Towards a new Semantic Social Navigation Paradigm for Autonomous Robots using CORTEX

Pedro Núñez[1], Luis J. Manso[1], Pablo Bustos[1], Paulo Drews-Jr[2], Douglas G. Macharet[3]

***Abstract*— Navigation is one of the most basic and widely studied problems in the field of autonomous robots. The scientific community assumes the navigation task as a combination of three fundamental robotics skills, which can be summarized in i) self-localization; ii) path-planning; and iii) map building. Traditionally, these problems have been addressed using a geometric world model, that is, a 2D or 3D map representation of the environment. This tendency is now changing, and the scientific community is experiencing an increasing interest in so-called semantic solutions, which integrate semantic knowledge and geometrical information. In addition, new generation of robots should be able to work also taking into account social conventions, which is commonly named social navigation. This paper describes the ongoing work of a new proposal for a navigation paradigm where the semantic knowledge of the robot's surroundings and different social rules are used in conjunction with the geometric representation of the environment. The proposal uses CORTEX, an agent-based Robotics Cognitive Architecture which provides a set of different agents in the deliberative-reactive spectrum. This paper introduces three cases of use that will be tested in two different social robots within the NAVLOC project[1].**

I. Introduction

In the not too distant future, social robotics will be helpful in everyday life. Social robots will perform typical human tasks in offices, hospitals, homes or museums. In these complex and dynamic scenarios people and objects usually move around the robot, complicating robot capabilities for navigating. Thus, in the social context where these robots are going to work, there exist different skills that are expected, such as human or object avoiding collisions, localization, path planning or map building. These robot's skills typically have been addressed in the literature using representations of the spatial structure of the environment, however this tendency is now changing, and the scientific community is experiencing an increasing interest in so-called semantic solutions, which integrate semantic knowledge and geometrical information.

In the last years, the term *social navigation* in robotics, which is expected to become an increasingly important task in next social robots generation [1], has been introduced as a way to relate the robot navigation in human scenarios and human-robot interaction. New generations of social robots should be able to generate different socially accepted routes during an interaction with humans and also exhibit proactive social behaviors during the navigation [2] (*e.g.*, to gracefully approach people, or to wittily enter and exit from a conversation).

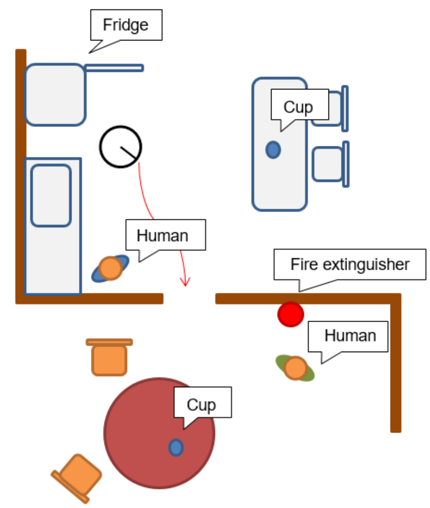


Fig. 1. Brief description of the semantic social navigation problem. The robot has to choose the best route and navigate from the kitchen to the living-room by using its semantic knowledge (*e.g.*, cups, fire extinguisher and fridge) and social rules (*e.g.*, humans along the path)

Fig. 1 illustrates the problem to solve: the robot located in the kitchen has to choose the best route and navigate from its pose to the living-room along a complex dynamic environment with people. The semantic social navigation approach described in this paper firstly introduces a high-level and long-life knowledge captured by the robot from the environment (cups, fridge or humans in the figure), in a similar way that the human point-of-view, and after, it introduces socially accepted rules to the semantic knowledge during the planning and navigation tasks.

The navigation paradigm described in this paper uses the cognitive architecture CORTEX [3]. CORTEX is based on a set of agents (*i.e.*, semi-autonomous functional units that collaborate by means of a common representation in their pursue of a common goal) that can run anywhere in the deliberative-reactive spectrum. In this cognitive architecture there are navigation, perceptual and human-robot interaction

*This work has been partially supported by the MICINN Project TIN2015-65686-C5-5-R, by the Extremaduran Goverment project GR15120, by MEC project PHBP14/00083 and by CAPES-DGPU 7523/14-9.

[1] Pedro Núñez, Luis J. Manso and Pablo Bustos are with University of Extremadura, Escuela Politécnica, Cáceres, Spain. pnuntru@unex.es

[2] Paulo Drews-Jr is with Federal University of Rio Grande, Rio Grande, Brasil.

[3] Douglas G. Macharet is with Federal University of Minas Gerais, Belo Horizonte, Brasil.

[1]Brazil-Spain cooperation on navigation and localization for autonomous robots on underwater and terrestrial environments

agents, among others, thereby facilitating the combined use of semantic knowledge and social rules. In the new proposal of navigation, perceptual agents are used to acquire information from the environment and to detect objects from it (semantic knowledge), human-robot interaction agents are used to infer or apply social rules, and navigation agents are used to provide skills to navigate in a secure way.

This article is structured as follows: Section II provides a brief summary about similar works in this field of research. In Section III, a description of CORTEX cognitive architecture and the hybrid representation are made. Section IV provides the proposal description along the main agents involved. The description of the semantic social navigation paradigm is described in Section V. Finally, the main conclusions are detailed in Section VII.

II. Related Work

Classical navigation algorithms use spatial representation of the robot's surrounding, that is, the path-planning or the localization problems require a geometric map of the environment. Recently, several advances in semantic navigation have been achieved. In fact, social robots that incorporate skills for task planning and storing semantic knowledge in their maps are commonly used (*e.g.*, classification of environments, indoor or outdoor, spaces, such as rooms, corridors or garden, and labels of places and/or objects) [4], [5]. By using this semantic knowledge, robots are able to navigate or planning other tasks. Some works autonomously acquire this semantic information by analyzing data from robot's sensors, or by using voice interactions with robots (see Kostavelis's survey [4]).

Social navigation started being extensively studied in the last years and several methods have been proposed from then. Most of the solution are based on using a classic navigation algorithm, and therefore adding social conventions and/or social constraints. According to this paradigm, some authors have proposed models of social rules by using cost functions. In [6], for instance, the authors use a classical A* path planner in conjunction with social conventions, such as to pass humans on the right. In the Traberg et al.'s work [7], they use potential fields and a proxemics model[2]. Other solutions for social navigation use the detection of human intentions in order to model the social navigation. In [9], authors propose the Modified Social Force Model (MSFM), basically a local navigation method where the path is able to be modified after analyzing the human intention.

Different proposals can be found in the literature about cognitive architectures and the kind information to use, most of them by separately storing symbolic and metric information. Symbolic knowledge representation, such as the works proposed in [10] and [11], has been at the core of Artificial Intelligence since its beginnings. Most recently, solutions that integrate spatial and symbolic knowledge in a unified representation are studied in the literature (*deep representation* [12]). Examples of these deep representations are the works in [12], [13], [3].

[2]Proxemics is defined as the study of humankinds perception and use of space [8]

III. Deep State Representation and Cortex

The concept of *deep representations* was initially described by Beetz et al. [12] and it advocates the integrated representation of robots knowledge at various levels of abstraction in a unique, articulated structure such as a graph. Based on this concept, a new shared representation, *Deep State Representation (DSR)*, to hold the robots belief as a combination of symbolic and geometric information, is proposed in [3]. This new structure represents knowledge about the robot itself and the world around it, in a flexible and scalable way. More formally, in [3] DSR is defined as *a directed multi-labelled graph where nodes represent symbolic or geometric entities and edges represent symbolic and geometric relationships*.

The robotics cognitive architecture CORTEX is defined structurally as a configuration of software agents connected through DSR. An agent within CORTEX is defined as *a computational entity in charge of a well defined functionality, whether it be reactive, deliberative of hybrid, that interacts with other agents inside a well-defined framework, to enact a larger system*. The CORTEX architecture, which has been compared with other similar architectures in [3], is implemented on the top of the component-oriented robotics framework RoboComp [14]. In CORTEX, higher-level agents define the classic functionalities of cognitive robotics architectures, such as navigation, manipulation, person perception, object perception, dialoguing, reasoning, planning, symbolic learning or executing. These agents operate in a goal-oriented regime [10] and their goals can come from outside through the agent interface, and can also be part of the agent normal operation.

In Fig. 2 an overview of the DSR and its location within the cognitive architecture CORTEX is drawn. Different agents, such as navigation, person detector or planner are also shown. DSR is illustrated as a a graph where all the robot knowledge about its surrounding is represented. The next sections describe the design of the semantic social navigation system using the cognitive architecture CORTEX, analyzing the agents involved and the relationships between them in three different cases of study.

IV. Agents

In the proposal of a new social and semantic navigation paradigm for robots, different specific agents within CORTEX are involved. First, in the higher layer of the architecture the robot must have the capability of detecting objects in the path and updating the symbolic model accordingly. Additionally, the skill of detecting humans is also mandatory because robots need to know about humans to get commands, avoid collisions and provide feedback. The final, and most important agent for social navigation, is the one implementing the navigation algorithms that allows robots to navigate from a point to another in a secure and social manner (implementation of the path-planning,

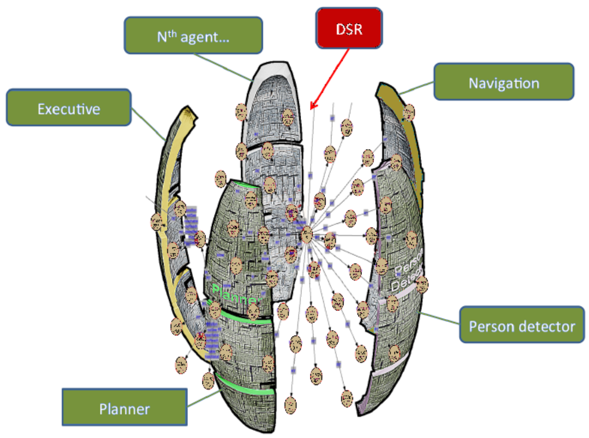


Fig. 2. An overview of the DSR and its location within the cognitive architecture CORTEX [3]

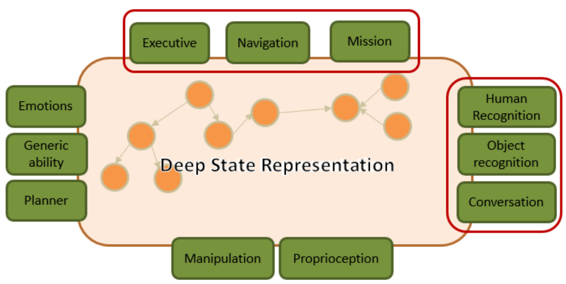


Fig. 3. Main agents within CORTEX involved in the semantic social path planning described in this proposal are highlighted in red.

localization and SLAM algorithms, among other). In the next subsections, a brief description of the main agents involved in the proposal is provided. These agents are highlighted in Fig. 3, which illustrates the current CORTEX cognitive architecture [3].

A. Object detection and representation

The object perception agent is in charge of recognizing and estimating the pose of objects and visual marks in the environment. For each object or mark detected it describes within the model (DSR) not only the pose but also its type.

These kind of elements are useful for the robot in several social scenarios. For instance, humans will seldom ask the robot to go to a coordinate because they do not necessarily need to know the reference frame used by the robots and, more importantly, because it is not comfortable for humans to provide targets in such a way.

Synthetic visual marks are detected using the AprilTags library [15]. Arbitrary visual marks will be detected using the OpenCV library [16] and 3D objects are currently being detected using an object recognition pipeline based on the PointClouds library [17]. The poses of the detected objects are referenced to known objects (in the DSR) that support them, such as a known table under a target cup. Once referenced, the kinematic relations embedded in DSR allow the computation of any object's pose from any reference frame easily.

B. Person detector

Person detector is the agent responsible for detecting and tracking the people in front of the robot. Humans do not usually enjoy its personal space being invaded by robots. The presence of humans in the robots' path or in their environment may determine changes in the navigation route in order to make it socially acceptable.

The person detector agent acquires the information using an RGBD sensor. For each detected person the agent inserts in the DSR the pose of its torso, its upper limbs, and the head. The lower limbs are ignored because they do not provide as much social information as the head, the upper limbs and the torso do [3]. These elements can be used to infer the objects referenced by the humans when they point or look at them. The torso is used to avoid entering the personal space of humans and as an indicator of the possible directions in which they might walk.

C. Conversation

The conversation agent performs human-robot interaction (HRI). In social environments, HRI provides tools to the robot and/or human to communicate and collaborate. Therefore, this agent is used to include information in the model when humans tell robots about unknown objects and to properly acquire commands. Automatic Speech Recognition and Text-to-Speech algorithms allow robot to both send and receive information to/from humans in the environment during its social navigation.

D. Mission

This agent is used as a means to provide missions to the executive agent and to visualize the DSR. It has two graphic views. A graph-like view and a 3D geometric view [3].

E. Executive

The Executive is responsible of planning feasible plans to achieve the current mission, managing the changes made to the DSR by the agents as a result of their interaction with the world, and monitoring the execution of the plan. The active agents collaborate executing the actions in the plan steps as long as they consider them valid (it must be taken into account that agents might have a reactive part). Each time a structural change is included in the model, the Executive uses the domain knowledge, the current model, the target and the previous plan to update the current plan accordingly. The Executive might use different planners. Currently AGGL [18] and PDDL-based [19] planners are supported.

F. Navigation

Navigation is in charge of performing local navigation complying with social rules and including the location of the robot in the DSR. Global path planning is performed by the symbolic planner used by the executive.

Two poses are maintained by the robot: the pose obtained from the local odometry, and the pose provided by a localization algorithm based on external geometric laser features. Given their properties, each of these poses is useful for a particular purpose. Odometry provides good information relative to the robot's position in the short term, while localization provides good information for mid and long term positioning. Additionally, the space walked by the robot in the last seconds is also included.

Regarding localization algorithms, the navigation agent is algorithm-independent. It has been used with different algorithms showing different properties, which can be selected to fit different kinds of environments.

While it can be used with any local navigation system, the navigation agent has been only tested with the path-planning algorithm proposed in [20], an algorithm based on the elastic-band representation, with successful results. The navigation paradigm presented in this paper extends the geometrical path-planning to a social semantic algorithm, which is described in the next section.

V. Social Semantic Navigation in CORTEX Cognitive Architecture

In this section the semantic social navigation paradigm is described. An overview of the system is shown in Fig. 4. On the top of the architecture is the global semantic path planner, followed by a local geometrical path planner. Both of them are affected by the social navigation model. The semantic path planner chooses the optimal route, that consists of a list of waypoints. These waypoints along the path are characterized by a set of labeled objects in the map that the robot should perceive. Then, the robot plans a local geometrical navigation from its current pose to the next waypoint in the list, and looks for the objects in the path. Finally, this path is affected by the social navigation model, and if necessary, the local (or global) route is re-planned. All the agents within CORTEX described in the previous section are concurrently running in the navigation process.

A. Semantic Path Planning

Global path planning at a symbolic level is performed by the planner included in the executive. The semantic path planner is based on the use of a two-hierarchies architecture, similar to that one presented in [5]. Both, the spatial and semantic properties of each object within DSR allow the planner to choose the best global route. Let $O = \{o_1, o_2, ..., o_n\}$ being the list of n objects o_i within the semantic map of the robot, that is, its high level and long-term knowledge. Each object o_i is characterized as $o_i = \{m_i, s_i, l_i\}$, where m_i is the metric representation of the object (*i.e.*, rotation and translation matrices from its parent node), s_i is the semantic information associated to the object (*i.e.*, label) and l_i is the mesh (*i.e.*, 3D point cloud). Each object has a parent node, which usually represents the room where the object is located. Rooms are also nodes of the graph, that are connected if they are sharing a door. Thus, the semantic path planning algorithm chooses the best route from the graph,

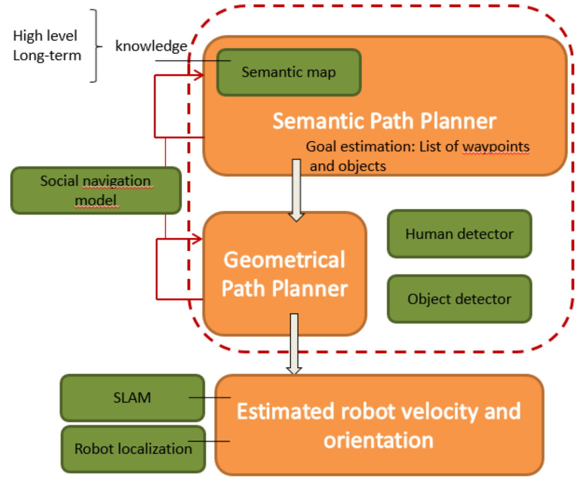


Fig. 4. The overall blocking diagram of the proposed system.

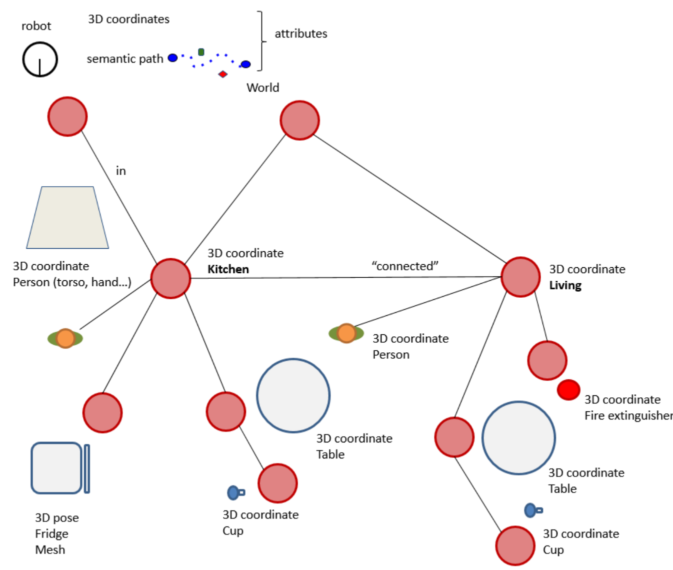


Fig. 5. Brief description of the DSR for the study cases described in this paper.

that is, the list of rooms that the robot has to visit. Therefore, the planner generates a list of waypoints and includes the path within robot's attributes in DSR (i.e., all the agents in the architecture are able to visualize it). This path is characterized by a set of n waypoints, $\Omega_R = \{\rho_1, \rho_2, ...\rho_n\}$ and j ordered objects, $\Gamma_R = \{o'_1, o'_2, ...o'_j\}$, being $o'_j \in O$, that the robot should perceive during the navigation. Thus, the global navigation is achieved between consecutive rooms according to the path Ω, where the robot has to detect the objects from Γ_R. In Fig. 5, an example of DSR is shown, where both semantic and geometrical information of objects, humans, rooms and the robot are also illustrated.

B. Geometrical path Planning

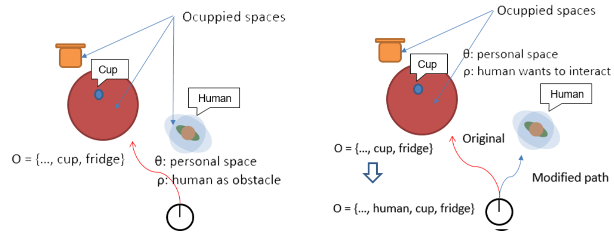
Once the robot is assigned the path and Γ_R, the geometrical path-planner should accomplish the navigation between two consecutive waypoints and look for two consecutive objects, o'_{k-1} and o'_k. The geometrical path-planning algorithm of the proposal is based on the use of graphs as a representation of free space and of elastic-bands [20] as an adaptable representation of the current path. Elastic bands work as a glue filling the gap between the internal representation of the path and the constraints imposed by the world physics. In order to build a graph representing the overall free space, the probabilistic road map algorithm, PRM, is used [21] along with a preexisting map and a collision detection algorithm. To complete this schema, the RRT algorithm [22] is also included in the system to complete the paths when unconnected islands remain in the PRM graph or to connect the robot's current position and robot's final position with nodes in the graph. The object perception agent is directly involved in the path following process: when the robot detects the object o'_k (or a list of objects) at the waypoint ρ_l, a new target is generated ρ_{l+1}, being the new local route defined by the nodes ρ_l and ρ_{l+1}.

C. Social Navigation Model

In order to mathematically formulate the socially-acceptable navigation algorithm, let denote $H = \{H_1, H_2, ...H_n\}$ the set composed by n humans in the environment. Each human, H_i, in the DSR is represented by the pose of its torso, its upper limbs, and the head (see Fig. 5). These elements can be used for defining a personal space θ_i and a social interaction intention ρ_i. Both θ_i and ρ_i are detected by the human detector agent, and are included in the DSR as information associated to human H_i. On one hand, and similar to the work presented in [1], θ_i is defined as Gaussian Mixture Model of two 3D Gaussian functions, one for the front of the individual, and other for its rear part. By adjusting the covariance matrices of these two gaussians, one can modify the personal space model. On the other hand, ρ_i describes the different cases where a human wants or not to interact with the robot during the path: i) human does not want to interact (*i.e.*, human is considered as obstacle); ii) human wants to interact with the robot, and then, the robot has to approach human, interact and finish the communication. In this respect, depending of the ρ_i value, the final path may be modified. For instance, if the human is considered as obstacle, the graph in the geometrical local navigator has to be updated in order to avoid this new obstacle (see Fig. 6(a)). On the contrary, if the human wants to interact with the robot, a new object $o'_k = H_i$ is included in the list of nodes to reach, being H_i the next target (see Fig. 6(b)).

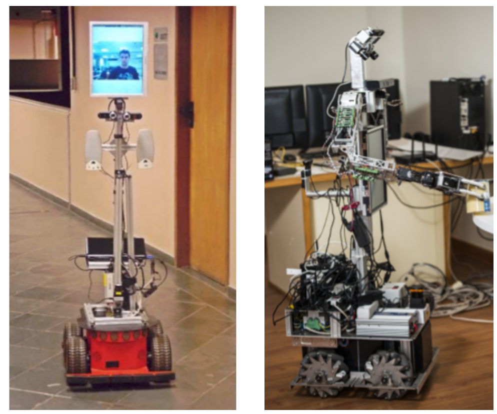
VI. Cases of Study

Within the 'BS-NAVLOC' project, this paper proposes three different cases of study. All of them are examples of robots navigating in indoor environments, and the main goal is to demonstrate that the semantic social path planning



(a) (b)

Fig. 6. Social navigation model proposed in this paper: a) human does not want to interact with the robot; and b) the human wants to interact.



(a) (b)

Fig. 7. The semantic social navigation within CORTEX is going to be integrated in two different robots: a) DOKBot robot, from VerLab research group at the University Federal of Minas Gerais; b) Shelly robot, from RoboLab research group at the University of Extremadura.

algorithm proposed in this paper, using the CORTEX cognitive architecture, may be performed in different robotics platforms in a near future and with successful results. In this section, the ongoing work is presented, describing briefly the DSR and the relationships between the involved agents.

The proposal of semantic social navigation paradigm is going to be tested in two different robots (Fig. 7). The first robot is the DOKBot, from VerLab at the University Federal of Minas Gerais, which consists of a Pioneer 2-AT robotics platform equipped with different perceptual sensors, such as laser, RGB cameras and microphones (see Fig. 7(a)). This robot was originally designed for semiautomatic telepresence purposes. The second autonomous system is Shelly, an anthropometric social robot that is currently being used in RoboLab, at the University of Extremadura. This robot was designed to help in daily life tasks. It is composed of an omnidirectional base, two 7-DOF arms with two-fingered grippers and a RGB-D camera attached to a pan-tilt-yaw structure. It has another RGB-D camera on the upper part of the torso which is used to detect human bodies and a lidar for navigation. This robot is illustrated in Fig. 7(b).

Next, the experimental scenarios are described. They have been designed from low-complexity to high-complexity levels:

- Semantic Navigation using CORTEX. In this experi-

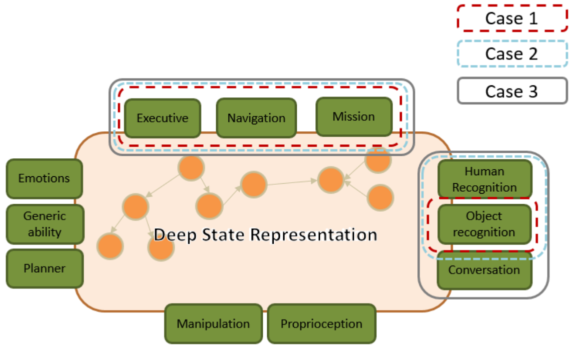


Fig. 8. Agents involved in the cases of study described in this paper. a) semantic navigation; b) semantic social navigation; and c) semantic social navigation with HRI

ment, the robot chooses the best route from a room to another. In this scenario, there is not people in the path, and thus, only the semantic knowledge of the rooms is used. Fig. 8 shows the agents involved in this scenario (red colour in the figure).

- Semantic social Navigation. This experimental scenario consists on a semantic social navigation. Similar to the previous case of study, the robot has to navigate between two different rooms in an indoor and human environment. In this respect, people walk or stand in the robot path, and thus, the robot has to modify the route in order to be socially accepted. The set of agents involved in this case of study is highlighted in blue colour in Fig. 8.
- Semantic social Navigation with HRI In this case of study, the robot first interacts with the human in order to know what is the next room to visit, and also, other humans interact with the autonomous agent during the path. In this HRI, the robot may modify partial or fully its route. Finally, in Fig. 8, the agents involved in CORTEX are highlighted in gray colour.

VII. Conclusions and Future Works

This paper presents the ongoing work, within the NAVLOC project, of a proposal for the design of a semantic social navigation system. The approach is based on the use of a global semantic path-planner in conjunction with a social navigation model. The theoretical proposal achieves the main goal of this kind of algorithm, that is, the robot is able to choose the best route from its current position to another position in a dynamic and complex scenario by using its high level knowledge and by applying social rules in order to be socially accepted. High functionality and robustness are guaranteed by using the cognitive architecture CORTEX and the Deep State Representation.

As it was aforementioned, this paper describes the ongoing work, where three different experimental scenarios are also described in order to test the proposed social navigation algorithm in future works. Currently, both spanish and brasilian researching teams, are working in integrating CORTEX in the two robots presented in this paper, Shelly and DOKbot.

Acknowledgments

This work has been partially supported by the MICINN Project TIN2015-65686-C5-5-R, by the Extremaduran Goverment project GR15120, by MEC project PHBP14/00083 and by CAPES-DGPU 7523/14-9.

References

[1] Gómez, J., Mavridis N. and Garrido, S. Social Path Planning: Generic Human-Robot Interaction Framework for Robotic Navigation Tasks, in Workshop on Cognitive Robotics Systems: Replicating Human Actions and Activities at IROS 2013.
[2] Lichtenthaler, C., Peters A., Griffiths, S. and Kirsch, A. Social Navigation - Identifying Robot Navigation Patterns in a Path Crossing Scenario, in Lecture Notes in Computer Science, Volume 8239, pp 84-93
[3] Luis Vicente Calderita, Deep State Representation: a Unified Internal Representation for the Robotics Cognitive Architecture CORTEX, PhD Thesis, University of Extremadura, 2015.
[4] Kostavelis, I. and Gasteratos, A. Semantic mapping for mobile robotics tasks: A survey in Robotics and Autonomous Systems 66 (2015) pp. 86-103, 2014.
[5] Galido, C., Fernández-Madrigal, J.A., González J., and Saffiotti, A. Robot Task Planning using Semantic Maps in Robotics and Autonomous Systems, 56(11), pp. 955-966, 2008.
[6] Kirby, R., Simmons, R., Forlizzi, J.: COMPANION: A Constraint-Optimizing Method for Person-Acceptable Navigation, in IEEE International Symposium on Robot and Human Interactive Communication, 2009
[7] Tranberg Hansen, S., Svenstrup, M., Andersen, H.J., Bak, T.: Adaptive Human Aware Navigation Based on Motion Pattern Analysis, in IEEE International Symposium on Robot and Human Interactive Communication, 2009
[8] Hall, E. "Proxemics", Current Anthropology, vol. 9, no. 2-3, pp. 83108, 1968.
[9] Photchara R., Mae, Y, Ohara K., Kojima M., and Arai T. Social navigation model based on human intention analysis using face orientation, in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1682-1688, 201.
[10] Rusell, S., Norvig, P.: Articial Intelligence: A Modern Approach (3rd Edition). Pearson (2009)
[11] Poole, D., Mackworth, A.: Articial Intelligence: Foundations of Computational Agents. Cambridge University Press (2010), http://artint.info/index.html
[12] Michael Beetz, Dominik Jain, Lorenz Mösenlechner and Moritz Tenorth. "Towards performing everyday manipulation activities". In Journal of Robotics and Autonomous Systems, vol. 58, n. 9, pp. 1085–1095. Elsevier. 2010.
[13] Ali, M.: Contribution to decisional human-robot interaction: towards collaborative robot companions, PhD Thesis, Institut National de Sciences Appliquees de Toulouse, France (2012).
[14] L. Manso, P. Bachiller, P. Bustos, P. Núnez, R. Cintas and L.Calderita. RoboComp: a Tool based Robotics Framework, In Proceedings, SIMPAR Second International Conference on Simulation, Modeling and Programming for Autonomous Robots. pp 251-262. 2010.
[15] Olson, E. AprilTag: A robust and flexible visual fiducial system. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pages 3400-3407.
[16] http://opencv.org/
[17] http://pointclouds.org/
[18] L.J. Manso, P. Bustos, P. Bachiller and P. Núñez. "A Perception-aware Architecture for Autonomous Robots" In International Journal of Advanced Robotic Systems (ISSN 1729-8806), InTech, Vol. 12, No. 174, 2015. DOI: 10.5772/61742.
[19] McDermott, Drew and Ghallab, Malik and Howe, Adele and Knoblock, Craig and Ram, Ashwin and Veloso, Manuela and Weld, Daniel and Wilkins, David. "PDDL-the planning domain definition language". 1998.
[20] Haut, M., Manso, L.J., Gallego, D., Paoletti, M., Bustos, P., Bandera, A. and Romero-Garces, A. A Navigation Agent for Mobile Manipulators, in Robot 2015: Second Iberian Robotics Conference Volume 418 of the series Advances in Intelligent Systems and Computing pp 745-756.

[21] Kavraki, L.E., Svestka, P., Latombe, J.C., Overmars, M.: Probabilistic roadmaps for path planning in high-dimensional configuration spaces. IEEE Transactions on Robotics and Automation 12(4), 566580 (1996).

[22] LaValle, S.: Planning Algorithms. Cambridge University Press, Cambridge (2006).