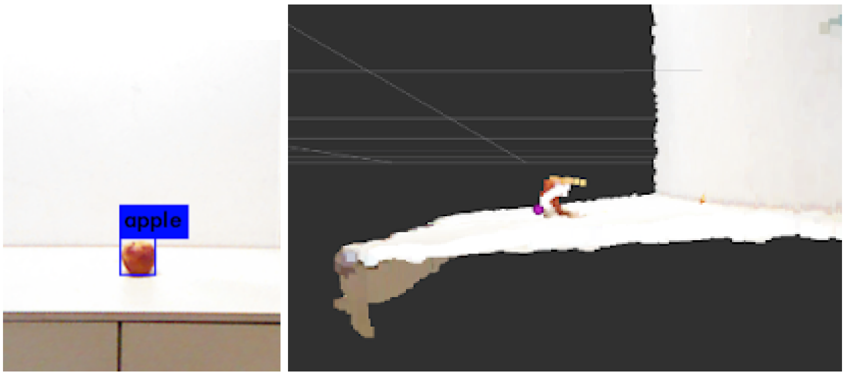

Fig. 2. Example of object 3D point estimate in RViz with the proposed Spiral Filtering algorithm (Pink dot represents the Point message)

2.4 Face Recognition Constraints in Robocup@Home

This work [7] was published in the 2023 Latin American Robotics Symposium and explores strategies to deal with the limits of facial recognition in @Home service robots. A facial recognition system designed for a service robot present in a home environment is unlikely to be required to recognize hundreds of different people in any given time. The system, however, must be able to accurately recognize the Operator and guests with minimal computational power, allowing for speed and reduced energy usage. With these constraints in mind, a KNN algorithm was used and a system was developed to allow for training with multiple datasets. When dealing with multiple users a more complete dataset can be used and when looking for a specific person or small group of people the system can increase it's accuracy and reduce the system load by using a model with less points in the multidimensional feature space. Following our task orchestration research, we tested with success the management of which datasets to perform training on and which ones should be used for a given recognition action with Behavior Trees.

2.5 Ontology

An ongoing PhD work aims to develop an ontology for the household ambient in which @Home robots are tested in order to formalize concepts and ease the communication with an unified language in a field where multiple communities are involved in and contribute to. With this in mind, this ontology aims to encompass all different areas of service robotics, including robot parts, robot actions, capabilities, behaviors and the features of the world the robot may interact with. Through communication with other Brazilian teams we are adjusting the concepts and its rules in order to achieve an universal system. Currently our task orchestration system follows the principles of the ontology and so does the TDP. In the future we plan to use this system as a guide on the development of the CBR rulebook, minimizing task and score interpretation problems.

3 Contributions

Team UTBots has been involved in the development of the @Home initiative in the Latin American scene since its first edition in 2014 and, since then, has contributed to the field of robotics with multiple research projects ranging from undergraduate to PhD students. Our contributions also include participation in the organization of the category, with a seat on the technical committee of the Brazilian league. We also hosted the Latin American Robotics Competition (LARC) in 2017 and, in 2024, both the V Brazilian Workshop on Service Robotics and the VI Brazilian Workshop on Humanoid Robotics. All of our software – and some of our hardware – is open-source and available on GitHub (`https://github.com/UtBotsAtHome-UTFPR`), with documentation on how to use it with and without our robot, HESTIA. An important contribution to the Brazilian league is our extensive video course on ROS[2], in Portuguese, which has guided the education of several researchers in robotics and new members of various teams in the league.

3.1 Theseus: Low-cost Open-Source Robotic Manipulator

The most popular solution for robotic manipulators in the Latin American scene is the use of high torque servo motors, more precisely Robotis' Dynamixel servo

[2] Our ROS Course: `https://www.youtube.com/watch?v=ZDF9_NnBu1A&list=PLaJW7rgXka310KR211BFWxLk-8qZ5HaDV`

motors. Some teams buy manipulators and others build considering the servo motors specifications, but even building them is highly expensive because of the cost of the motors and has big limitations of payload. Robotis OpenManipulator-X (RM-X52-TNM) [14] and Trossen Robotics ReactorX 200 [15] are examples of such commercial manipulators. The former has 5 DOF, supports a payload up to 500g and currently is sold for $1,416.60; the latter has 5 DOF, supports a payload of only 150g and is currently sold for $2,099.95. Specially in Brazil, this values become more expensive because of the exchange rate from USD to Brazilian Real, currently $1,000 USD = R$5,826 BRL.

Our team's proposal with Theseus was to design a manipulator capable of upholding up to 1.5kg with 5 degrees of freedom (DOF). In comparison with the cited manipulators, it is also built with aluminum extrusions and 3D printing, but, instead of using servo motors, we decided to apply outrunner brushless motors, reduced with custom cycloidal 25:1 gearboxes controlled by magnetic encoders with position control drivers based on the open-source Odrive [19] hardware. With this choice, the idea was to reduce the cost to less than $1000.00 for a bigger payload capacity.

For Theseus' first version we estimate a cost of around $700.00. Comparing the cost of $269.90 for each Dynamixel MX430 servo-motor [16] used in RM-X52-TNM and ReactorX 200, we were capable of reducing the cost of each motor with controller to around $100.00 – and estimate made with the SEQURE ODESCv3.6 [17] two-motor controller, Odrive's D5312s [18] brushless motor and the as5048a magnetic encoder. The idea is to share Theseus as an open-source project[3] in order to aid other teams with resource constraints to build a low-cost higher payload capacity manipulator.

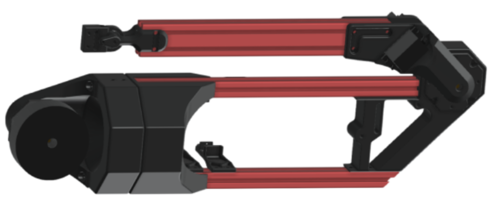

Fig. 3. Theseus open-source manipulator 3D model.

[3] Theseus Project Source: `https://github.com/UtBots-Home/theseus-project`

4 Conclusions and future work

Various research interests guide our future development, and involve current innovation for the robot's perception and cognition.

In the robot's vision domain, some improvements include adding the capability to adjust face orientation to improve results with our KNN solution for face recognition, more robust and reliable object 3D point estimation with more sophisticated voxel neighbors data structures, such as KDTrees, and pointcloud segmentation learning models. Few-shot learners are also recent advancements in machine learning that seem promising specially in adapting our object manipulation systems for novel objects for every new edition of the competition.

For our voice and human-robot interface, language models and other natural language processing techniques are promising tools to deal with social reasoning. Also, considering our approach with behaviour trees has simple syntax for structuring tasks, language models are promising solutions for tasks such as GPSR and Enhanced GPSR.

Following our recent hardware development and initiative in task orchestration with behaviour trees, we intend to validate both of these systems extensively in Stage I and II tasks, aiming for the consolidation of HESTIA's reliability and capability to deal with reactive ambients in the competition setting.

References

1. J. Cerbaro, D. Martinelli, A. S. Oliveira, and J. A. Fabro. Waiterbot: Comparison of fuzzy logic approaches for obstacle avoidance in dynamic unmapped environments using a laser scanning system (lidar). In *2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE)*, 2020.
2. A. L. Sordi Filho, L. P. Oliveira, A. S. Oliveira, J. A. Fabro, and M. A. Wehrmeister. Simultaneous navigation and mapping in an autonomous vehicle based on fuzzy logic. In Nadia Nedjah, Heitor Silvério Lopes, and Luiza de Macedo Mourelle, editors, *Studies in Computational Intelligence*, volume 664, pages 53–68. Springer International Publishing, 1 edition, 2017.
3. T. Becker, A. S. Oliveira, J. A. Fabro, and R. L. Guimaraes. Lida bridge: A ros interface to the lida (learning intelligent distribution agent) framework. In Anis Koubaa, editor, *Studies in Computational Intelligence*, volume 625, pages 703–728. Springer International Publishing, 2016.
4. D. Santos, P. Palar, A. S. Oliveira, J. A. Fabro, and T. Becker. Adding conscious aspects and simulated emotions through facial expressions in virtual robot navigation with baars-franklins cognitive architecture. In *2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE)*, João Pessoa-PB, 2018.
5. Jo ao Fabro, Marlon de Oliveira Vaz, and Andre Schneider de Oliveira. Design and development of an automated system for creation of image datasets intended to allow object identification and grasping by service robots. In Bruno José Torres Fernandes and Antônio {Pereira Júnior}, editors, *Proceedings of the 14th Brazilian Conference on Computational Intelligence*, Curitiba, PR, 2019. ABRICOM.

6. G. F. Armênio, J. A. Fabro, R. R. Tognella, F. P. Conter, M. V. Oliveira, and E. S. Silva. Estimating the 3d center of an object with kinect sensor rgb-d images. In *Latin American Robotics Symposium (LARS), XX Simpósio Latino Americano de Robótica (LARS 2023)*, Salvador, Bahia, Brazil, 2023.
7. D. Segalle, G. F. Armênio, G. F Lewin, E. S. Silva, and J. A. Fabro. Identifying new persons in the context of the robocup@home competition using knn. In *Latin American Robotics Symposium (LARS 2023)*, Salvador, Bahia, Brazil, 2023.
8. Nicolas Freitas Dias, Jardel dos Santos Dyonisio, Luis Milczarek Quadros, Gabriel Amaral Dorneles, Marina Zanotta Rocha, Murilo Castanheira Bicho, João Francisco de Souza Santos Lemos, Igor Pardo Maurell, Jardel Prado, Henrique Bezerra, Cassio Teixeira, Rodrigo da Silva Guerra, and Paulo Lilles Jorge Drews Junior. Shark: Stable hoverboard-driven autonomous robot kit. In *BRAHUR/BRASERO 2024 - Brazilian Workshops on Humanoid and Service Robotics 2024, Curitiba/PR, Brazil, April 25-27*, 2024.
9. Jean-Marc Valin. A hybrid dsp/deep learning approach to real-time full-band speech enhancement, 2018.
10. Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. Robust speech recognition via large-scale weak supervision, 2022.
11. Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. Llama: Open and efficient foundation language models, 2023.
12. Open source natural language processing (nlp). `https://rasa.community/open-source-nlu-nlp/`. Accessed Jul., 2023.
13. J. Redmon and A. Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.
14. ROBOTIS. Openmanipulator-x rm-x52-tnm. https://www.robotis.us/openmanipulator-x-rm-x52-tnm/. Accessed: 2024-11-21.
15. Trossen Robotics. Reactorx 200 robotic arm. https://www.trossenrobotics.com/reactorx-200. Accessed: 2024-11-21.
16. ROBOTIS. Dynamixel xm430-w350-t. https://www.robotis.us/dynamixel-xm430-w350-t/. Accessed: 2024-11-21.
17. Sequre Mall. Odesc dual drive. https://sequremall.com/collections/odesc/products/odesc-dual-drive. Accessed: 2024-11-21.
18. ODrive Robotics. Dual shaft motor d5212s 300kv. https://shop.odriverobotics.com/products/dual-shaft-motor-d5212s-300kv. Accessed: 2024-11-21.

HESTIA Hardware Description

UTBots' Household Efficient Service Task-Integrated Assistant (HESTIA) specifications are as follows:

- Base: closed environment terrain base (differential pair) with Field Orient Controlled 2 brushless motors, 2.5m/s max speed.
- Torso: aluminum extrustions, acrilic sheets and plastic sheets for structure, component fixations and external protection.
- Manipulator: Mounted on torso. 5DOF, built of aluminum, 3D printing, 5 outrunner brushless motors, cycloidal 25:1 gearboxes, magnetic encoders and Field Oriented Control drivers. Maximum load: 1,5kg.
- Neck: fixed aluminum extrusion.
- Robot dimensions: height: 1.3m, width: 0.7m depth 0.8m
- Robot weight: 20kg.

Our robot incorporates the following sensoring devices:

- LIDAR: 1 Slamtec RPLidar A1 and 1 SICK S300.
- RGB-D: 1 Intel Realsense d435i.
- RBG Camera: 1 Logitech C920.
- Magnetic Encoders: 5 Magnetic Encoders as5048a.
- Microphone: 1 Boya VM01.

Fig. 4. UTBots' Robot HESTIA

HESTIA Software Description

For our robot we are using the following software. Our wrappers and packages are all located in out Github [4]:

- Platform: Linux Ubuntu 20.04 and Linux Ubuntu 22.04.
- Navigation: nav2 amcl
- Mapping: slam_toolbox
- Face recognition: our package built with actions interface and a fast training algorithm, using Python's face_recognition and dlib.
- Audio filtering: our wrapper for Silero Voice Audio Detection with RNNoise.
- Speech recognition: our wrapper for whisper.cpp.
- Speech generation: our wrapper for Mimic3 TTS and Coqui TTS.
- Natural Language Understanding: our wrapper for RASA.
- Natural Language Reasoning: our wrapper for llama.cpp, for human-robot interactions that require previous context.
- Object recognition: our wrapper for YOLOv11.
- Pose Recognition and People Tracking: our Mediapipe Pose, with our package mediapipe_track.
- Manipulator control: our custom control package with moveit.

[4] UTBots Team's Github: `https://github.com/UtBotsAtHome-UTFPR`

External Devices

- Intel Nuc with Linux Ubuntu 20.04
- NVIDIA Jetson Nano with Linux Ubuntu 20.04
- Dell Inspiron 15 Notebook with Linux Ubuntu 22.04
- Raspberry Pi 3B+
- 2 SEQURE ODESC v3.6 for Field Orient Control of the robotic arm